Exploiting Multicontact Nerve-Cuff MIMO Modeling for Neural Network-based Classification of ENG Signals

Antonio Coviello[1], Francesco Linsalata[1], Umberto Spagnolini[1], Atul Kumar[2], Maurizio Magarini[1]

[1] Department of Electronics Information and Bioengineering, Politecnico di Milano, Milan, Italy
[2] Department of Electronics Engineering Indian Institute of Technology (BHU) Varanasi, India
antonio.coviello@polimi.it

Abstract. Peripheral neuropathies represent a significant medical challenge, severely impacting health and quality of life. This study leverages a Multiple Input Multiple Output (MIMO) model to classify electroneurographic (ENG) signals recorded via multi-contact cuff electrodes. By optimizing neural network architectures, including the Electroneurographic Network (ENGNet) and Inception Transformer (IT), the study demonstrates superior classification performance, with ENGNet achieving the highest accuracy and robustness. An extensive ablation study offers detailed insights into the roles of individual network components. These results pave the way for integrating these models into neuroprosthetic systems, potentially enabling real-time sensory and motor restoration with high precision and efficiency.

Keywords: Electroneurographic (ENG) Signal, Neural Networks Classifications, Peripheral Nerve Interface, Real time application.

1 Introduction

Peripheral neuropathies (PNs) represent an important challenge in the medical field, becoming one of the major health issues. Nerve deficits can significantly impact an individual's quality of life [1]. Several studies aimed at addressing the issue of PNs. Both non-surgical approaches, such as medication use, and surgical ones, such as autografts or regenerative medicine techniques, have been proposed. However, none of them can guarantee satisfactory recovery. Therefore, it is important to promote multidisciplinary approaches to understand, treat and, effectively manage PNs. In contrast to commercial methods, which require complete nerve healing for functional recovery, implanted devices utilize the residual nerve signal to perform neuromodulation and restore functionality [2]. By employing a fully implanted device, complete recovery is not necessary. The device itself is capable of interpreting and completing the intended action correctly, acting as a support for the patient. Numerous challenges remain to be overcome to ensure the safe commercialization of implanted devices. One of these challenges involves extracting information and effectively analyzing signals to facilitate neuromodulation aimed at restoring lost functionality.

The nerve signal extracted from peripheral nerves, named as the electroneurographic (ENG) signal, consists of a wide range of simpler signals, called spikes or action potentials (AP), which represent the fundamental unit of information transmission in the body. These signals, traveling along the axons of neurons, are capable of reaching the desired target, such as a muscle. Such a system can be formulated as a multiple-input multiple-output (MIMO) ENG model that accounts for the aggregate propagation of the nerve motor and sensory signal nano activities. The model, which was pioneered in [2], arises from the deployment of highly spatial selective multi-contact cuff electrodes to measure the ENG signals. The temporal and spatial summation of different APs generate the ENG signal. Multiple APs are needed to generate complex body commands, such as arm movement. When neuropathy occurs, signal interaction is limited by the damaged axons. This results in the loss of part of the ENG signal and thus compromises the correct execution of intended movements [3]. Therefore, proper modelling, acquisition and classification of the ENG signal are the essential aspects that need to be assessed to ensure the restoration of the lost functionality. Those are enrolled by the use of implanted devices [3]. One of the main issues related to the implanted device is the signal analysis. Several studies have been conducted to evaluate its feasibility. In [4], selective analysis of APs is proposed using a thresholding method. Convolutional classifiers are then used to classify the signal. In this case, a group of 3 classes is considered, achieving mean accuracy values of 80% and an F1 score of 75%. This approach focuses the attention about single spikes exalting spatio-temporal relationships. Alternatively, there are methods that enhance global information. In [5], a feature extraction framework is performed using moving observation windows and different types of machine learning techniques. In this case, the classification is performed based on intensity levels without changing the class. Windows with variable ranges, between 50 and 350 ms, have been considered, achieving accuracy values above 90%. The implementation of classifiers based on neural networks (NNs), which use kernels for feature selection and subsequent classification, appears to be a promising approach [2]. In [6] they attempted to decode the motor intent of an individual, showing that this approach is one of the most suitable for the ENG signal classification. Others convolution classifiers, as inception time (ITs) where tested. Its building block consist in repetition of inception module. The purpose is to enhance the intrinsic features of the signal. An example of an IT network is [7], where image classification was performed. The ENG network (ENGNets) are another example. ENGNets utilize a combination of convolutional and clustering layers, leveraging depth and separable convolutions to efficiently process ENG data by capturing significant temporal and spatial patterns for classification. This network was inspired by the Electroencephalographic Network (EEGNet) [8], originally used for brain-computer interface applications. Lastly, the long short-term memory (LSTMs) are recurrent-based models capable of storing information for a longer period compared to standard recurrent neural networks. Examples of LSTM networks applied to ENG signals are reported in [2], where different signal windows were processed to evaluate real-time applications.

