Recognition of Foot Gestures through Time-Frequency Analysis and Deep Learning Techniques Using sEMG Data

Oğulcan Çelik, İbrahim Kaya, Özlem Karabiber Cura*
İzmir Kâtip Çelebi University, Department of Biomedical Engineering, Çiğli, 36520, İzmir, Türkiye
**Corresponding author: Y230238001@ogr.ikcu.edu.tr / ogulcancelik0011@gmail.com*

Abstract

Most existing approaches for recognizing foot gestures rely on manually extracted features from sEMG recordings. This study aims to develop a deep learning-based method that automatically utilizes the time-frequency representation of sEMG signals from a single-channel sEMG, eliminating the need for manual feature extraction. To achieve this, the continuous wavelet transform (CWT) is used to generate time-frequency RGB images from sEMG signals. Transfer learning is then applied using convolutional neural networks, leveraging pre-recorded foot gesture signal data from healthy individuals. These CWT images are used to classify different foot gestures. The proposed method is evaluated on a dataset of foot gesture sEMG signals, and the results demonstrate that even with single-channel sEMG data, this approach can achieve near state-of-the-art accuracy.

Keywords: CWT, STFT, foot gestures, deep learning, sEMG

Introduction

Electromyography (EMG) is a widely used technique for assessing and recording the electrical activity of skeletal muscles. By detecting electrical signals generated during muscle contractions, EMG provides valuable insights into muscle function and neuromuscular health. Surface electromyography (sEMG), a non-invasive method, captures these electrical signals from the skin's surface, making it highly applicable in both clinical and research settings [1]. Recent advancements have sparked interest in using sEMG for classifying various motor activities, particularly foot gestures [2]. These gestures, including dorsiflexion, plantarflexion, inversion, and eversion, play a crucial role in maintaining balance and mobility and performing daily activities. Accurate recognition and classification of foot gestures have significant implications for assistive device development, rehabilitation strategies, and human-computer interaction systems.

Traditional approaches for classifying foot gestures with sEMG signals often rely on manually extracted features, which can be time-consuming and prone to variability. However, the advent of deep learning has led to the automation of this process. Convolutional neural networks (CNNs) have shown remarkable success across various fields by automatically learning and extracting essential features from raw data [2]. Robotic technologies are increasingly being used in diverse applications, including prosthetic limb control and the remote operation of robotic arms and legs. EMG signals have emerged as a powerful tool for both controlling and diagnosing fundamental movements [3]. These signals, originating from biopotential activity generated by muscle contractions, result from complex electrochemical and electromechanical interactions within the body. When action potentials generated in the brain travel through nerves to reach muscles, spontaneous movements occur. Electrodes placed on the skin capture these EMG signals, which can then be analyzed in either the time or frequency domain. In recent years, there has been growing interest in using sEMG for classifying motor activities, particularly foot gestures. The ability to accurately classify these movements is crucial for developing advanced assistive technologies. Understanding EMG signals and foot movement dynamics requires acknowledging the complexity of muscle-generated signals and employing advanced techniques for their detection, decomposition, processing, and classification [3]. Analyzing biomedical signals like EMG is a challenging task that necessitates sophisticated signal processing techniques. Continuous Wavelet Transform (CWT) and Short-Time Fourier Transform (STFT) are widely used in EMG analysis, offering fast and precise methods for movement classification. Wavelet transform analysis facilitates feature extraction from EMG signals, converting complex data into useful control signals for prosthetic devices [4]. Applications of CWT and STFT in foot gesture analysis include developing multi-channel electromyogram acquisition systems and integrating deep learning (DL) methods for

gesture recognition based on sEMG signals. These techniques enhance recognition accuracy by extracting informative features from sEMG data.