Contributions: To better understand the relationship between kernels of the NN and ENG signal modelling, in this article we leverage on the mathematical modelling proposed in [2] to correlate the waveform of individual spikes with the optimal kernel size for classification. Indeed, the use of kernels enables a comprehensive analysis, encompassing both global and local information. This study is obtained thank to the emulation of ENG signal propagation by means of Comsol multiphysics [9] simulator, which is here highlighted. Different types of convolutional networks are. By contrast to NN studied in [2], in this paper we investigated also the use of Convolutional Tiny Transformer (CTTs) for ENG classification. Moreover, we perform an ablation study to determine which convolutional block has the greatest relevance in the ENG classification, providing insights for the development of future approaches. This exploration helps in identifying the best strategies to enhance the effectiveness of NNs in classifying ENG signals, paving the way for new developments in restoring lost function in real-time applications.

Organization: The rest of the paper is organized as follows. Section II introduces the MIMO model of the ENG signal propagation. In Sec. III we show how this modelling can be used within COMSOL to emulate an ENG signal. Section IV reports the NN classification results and the ablation study. Section V concludes and summarizes the work.

2 Neural network-based ENG signals classification

Nervous signals propagate through axons, elongated extensions of neurons, which conduct APs, facilitating communication and coordination within the body. The measurement of the spatial and temporal linear combinations of the individual APs, as shown in Fig. 1, can be achieved by placing a cuff electrode around the nerve to measure the ENG signal. However, given the presence of thousands of axons within a nerve, there arises the need to develop a model that accurately and simply describes the behavior of the ENG signal. This will improve the extraction of information contained within the nerve signal using appropriate kernels. We recently showed that a MIMO system can be devised to accurately model the transmission of the signal captured through a cuff electrode [2].

2.1 Nerve-cuff MIMO modelling

The MIMO model is developed by considering a simplified representation of the nerve, as reported in Fig. 1. A nerve section is shown with N axons. The number of axons is not known a priori but can be estimated as proposed in [10]. Each axon controls a specific body region. Consequently, performing complex movements requires the coordination of multiple axons. To describe the ENG signal, we start from its building block element, i.e. the spike [11].

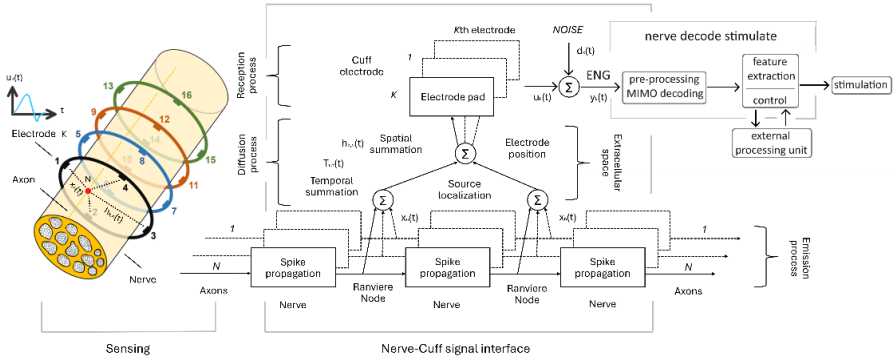


Fig. 1. Schematization of the nerve/axon-cuff electrode interface. From the ENG measurement to the data processing.

A spike associated with the activity of the axon is typically modeled as a unipolar voltage:

$$x_n(t) = \begin{cases} A_n t^{m_n} e^{-B_n t}, & t \geq 0 \\ 0, & t < 0 \end{cases} \tag{1}$$

with $n = 1, \dots, N, A_n \left(\frac{V}{s^m}\right), B_n \left(s^{-1}\right)$, and m_n represent the amplitude of the biological signal, the decay time, and the temporal consistency dependence. It is important to emphasize that different axons exhibit different waveform characteristics. The ENG is not comprised of a single spike but is the sum of multiple APs in space and time. The firing rate of the APs determines the intensity of the selected stimulus. The interspike interval over time is typically modeled as a Poisson process, providing a coding mechanism for information transmission in the nervous system.

The ENG signal observed at the kth electrode of the cuff can be modeled as:

$$u_k(t) = \sum_{n=1}^{N} h_{k,n} \left(\sum_{i=1}^{I_n} x_{i,n}\left(-t + T_{i,n}^{aff}\right) + \sum_{j=1}^{J} x_{j,n}\left(t - T_{j,n}^{eff}\right) \right), \tag{2}$$

where I_n and J_n represent the number of spikes in the n th afferent and efferent axons respectively, and $T_{i,n}^{aff}\left(T_{j,n}^{eff}\right)$ indicates the time corresponding to the i_{th} (j_{th}) pulse in the afferent (efferent) direction. $h_{k,n}$ represents the lead field matrix that models the contribution of the nth axon to the k_{th} electrode, and is determined by the spatial distance and conductivity between them.

The formulation of $h_{k,n}$ is:

$$h_{k,n} = \frac{1}{4\pi\sigma}\left(-\frac{1}{||\boldsymbol{n}-\boldsymbol{k}||^2}\right), \tag{3}$$

where σ represents the temporal conductivity, $||\cdot||$ is the Euclidean norm, and n and k are the tridimentional position vectors of the axon and the electrode, respectively. Finally, the ENG signal can be represented in matrix form as $y(t) = u(t) + d(t)$, where d(t) is the distortion vector that accounts for the interference of the electromyographic (EMG) signal and the thermal noise. This MIMO model allows for the evaluation of the methods for extracting the spatio-temporal characteristics of the ENG signal, which are essential for NN-based classification through the implementation of a finite element simulation, such as COMSOL software [9] as discussed in the next section. The MIMO model proposed in this work provides a robust framework for modeling ENG signals and optimizing neural network kernels, enabling a detailed representation of spatio-temporal ENG characteristics. Unlike traditional feature extraction methods, such as thresholding or manual segmentation [4], or simpler linear models used in neural prosthetics, the MIMO approach captures the complexity of neural signal propagation. For instance, threshold-based classifiers [4] achieved F1-scores of 75%, significantly lower than the ENGNet and CTT models tested in this study. Furthermore, MIMO modeling integrates seamlessly with advanced neural networks, providing high-resolution, multi-dimensional inputs that surpass the limitations of older technologies like single-channel nerve cuffs or basic machine learning algorithms [16]. By combining spatial and temporal interactions in high resolution, the MIMO model bridges the gap between signal acquisition and real-time classification. This integration with deep learning architectures marks a substantial advancement in neuroprosthetic applications, paving the way for more effective and generalizable solutions for mixed or complex nerve signals.

2.2 Neural Network kernels and parameters tuning

The proposed ENG MIMO model in previous subsection can be used for optimizing and tuning the kernels in NNs. Acting as mathematical filters, kernels are applied to the data to extract relevant characteristics, emphasizing specific frequencies present in ENG signals. Utilized as convolutional layers, kernels enable the identification and extraction of significant patterns, facilitating effective differentiation of AP events within the ENG signal and their recurrence over time. The synergistic interaction between the proposed MIMO model and NNs guides the selection of parameters and kernel configurations, facilitating the optimized classification of ENG signals. Indeed, the model described in (2) plays a vital role in extracting features embedded within nervous signals. Defining optimal kernels enables accurate extraction of spikes from measurements acquired using a multi-contact cuff electrode, providing robust parameters for classifying ENG signals through NNs. Equation (2) allows to identifies the key parameters,

including: $T_{i,n}^{aff}$,$T_{j,n}^{eff}$,$x_{i,n}$,$x_{j,n}$, and $h_{k,n}$, guiding kernel determination to maximize NNs performance. By selecting appropriate parameters A_n, B_n, and m_n through Eq. (1), the desired spike waveform generated by the nth axon can be obtained, thereby defining the optimal kernel length. Analyzing the dataset [12], we observe that the typical duration of APs varies up to a maximum of 3ms, as shown in the Fig. 2. However, since the signal is contaminated by external noise, only the peak can be observed, not the entire waveform. With a peak duration ranging from 0.5ms to 1ms, the relevant spectral content for APs lies in the range of $1 - 2kHz$. Considering a theoretical total spike duration of 3ms and a sampling frequency of $f_s = 5kHz$, 15 samples are required to define spike characteristics, representing a starting point for optimal kernel length. However, subsequent corrections must be implemented to account for factors such as subject variability and different network architectures. The length correction can be done empirically by choosing values within 15 and observing if accuracy improves, or by plotting the spectrum of the first kernel filters used in NNs. If a peak is present in the range of $1 - 2kHz$, as reported in Fig. 4 processing has been done correctly; otherwise, a different length should be selected. Performing this operation on all NNs, the optimal kernel lengths for CNN and LSTM have been found to be 9 samples, while for IT, 14 samples are optimal. Regarding ENGNet and CTT, a window size of 100 samples was chosen as the recurrence parameters $T_{i,n}^{aff}$, and $T_{j,n}^{eff}$ are more prominent in the network. Additionally, for the spatial dimension defined by the parameter $h_{k,n}$, a spatial kernel equal to the number of channels K, or double that, was defined to capture correlations between signal source positions.

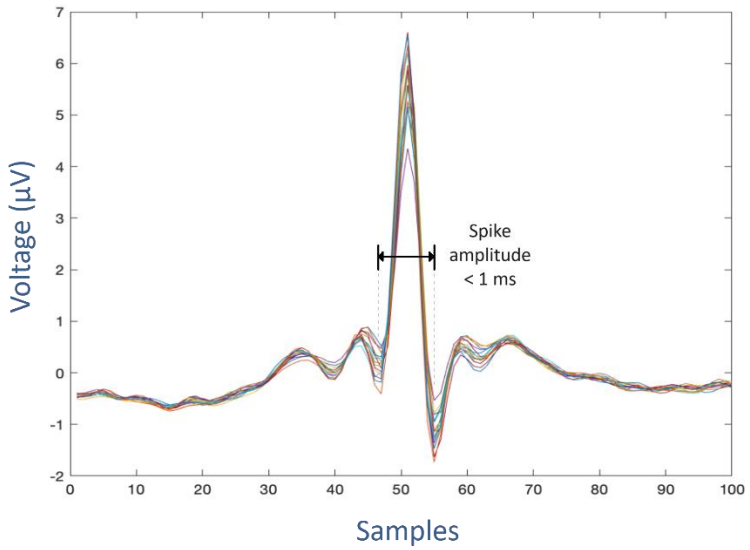


Fig. 2. Real AP signal averaged to provide noise cancellation over the 16 channels, f_s equal to 30kHz.