Recent studies have explored machine learning-based approaches, such as SENIAM, to improve EMG signal onset/offset detection, thereby enhancing prosthetic device performance [5]. Despite the challenges of assessing surface EMG signals due to the complexity and overlap of muscle activity, CWT and STFT have proven valuable in foot gesture research. Investigating foot gestures is essential for understanding the neuromuscular system, with sEMG serving as a critical tool for this purpose. The combination of sEMG data with CWT and STFT analysis provides a robust framework for analyzing foot movement patterns. The significance of foot gesture analysis extends beyond assistive technologies to diagnosing musculoskeletal conditions, such as diabetic neuropathy and knee muscle exertion estimation [6]. Signal processing techniques like CWT and STFT facilitate the evaluation of temporal and spectral variations in muscle contractions, offering valuable insights for medical applications [7]. The potential applications of foot gesture recognition are extensive, promising advancements in diagnostics, rehabilitation, and neuromuscular research. Machine learning plays a crucial role in enhancing foot gesture analysis by extracting key features from sEMG signals through wavelet-based methods [8]. Emerging technologies like SENIAM demonstrate the effectiveness of machine learning in detecting muscle activation timing [9]. Additionally, real-time hand gesture recognition models leveraging deep learning and EMG analysis showcase the transformative potential of artificial intelligence in this field. By integrating machine learning with EMG signal processing, researchers can significantly improve the analysis of CWT and STFT data, ultimately enhancing the understanding and diagnosis of musculoskeletal and neuromuscular conditions [10].

This study explores the application of deep learning to analyze electromyography (EMG) data from healthy individuals, specifically focusing on foot gesture classification using single-channel surface EMG (sEMG) signals. We propose a novel method that transforms sEMG signals into RGB images using the Continuous Wavelet Transform (CWT) to capture both temporal and spectral characteristics. To classify these images into specific foot gestures, we leverage transfer learning with several pre-trained CNNs, including InceptionV3, ResNet18, ResNet101, InceptionResNetV2, and ResNet50. We also explore the use of Short-Time Fourier Transform (STFT) for image generation. Our method is evaluated using a comprehensive dataset of foot gestures from 12 healthy participants. The extensive signal processing and model optimization within our deep neural network framework demonstrate the meticulous analysis required for reliable results. Our findings indicate that this approach achieves near state-of-the-art accuracy, highlighting deep learning's significant potential in advancing bioelectric signal analysis and foot gesture recognition.

Methodology

A dataset consisting of EMG (Electromyography) data for five different foot gestures was generated by collecting data from 12 healthy individuals. The collected data was organized into cell-shaped datasets. Subsequently, signal processing and analysis were performed on these datasets using Continuous Wavelet Transform (CWT) with specific parameters, utilizing Morse, Amor, Bump, Morlet, and Paul wavelet methods. Additionally, the Short-Time Fourier Transform (STFT) was employed to further analyze the EMG signals. These analyses aimed to examine and evaluate the characteristics of the signals and data points derived from the EMG recordings associated with the foot gestures. This comprehensive approach allowed for a detailed investigation into the muscular activity patterns associated with the specified foot gestures of healthy subjects. In the conducted analyses, both the CWT and STFT techniques were applied to the datasets with predefined parameters. This involved scrutinizing the time-frequency representation of the EMG signals and identifying relevant features and patterns [11]. The CWT analyses facilitated the extraction of essential information regarding the temporal and frequency characteristics of the signals and provided insights into the nuances of muscle activity during the specified foot gestures. Similarly, the STFT analyses provided complementary insights by offering a different perspective on the time-frequency domain, allowing for the identification of transient features and changes in frequency content over time [12]. Together, these analyses contribute to a better understanding of the muscular dynamics associated with the selected foot gestures and offer valuable insights for potential applications in areas such as biomechanics and movement science. The resulting CWT spectrograms for each wavelet method (Morse, Amor, Bump, Morlet, and Paul) and the STFT spectrograms were categorized and stored as images labeled with the specific foot gestures: dorsiflexion, plantarflexion, inversion, eversion, and toe spreading. To facilitate the use of these images for machine learning purposes, the first 7 out of the 10 repetitions of each foot gesture were initially used as the