3 COMOSL-based ENG signal emulation

The ENG MIMO model, discussed in previous section, provides a unique opportunity to investigate the interactions among signals produced by different axons. These interactions constitute the ENG signal measured by the cuff electrode surrounding the nerve. In this section we utilize the multiphysics software COMSOL [9] to conduct numerical simulations. An example of simulation is reported in [13]. We generate realistic synthetic data and meaningfully compare them in terms of signal amplitude and intrinsic characteristics, as report in Fig. 3. ENG signals typically exhibit a frequency band ranging from 500Hz to 7kHz, with an energy peak around 2kHz [14]. It is worth noting that the frequency components depend on the type of electrode used [15]. These sampled and measured signals demonstrate weak amplitudes, typically around 50µV [14]. The primary noise added to the ENG signal, is the EMG, that contain the most of its power below 800Hz [16, 17]. Given that the spectral content of ENG signals mostly falls below 2.5kHz [2], we can mitigate the impact of these impairments in digital preprocessing by applying a band-pass filter with cutoff frequencies between 0.8 and 2.5kHz. Simulations enable us to explore the propagation mechanisms within the nerve, providing a deeper understanding of its functioning. The use of COMSOL in our context aims to create a nerve model for generating neural signals, allowing for the evaluation of APs variations based on the axon model used. The generated simulation features a nerve with a diameter of 1 cm and a length of 6 cm. Various geometries can be employed for

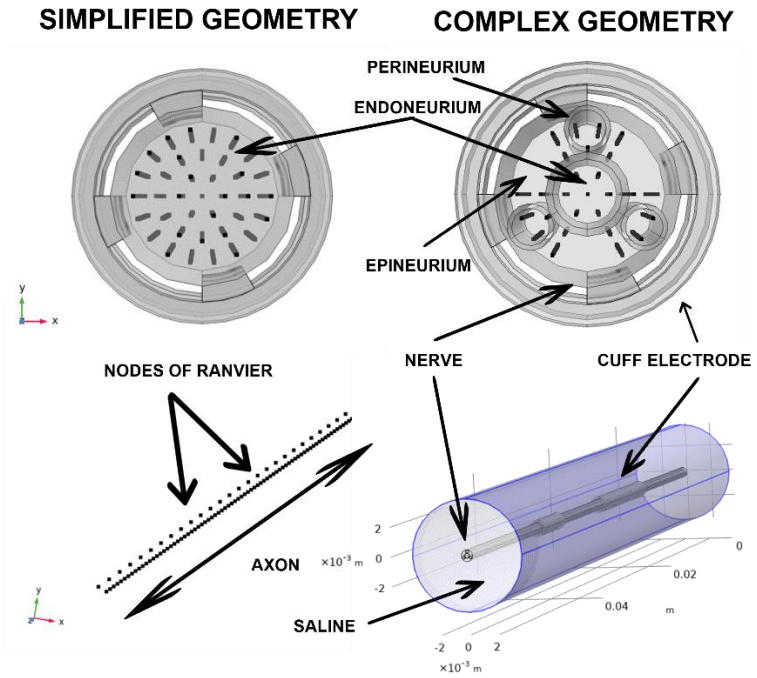


Fig. 3. Nerve geometry used for the COMSOL simulation.

the nerve, ranging from a simplified geometry of a simple cylinder to a more complex one reflecting hierarchical structures or real images. The different axons in the nerve have different diameters, which result in variations in propagation velocity, waveform, and fiber directionality. Axons are modeled as points within the nerve, generating action potentials that need to be recorded by electrodes, with each point representing a Node of Ranvier [13]. Once the waveform is generated, it propagates combines with others until reaching the electrode surface, producing the ideal ENG signal. The waveform undergoes attenuation proportional to the impedance of the saline solution in which the nerve is immersed. Cuff electrodes are placed around the nerve, and the measured signal is the average potential value recorded at each electrode. For signal sampling, an ideal ground was set, while the two bases were set as insulators. Conductivity parameters were set for the conductive (platinum) and insulating (PDMS) parts of the cuff electrode. The simulation includes the physics of AC/DC electrical currents, with specific conditions to ensure correct neural signal recording and localization. After signal recording, Gaussian noise with zero mean and a certain variance is generated for each channel and added to analyze AP variations, with a signal-to-noise ratio (SNR) set at 7.5 dB. Simulated results show that only the peak-associated part is preserved in the noisy ENG signal. Since the observed maximum duration varies from 0.5ms to 1ms, the spectral content is expected to lie in the range of 1-2kHz. This was empirically confirmed by plotting the spectrum of kernel classifiers, as reported in Fig. 4. The system serves as a good model for simulating signal sensing for both split cylinder and spiral cuffs but has some approximations, such as the lack of consideration for cuff electrode closure and any wires exiting the cuff electrode. For further details and insights, refer to [13].

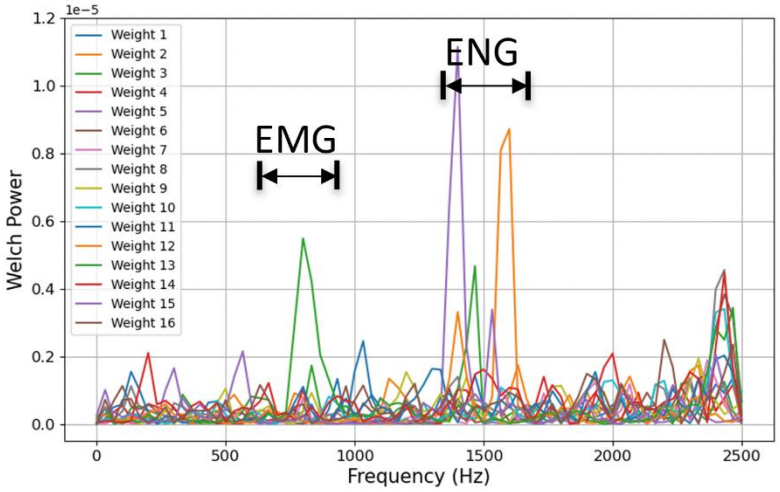


Fig. 4. ENGNet first layer kernel spectrum.

4 Numerical results

Building upon the insights gained from COMSOL simulations, we have defined the kernel parameters for each single network. We have proceeded with the classification of the real data available in [12], evaluating the NNs performance.

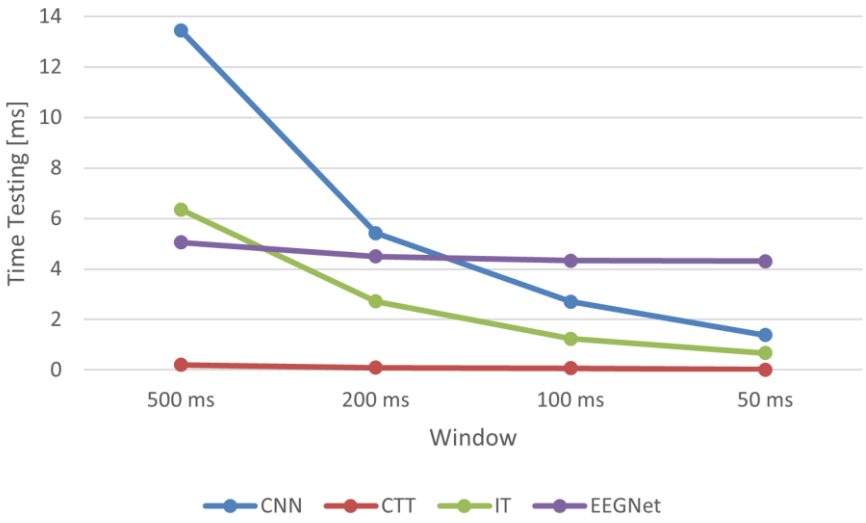
4.1 Signal elaboration and classification

The analysis was conducted on somatosensory ENG signals extracted from three rat on 4 different stimuli: dorsiflexion, plantarflexion, touch, and pain. The dataset is provided in [12]. The ENGsignals were recorded at 30kHz using a multicontact cuff electrode with 16 read channels arranged in a matrix configuration of (4 rings, 4 contacts). The pipeline reported in [2] consist in: bandpass filtering, downsampling, thresholding and signal windowing. The best window duration should be chosen to potentially evaluate real-time clinical operations of the investigated classifiers. The obtained windows are then classified using five different types of NNs architectures with various complexities. Further details about the networs are reported in [2]. A study was conducted to optimize the training time of the networks, using early stopping to mitigate overfitting. Training stops when the validation accuracy does not improve for 8 consecutive epochs. Additionally, network parameters were calibrated through an ablation study to maximize the performance.

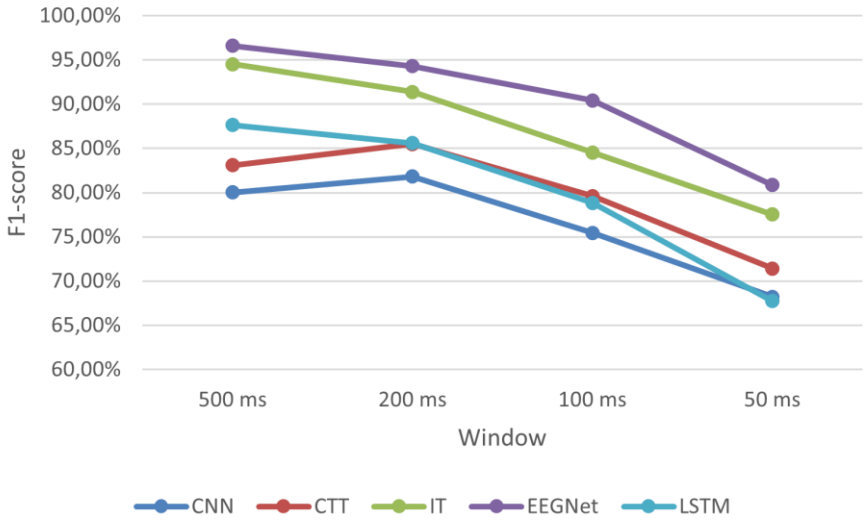
4.2 Neural networks perfomance comparisons