training set, while the remaining 3 repetitions were designated as the test set. However, this method did not yield the desired results in terms of performance and accuracy. By utilizing MATLAB's Deep Neural Network application, we observed that configuring the data with a randomized split, where 70% was allocated for training and 30% for testing, produced significantly better outcomes. This method allowed for an effective evaluation and validation of the muscular activity patterns and the classification of foot gestures among healthy individuals. By leveraging both CWT and STFT analyses, this study not only highlights the advantages of combining different time-frequency analysis methods but also demonstrates their effectiveness in capturing the complex dynamics of EMG signals associated with foot gestures. These insights can significantly enhance the development of assistive devices, rehabilitation protocols, and human-computer interaction systems by providing a deeper understanding of muscle activity and movement patterns [12].

CWT & STFT Analysis on EMG Signals

The CWT analyses were conducted based on the organization of EMG signals associated with foot gestures. Within the CWT framework, the impact of mother wavelet on EMG signal data accuracy was investigated to shed light on their contributions to deep learning [13]. Mother wavelets such as Amor, Morse, Bump, Paul, and Morlet were considered within the CWT, with a specific focus on discerning the differences among them. Each wavelet method was chosen for its unique ability to capture various characteristics of the EMG signals, thereby enhancing the robustness of the analysis. In addition to CWT, Short-Time Fourier Transform (STFT) was also employed to analyze the EMG signals [14]. STFT provides a time-frequency representation of the signal, making it useful for identifying transient features and changes in frequency content over time. This method complements CWT by offering a different perspective on the EMG signal characteristics, enabling a more comprehensive analysis. To assess the prediction accuracy between training sets and test data, models were constructed using various deep learning and transfer learning techniques [15]. Specifically, deep learning applications employed included Inceptionv3, ResNet18, ResNet101, InceptionResNetV2, and ResNet50. These models were trained and tested using the spectrogram images generated from both CWT and STFT analyses of the EMG signals. The selection of these models was based on their proven effectiveness in handling image data and their capacity to learn complex patterns from the input data [16]. The constructed models were then compared and evaluated to determine the influence of different CWT parameters and the use of STFT on EMG data accuracy, and to explore the efficacy of deep learning methods in classifying foot gestures. This comprehensive approach aimed to elucidate the impact of CWT parameters on EMG data accuracy and to examine the performance of different deep learning architectures, particularly in the context of distinguishing between various wavelet methods like Amor, Morse, Bump, Paul, and Recker, as well as the contribution of STFT, within the specified foot gestures. By leveraging these advanced techniques, the study provides valuable insights into the nuances of muscle activity during foot gestures and demonstrates the potential of combining CWT and STFT with deep learning to enhance the accuracy and reliability of EMG signal analysis [17]. Moreover, the findings highlight the importance of selecting appropriate wavelet parameters, time-frequency analysis methods, and deep learning models to achieve optimal classification performance. This research not only advances our understanding of EMG signal processing but also opens up new avenues for developing more effective prosthetic control systems, rehabilitation tools, and human-computer interaction interfaces [18]. Overall, the integration of CWT, STFT, and deep learning presents a powerful methodology for the analysis and classification of EMG signals, paving the way for future innovations in biomedical engineering and movement science. The combination of these techniques allows for a more detailed and accurate representation of EMG signals, leading to improved diagnostic and therapeutic applications [19].

Continuous Wavelet Transform (CWT)

The Continuous Wavelet Transform (CWT) is used to analyze the time-frequency content of a signal [20]. The CWT of a signal x(t) is defined as:

$$\boldsymbol{CWT(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t)\varphi^{*}\left(\frac{t-b}{a}\right)\ \mathbf{dt}}$$

- **x(t):** The input signal.
- $\boldsymbol{\varphi}$**(t):** The mother wavelet.
- **a:** The scaling factor (dilates or compresses the wavelet).