In this section, we present the numerical results for ENG classification by using different NN architectures. The selected architectures are CNN, LSTM, ENGNet, IT given in [2]. As a contribution of this work a CTT (or transformer) is also tested. It represents an innovative neural architecture based on attention, capable of capturing long term relationships regardless of position. Its key feature is the use of an internal attention mechanism that allows it to detect relationships between various parts of a sequence, eliminating the constraint of sequential dependence typical of other architectures such as IT and CNN. Additionally, the transformer adopts an encoder-decoder architecture, ensuring independent processing of different input elements. The addition of kernels to this network allows for the extraction of additional spatiotemporal information from the ENG signal, improving overall performance. This makes it particularly suitable for complex signals such as ENG signals [18]. Three performance metrics were used to evaluate the NNs: F1-score, testing time, and number of parameters. The ENG signal was divided into temporal windows (50, 100, 200, 500 ms) to examine the variation of the F1 score metric in relation to the input length. Each network was evaluated using a 5fold cross-validation setup, where 80% of the data was used for training and the remaining 20% for testing.

Average F1-score: As observed in Fig. 5a, the F1-score decreases with the observation window duration. This is due to the reduction in input data length and the associated information content. However, there are some exceptions for the 500ms window. Both CTT and CNN exhibit lower F1 score values compared to the 200ms window. The reason is that both networks, particularly CTT, tend to overfit due to the limited amount of data available for 500ms windows. Compared to networks like IT and ENGNet, CTT rapidly goes into overfitting if there is not a high number of input samples. The CTT excels in temporal performance, achieving classification times as low as 0.02 ms for 50 ms input windows (Fig. 5b). However, its accuracy lags behind other models, such as IT and ENGNet, particularly for longer input windows (Fig. 5a). This trade-off between speed and accuracy is evident across all tested input lengths. While the CTT's lightweight design and minimal computational requirements make it appealing, its overfitting tendencies and lower accuracy limit its suitability for clinical applications. Enhancing the network's architecture to balance speed with accuracy could make it a robust candidate for real-time use. The CTT's overfitting behavior applies to all tested input lengths, particularly for longer observation windows (e.g., 200ms and above, Fig. 5b). This results from the CTT's limited capacity to generalize across high-variability datasets, as longer windows amplify noise and redundant information, compromising the training process. Despite these challenges, strategies such as regularization, dropout, or reduced window lengths could help mitigate overfitting while preserving the network's efficiency. In [18], CTT demonstrates outstanding performance in EEG signal analyses. The discussion becomes particularly intriguing when comparing the Transformer with EEGNet, as the Transformer appears to prevail for EEG signal. However, if the same comparison is extended between CTT and ENGNet for ENG signal analysis reported in this paper, a completely opposite trend emerges. This shifts the focus to ENGNet, as previously highlighted in [2]. None of the windows associated with CTT has reached a value higher than 90%, making this classifier less appealing compared to IT or ENGNet for clinical and safe implementations on patients. We addressed the class imbalance and limited data in ENG datasets using a data augmentation strategy. This approach involved an 80% overlap of activity periods, as illustrated in [19]. Data augmentation techniques, as described in Section 4.1, significantly improve the classification performance of underrepresented classes, but they also reveal the CTT's dependency on high data variability. Achieving performance comparable to IT requires a substantial dataset, and even with augmented data, the training phase slows considerably, offsetting the advantage of a lightweight testing process. Despite these challenges, the overlapping strategy increased the training sample size without altering the test set, enhancing classification accuracy, especially for underrepresented classes. For example, the pain stimulus class showed an approximate 20% improvement in F1-score. These results highlight the importance of balancing training efficiency and data diversity to optimize classifier performance. Future work could explore strategies such as reducing the input window length or introducing additional convolutional blocks to improve feature extraction efficiency. All networks tested demonstrated improvements in performance, although to varying degrees. IT showed the greatest increase in F1-score, while CTT exhibited the least improvement. This indicates that, although CTT can effectively enhance its performance, it requires a significantly larger amount of data to achieve this



(a)



(b)

Fig. 5. Average F1-score a) and test time b) values per sample per window of the studied classifiers useful for real time applications. Time LSTM value are not included because of his too low performance.

goal, making it less suitable for real-time applications in this field. We can affirm that ENGNet emerges as the most suitable candidate for real-time implementation. Specifically, the windows of 100ms and 200ms could be suitable for realtime applications. Accuracy, which is not reported in this analysis, shows values 5% to 10% higher than F1 scores.

Testing Time: In implanted systems for real-time applications, response time is one of the critical and essential parameters for proper device implementation with temporal requirements. The CTT achieved remarkable results, as shown in Fig. 5b. From the 500 ms to the 50 ms of window size the total classification times were 0.2, 0.1, 0.08, and 0.02 ms, respectively. Compared to other networks, the CTT demonstrated temporal performance superior by one order of magnitude to IT, two orders to CNN and ENGNet, and three orders to LSTM. This leads us to conclude that if the classifier's performance for smaller windows could be improved, as ENGNet, this network could be the most suitable for medical applications with temporal constraints or limitations.

Number of Parameters: The last consider parameter is the computational load and consequently the implementability of the algorithm on devices, which often have limited computing capabilities. The parameter count for each network resulted in: 1,242,116 for LSTM, 577,956 for CNN, 50,948 for IT, 5,484 for CTT, and 4,964 for ENGNet. This highlights that CTT has a low number of parameters, and like ENGNet. If its performance could be improved, it could be a network of interest for devices with limited capacities.

Ablation Study: An ablation analysis was conducted to examine both the input characteristics and the individual components defining the classifiers. The input arrangement study and the optimal number of parameter where reported in detal in [2]. The LSTM study was not conducted due to its identified non-optimality for this type of signal. For the ENGNet and CTT networks, using a first line of kernels with length 100 samples was able to maximize F1 score. Smaller or larger lengths (50 and 150 samples, respectively) resulted in lower performances by approximately 5-10%. Removing one block for time we were able to study the contribution of each block. For ENGNet, as reported in Fig. 6, the 1D Convolutional Block, led to the greatest reduction, while Depthwise Convolution and Separable Convolution, reported a minor reduction. The CTT network consists of three distinct blocks: Patch Embedding, Temporal Convolution, and Spatial Convolution. As reported in Fig. 6, all three blocks contributed equally to the network. For IT, 3 kernels per block were used (three consecutive blocks repeated twice, for a total of six [2]). As reported in Fig. 6, eliminating the repeated blocks does not significantly change the performance, expect for the F1-score that results in less stability. Regarding the kernels, when their length are comparable to the ones of the AP significant results are obtained. The analysis of F1-scores, testing times, and parameter counts reveals the trade-offs inherent in different neural network architectures. ENGNet consistently outperforms other models in terms of accuracy, particularly for 100ms and 200ms windows, making it the most reliable for precise signal classification. IT offers a balance between performance and flexibility, excelling in scalability to various input

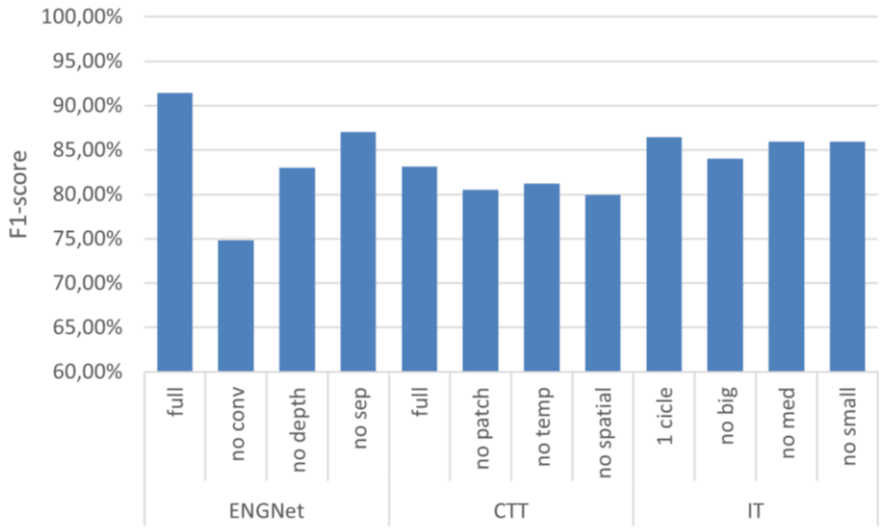


Fig. 6. Ablation study results for 100 ms input window length.

lengths. The CTT, while excelling in testing time due to its lightweight design, shows limitations in accuracy, especially for longer windows. These findings suggest that the choice of architecture should be guided by the specific requirements of the application, balancing speed, accuracy, and resource constraints to ensure optimal clinical outcomes.

5 Conclusion

This study addressed the real-time classification of sensory and motor stimuli delivered by ENG signals, aiming to overcome peripheral neuropathy health issues. A MIMO modeling approach was utilized to characterize ENG signals and to build the kernels of several neural network classifiers. This MIMO model provides a unique opportunity to investigate interactions among ENG signals produced by different axons, offering critical insights into neural activity. Several neural network-based methods were proposed and assessed using both measurements and simulations of ENG signals. Among the tested architectures, ENGNet emerged as the most reliable candidate for real-time clinical applications due to its superior classification accuracy and robustness across all tested windows. IT, while slightly less accurate, demonstrated excellent scalability and adaptability to varying input lengths, making it suitable for versatile configurations. Additionally, the remarkably low parameter count and fast execution speed of CTT, presented here for the first time, highlight its potential for resource-constrained applications, albeit with limitations in classification accuracy for longer observation windows. The adaptability of the proposed models to various types of nerves and neuroprosthetic devices represents a critical aspect for future research. While this study

focused on somatosensory ENG signals recorded from rat nerves, the architecture of the models can be extended to motor or mixed sensory-motor nerve signals. This extension would require modifications in training datasets, including the collection of new datasets incorporating these features. Furthermore, the proposed models are not limited to applications in rats but can be extended to human nerves, highlighting their potential for broader clinical applications. These findings demonstrate the feasibility of integrating ENGNet and IT into neuroprosthetic systems to restore sensory and motor functions with high precision and efficiency. Future research will focus on further optimizing the models and extending the datasets to ensure scalability and robustness in diverse clinical settings, paving the way for practical and reliable neuroprosthetic solutions.