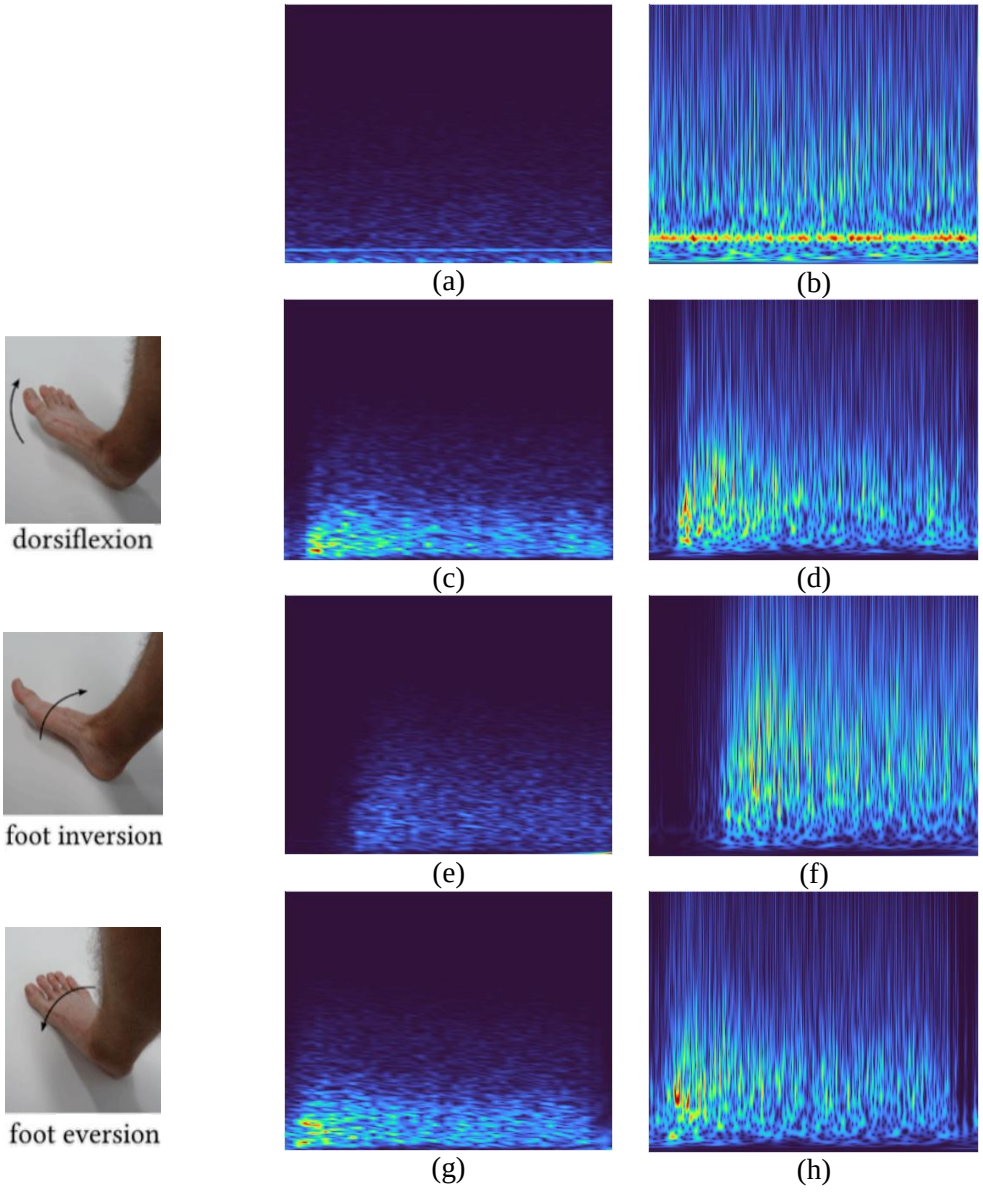
- **b:** The translation factor (shifts the wavelet).
- $\boldsymbol{\varphi^*(t)}$**:** The complex conjugate of the mother wavelet.

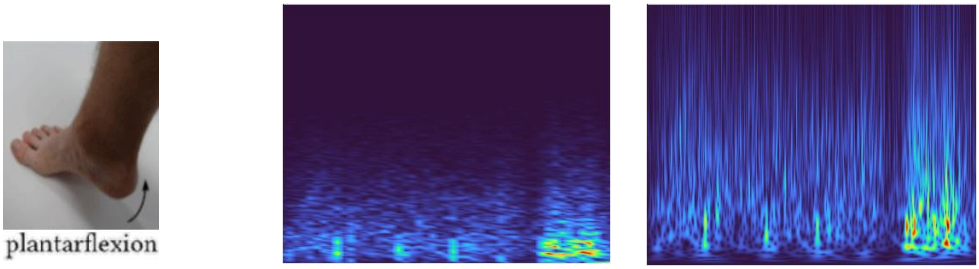
Short-Time Fourier Transform (STFT)

The Short-Time Fourier Transform (STFT) is used to analyze the frequency content of a signal over time [20]. The STFT of a signal x(t)x(t)x(t) is defined as:

$$\boldsymbol{STFT\{x(t)\}(t,f) = \int_{-\infty}^{\infty} x(\tau)\omega(\tau - t)e^{-i2\pi f\tau}\ d\tau}$$

- **x(t)**: The input signal.
- $\boldsymbol{\omega(t)}$: The window function, typically a Hamming or Gaussian window.
- $\boldsymbol{f}$: The frequency.
- **t**: The time shift.





(i) (j)

Figure 1. The spectrogram (a) and scalogram (b) of resting state EMG signals. Spectrograms of five distinct foot gestures, (c) dorsiflexion, (e) foot inversion, (g) foot eversion, and (i) plantarflexion, derived from STFT analysis of the EMG signals. Scalogram of five distinct foot gestures, (d) dorsiflexion, (f) foot inversion, (h) foot eversion, and (j) plantarflexion, derived from CWT-Morse analysis of the EMG signals.

Experimental Results and Discussions

In investigating the effects of CWT wavelet parameters on training data for five distinct foot gestures, our study implemented a Deep Learning design employing InceptionV3, ResNet18, ResNet101, InceptionResNetV2, and ResNet50. According to experimental results in Table 1, the Morse mother wavelet of CWT yielded the highest accuracy of 78.9% using ResNet101, showcasing its effectiveness. Similarly, the Amor mother wavelet of CWT achieved its best result with InceptionV3, reaching 73.9% accuracy. The Bump mother wavelet of CWT also performed exceptionally well with InceptionV3, attaining 73.9% accuracy. For the Morlet mother wavelet of CWT, the highest accuracy of 44.4% was observed using ResNet18, while the Paul mother wavelet of CWT achieved its best accuracy of 66.1% with InceptionV3. Finally, the STFT method recorded its highest accuracy of 73.9% with ResNet18. These findings underscore the critical role of parameter selection and deep learning model choice in enhancing accuracy for bio-signal processing, particularly in foot gesture recognition. The results demonstrate the possible advantages of using the Morse parameter in combination with ResNet101 and the effectiveness of substituting other factors with suitable models.

Table 1. Transfer learning results in terms of Accuracy (ACC)

		ResNet18	**ResNet50**	**ResNet101**	**InceptionV3**	**Inception+ ResNetV2**
Epoch-Iterations	**TF-Method**	**ACC**	**ACC**	**ACC**	**ACC**	**ACC**
20-210	**Amor**	%60.6	%43.6	%66.7	%73.9	%66.7
20-210	**Bump**	%67.2	%71.1	%57.8	%73.9	%66.1
20-210	**Morlet**	%42.8	%44.4	%40.6	%41.1	%42.2
20-210	**Morse**	%67.8	%77.2	**%78.9**	%73.3	%74.4
20-210	**Paul**	%64.4	%60.6	%41.1	%66.1	%61.7
20-210	**STFT**	%51.1	%73.9	%65	%71.7	%63.9