Acknowledgements

This study was supported by the AS23VARI19 ERASMUS+ project (KA171 CALL 2022 PROJECT) between the Polytechnic University of Milan and the Indian Institute of Technology (BHU) Varanasi.

References

1. Coviello, Antonio, et al. "Emerging peripheral nerve injuries recovery: advanced nerve-cuff electrode model interface for implantable devices." *GCWOT'24* (2024): 1-7.
2. A. Coviello, F. Linsalata, U. Spagnolini, M. Magarini et al., “Artificial neural networks-based real-time classification of eng signals for implanted nerve interfaces,” arXiv preprint arXiv:2403.20234, 2024.
3. G. Hussain, J. Wang, A. Rasul, H. Anwar, M. Qasim, S. Zafar, N. Aziz, A. Razzaq, R. Hussain, J.-L. G. de Aguilar et al., “Current status of therapeutic approaches against peripheral nerve injuries: a detailed story from injury to recovery,” International journal of biological sciences, vol. 16, no. 1, p. 116, 2020.
4. R. G. Koh, M. Balas, A. I. Nachman, and J. Zariffa, “Selective peripheral nerve recordings from nerve cuff electrodes using convolutional neural networks,” Journal of Neural Engineering, vol. 17, no. 1, p. 016042, 2020.
5. B. E. Silveira C., Khushaba R. N. and N. K., “Spatio-temporal feature extraction in sensory electroneurographic signals,” Phil. Trans. R. Soc. A., 2022.
6. D. K. Luu, A. T. Nguyen, M. Jiang, J. Xu, M. W. Drealan, J. Cheng, E. W. Keefer, Q. Zhao, and Z. Yang, “Deep learning-based approaches for decoding motor intent from peripheral nerve signals,” Frontiers in Neuroscience, vol. 15, p. 667907, 2021.
7. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.
8. L. Feng, L. Yang, S. Liu, C. Han, Y. Zhang, and Z. Zhu, “An efficient eegnet processor design for portable eeg-based bcis,” Microelectronics Journal, vol. 120, p. 105356, 2022.
9. “Comsol multiphysics.” [Online]. Available: https://www.comsol.it
10. L. N. Mishra, G. Kulkarni, and M. Gadgil, “Modeling the impact of the variation in peripheral nerve anatomy on stimulation,” Journal of Pain Research, pp. 4097–4111, 2022.
11. K. R. Grider MH, Jessu R, “Physiology, action potential.” StatPearls [Internet], Treasure Island (FL): StatPearls Publishing, 2024.

12. E. Brunton, C. Blau, C. Silveira, and K. Nazarpour, “Identification of sensory information in mixed nerves using multi-channel cuff electrodes for closed loop neural prostheses,” in 2017 8th International IEEE/EMBS Conference on Neural Engineering (NER), 2017, pp. 391–394.
13. D. D. F. Del Bono, A. Rapeaux and T. G. Constandinou, “Translating node of ranvier currents to extraneural electrical fields: a flexible fem modeling approach,” 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021.
14. J. C. S. Micera and S. Raspopovic, “Control of hand prostheses using peripheral information,” in IEEE Reviews in Biomedical Engineering, pp. 48–68, 2010.
15. R. A. Mezzarane, L. A. Elias, F. H. Magalhães, V. M. Chaud, and A. F. Kohn, “Experimental and simulated emg responses in the study of the human spinal cord,” Electrodiagnosis in new frontiers of clinical research. Rijeka: InTech, pp. 57–87, 2013.
16. S. Raspopovic, J. Carpaneto, E. Udina, X. Navarro, and S. Micera, “On the identification of sensory information from mixed nerves by using single-channel cuff electrodes,” Journal of neuroengineering and rehabilitation, vol. 7, no. 1, pp. 1–15, 2010.
17. R. Rieger, J. Taylor, A. Demosthenous, N. Donaldson, and P. J. Langlois, “Design of a low-noise preamplifier for nerve cuff electrode recording,” IEEE Journal of Solid-State Circuits, vol. 38, no. 8, pp. 1373–1379, 2003.
18. Y. Song, X. Jia, L. Yang, and L. Xie, “Transformer-based spatial-temporal feature learning for eeg decoding,” arXiv preprint arXiv:2106.11170, 2021
19. C oviello, Antonio, et al. "Strategies for Efficient ENG Signal Classification using Data Augmentation and Data Balancing Techniques." *IEEE MetroXRAINE 2024*. 2024. 1-6