To demonstrate the efficacy of the TF-based foot gesture identification technique, more thorough analyses are carried out using the ResNet101 model, taking into account the outcomes listed in Table 1. It is evident from this table that the CWT method's best classification accuracy for foot gestures is provided by the Morse mother wavelet. The confusion matrices and ROC graphs derived from test images of the STFT and CWT-Morse approaches, utilizing the ResNet101 Deep learning approach, are thus displayed in Figure 2. Notably, the ResNet101 deep learning technique achieves a classification score of over 83% for the CWT-Morse images of the resting state, Dorsiflexion, and Plantarflexion foot gestures. The only aspect of the TF images that falls short of expectations is the detection accuracy of the foot eversion gesture. The ROC-AUC values (Given in Figure 2 (c, d)) for all foot gestures are above 87.23 for CWT-Morse, indicating that the proposed method is successful in detecting foot gestures.

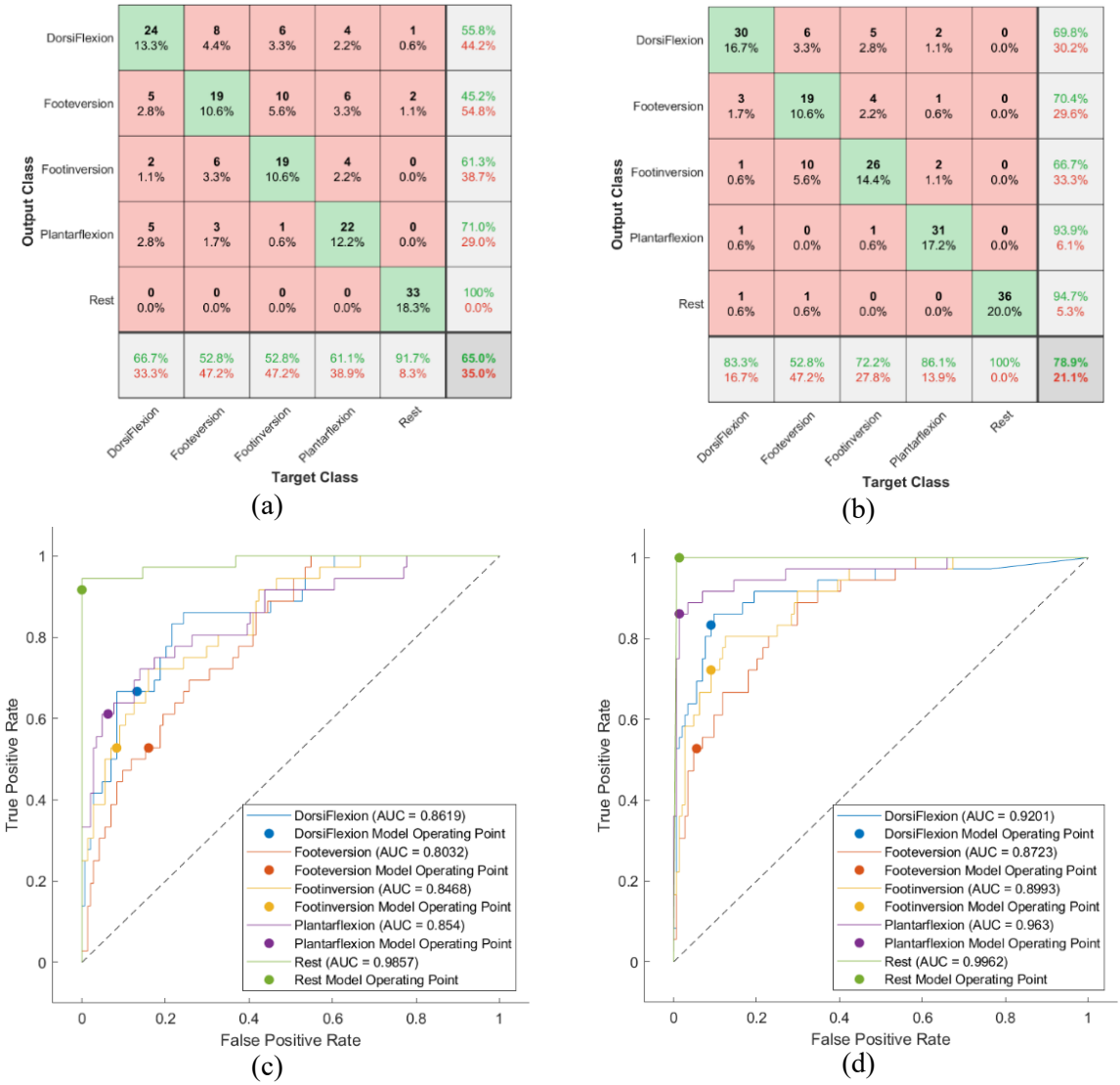


Figure 2. Confusion matrices and ROC graphs created by the ResNet101 deep learning model for evaluating the data of the (a, c) STFT and (b, d) CWT-Morse approaches

Conclusion

In the examination of sEMG foot data, the results are consistent with expectations, providing a solid foundation for analysis. Each dataset for the distinct foot gestures is systematically organized into separate folders, with each structured as a 12x10 matrix. This organization facilitated the application of Continuous Wavelet Transform (CWT) analyses, which are conducted to generate spectrograms corresponding to each foot gesture. These spectrograms are subsequently subjected to deep learning methodologies, where validation accuracies are calculated based on the recorded spectrograms. Our analysis explored various wavelet parameters, with the Morse mother wavelet emerging as the most effective. When combined with the ResNet101 architecture, it achieved the highest accuracy of 78.9%. This finding underscores the Morse wavelet's significant role in boosting the performance of deep learning models applied to sEMG data. Similarly, the Amor parameter achieved its best performance with InceptionV3, recording an accuracy of 73.9%. The Bump parameter, another prominent choice, also performed well with InceptionV3, achieving 73.9% accuracy. The Morlet parameter yielded its highest accuracy of 44.4% when analyzed using ResNet50, while the Paul parameter achieved an accuracy of 66.1% with InceptionV3. Lastly, the STFT method recorded its highest accuracy of 73.9% with ResNet50. These results underscore the importance of careful parameter selection in combination with the right deep learning architecture to optimize performance in foot gesture recognition. The Morse parameter paired with ResNet101 showed exceptional potential, further demonstrating that the choice of both wavelet parameter and network architecture plays a crucial role in enhancing the model's accuracy for sEMG-based classification.

Given the promising results observed, further research has been planned to refine and enhance the existing methodologies. The evaluation of foot gestures, despite being conducted with limited data, has provided valuable insights. Efforts have also been directed toward reducing training times in these computationally intensive studies. Future work will focus on optimizing deep learning architectures to further improve accuracy, particularly in distinguishing between complex movements. To advance this research, additional studies will explore the optimization of wavelet parameters and fine-tuning of deep learning networks to achieve even greater accuracy. Expanding the dataset and incorporating cross-validation techniques could enhance model robustness and generalizability. Real-time processing and deployment of these models in medical, rehabilitation, and prosthetic applications will also be considered, as precise interpretation of sEMG signals is crucial in these domains. Additionally, the application of transfer learning techniques could enable the utilization of pre-trained networks, leading to faster convergence and improved performance, especially when working with smaller datasets. The findings from this study reaffirm the significance of wavelet analysis and deep learning in bio-signal classification, setting the stage for further advancements in sEMG data interpretation. As research progresses, the integration of CWT with deep learning is expected to contribute significantly to various applications, from clinical diagnostics to the development of advanced prosthetic devices and rehabilitation technologies.

References

[1] Merletti, R., & Farina, D. (2016). Surface electromyography: Physiology, engineering, and applications. John Wiley & Sons.

[2] Phinyomark, A., Limsakul, C., & Phukpattaranont, P. (2011). Application of wavelet analysis in EMG feature extraction for pattern classification. Measurement Science Review, 11(2), 45-52.

[3] Hudgins, B., Parker, P., & Scott, R. N. (1993). A new strategy for multifunction myoelectric control. IEEE Transactions on Biomedical Engineering, 40(1), 82-94.

[4] Englehart, K., & Hudgins, B. (2003). A robust, real-time control scheme for multifunction myoelectric control. IEEE Transactions on Biomedical Engineering, 50(7), 848-854.

[5] Hermens, H. J., Freriks, B., Disselhorst-Klug, C., & Rau, G. (2000). Development of recommendations for SEMG sensors and sensor placement procedures. Journal of Electromyography and Kinesiology, 10(5), 361-374.

[6] Phinyomark, A., Nuidod, A., Phukpattaranont, P., & Limsakul, C. (2012). Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification. Electronics and Electrical Engineering, 122(6), 27-32.

[7] Karlsson, S., Yu, J., & Akay, M. (2000). Time-frequency analysis of myoelectric signals during dynamic contractions: A comparative study. IEEE Transactions on Biomedical Engineering, 47(2), 228-238.

[8] Boostani, R., & Moradi, M. H. (2003). Evaluation of the forearm EMG signal features for the control of a prosthetic hand. Physiological Measurement, 24(2), 309-319.

[9] Reaz, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: Detection, processing, classification, and applications. Biological Procedures Online, 8(1), 11-35.

[10] Oskoei, M. A., & Hu, H. (2008). Support vector machine-based classification scheme for myoelectric control applied to the upper limb. IEEE Transactions on Biomedical Engineering, 55(8), 1956-1965.

[11] H. J. Hermens, B. Freriks, R. Merletti, D. Stegeman, J. Blok, G. Rau, C. Disselhorst-Klug & G. Hägg, "European recommendations for surface electromyography," Roessingh Research and Development, vol. 8, no. 2, pp. 13-54, 1999.

[12] Kakarwal, S., & Paithane, P. (2022). Automatic pancreas segmentation using ResNet-18 deep learning approach. System research and information technologies, (2), 104-116.

[13] Cura, O. K., Akan, A., & Atli, S. K. (2024). Detection of Attention Deficit Hyperactivity Disorder based on EEG feature maps and deep learning. Biocybernetics and Biomedical Engineering, 44(3), 450-460.

[14] Jadhav, P., Rajguru, G., Datta, D., & Mukhopadhyay, S. (2020). Automatic sleep stage classification using time–frequency images of CWT and transfer learning using a convolutional neural network. Biocybernetics and Biomedical Engineering, 40(1), 494-504.

[15] Ozdemir, M. A., Cura, O. K., & Akan, A. (2021). Epileptic EEG classification by using time-frequency images for deep learning. International journal of neural systems, 31(08), 2150026.

[16] Escabí, M. A. (2005). Biosignal processing. In Introduction to Biomedical Engineering (pp. 549-625). Academic Press.

[17] Narin, A., & Isler, Y. (2021). Detection of new coronavirus disease from chest x-ray images using pre-trained convolutional neural networks. Journal of the Faculty of Engineering and Architecture of Gazi University, 36(4), 2095-2107.

[18] Kudithipudi, D., Saleh, Q., Merkel, C., Thesing, J., & Wysocki, B. (2016). Design and analysis of a neuromemristive reservoir computing architecture for biosignal processing. Frontiers in neuroscience, 9, 502.

[19] Ceglia, A., Verdugo, F., & Begon, M. (2023). Biosiglive: an open-source Python package for real-time biosignal processing. Journal of Open Source Software, 8(83), 5091.

[20] Prabakaran, B. S., Akhtar, A., Rehman, S., Hasan, O., & Shafique, M. (2021). BioNetExplorer: Architecture-space exploration of biosignal processing deep neural networks for wearables. IEEE Internet of Things Journal, 8(17), 13251-13265.