

RESEARCH ON A HAND GESTURE RECOGNITION MODEL TO OPTIMIZE ROBOTIC HAND CONTROL

Bui Thi Ngoc Han, Vi-Do Tran*

Ho Chi Minh City University of Engineering and Technology, Viet Nam

*Email: dotv@hcmute.edu.vn

Received: 18 December 2025; Revised: 6 March 2026; Accepted: 20 April 2026

ABSTRACT

Achieving accurate and responsive control of robotic hands remains a major challenge in human-machine interaction, particularly for applications requiring high dexterity such as prosthetic control and rehabilitation systems. Surface electromyography (sEMG) offers a non-invasive solution for capturing human motion intent; however, its nonlinear, non-stationary nature and susceptibility to noise significantly degrade recognition performance when using conventional methods. This paper proposes a robust sEMG-based hand gesture recognition framework that integrates systematic signal preprocessing, including noise filtering, normalization, and temporal segmentation, with a hybrid deep learning architecture. The proposed model combines a one-dimensional convolutional neural network (1D-CNN) for effective spatial and time–frequency feature extraction with a long short-term memory (LSTM) network to capture long-term temporal dependencies in multi-channel sEMG signals. Six fundamental hand gestures such as Rock, Paper, Scissors, Pointing, Hand Flexion, and Hand Relaxation are evaluated using a custom annotated sEMG dataset. Experimental results demonstrate that the proposed hybrid 1D-CNN–LSTM model achieves consistently high classification accuracies of 96.0% and 0.96 weighted F1-score, outperforming conventional 1D-CNN-based approaches relying solely on time-domain features. Furthermore, the proposed framework exhibits low computational complexity and strong inter-subject generalization, indicating its suitability for real-time robotic and prosthetic hand control applications.

Keywords: Hand gesture recognition, Surface electromyography (sEMG), 1D-CNN, EMG classification.

1. INTRODUCTION

Hand gesture recognition has emerged as a key enabling technology in modern human–machine interaction, providing an intuitive and efficient means for conveying human intent in applications such as prosthetic control [1-5], rehabilitation systems [6-9], and intelligent robotics. Among various sensing modalities, bio-signal based approaches have attracted increasing attention due to their ability to capture physiological information associated with muscle activation. In particular, surface electromyography (sEMG) offers a noninvasive [10] and direct measurement of neuromuscular activity, making it a promising candidate for developing responsive and natural gesture recognition systems [11-13] that are largely independent of external environmental conditions.

Despite these advantages, achieving accurate and robust sEMG-based hand gesture recognition under practical operating conditions remains a significant challenge. sEMG signals exhibit strong nonlinearity and non-stationarity and are highly sensitive to multiple sources of variability, including noise contamination, muscle fatigue, electrode displacement,

and inter-subject anatomical differences. These factors lead to substantial intra and inter-class variations, which complicate the reliable extraction of discriminative features and limit the generalization capability of recognition models.

Conventional sEMG recognition pipelines typically rely on handcrafted time-domain features, such as root mean square (RMS), mean absolute value, or waveform length, followed by shallow classifiers. While these approaches are computationally efficient, their limited representational capacity often restricts recognition performance, particularly for complex or highly similar hand gestures. More recently, deep learning-based methods [14-16], including one-dimensional convolutional neural networks (1D-CNNs) [17-20] and recurrent neural networks, have demonstrated superior performance by learning discriminative features directly from raw or minimally processed sEMG signals. However, many existing studies focus either on spatial feature extraction through convolutional operations or on temporal sequence modeling using recurrent architectures. As a result, the complementary information embedded in the frequency domain and the long-term temporal evolution of muscle activation patterns is not fully exploited, which constrains the discrimination capability of current models.

To address these limitations, this study proposes a hybrid 1D-CNN–LSTM architecture for sEMG-based hand gesture recognition that integrates RMS-based time-domain features with Fourier-domain spectral representations. The proposed framework leverages the feature extraction capability of 1D-CNNs to capture local spatial patterns across multiple sEMG channels, while long short-term memory (LSTM) networks are employed to model the temporal dependencies of muscle activation sequences. By jointly learning time frequency characteristics and long-term temporal dynamics, the proposed method provides a more comprehensive representation of sEMG signals, thereby enhancing robustness against signal variability and improving discrimination among similar gestures.

Experiments on six hand gestures (Rock, Paper, Scissors, Pointing, Hand Relaxation, Hand Flexion) demonstrate that the proposed hybrid model achieves 96.0% accuracy and 0.96 weighted F1-score, substantially outperforming a baseline 1D-CNN using RMS features alone (86.0% accuracy, 0.86 F1-score). The proposed framework maintains low computational complexity suitable for real-time robotic hand control applications.

The main contributions of this study are threefold: (i) the design of a hybrid 1D-CNN–LSTM framework that explicitly integrates RMS-based time-domain features with Fourier-domain spectral representations to jointly exploit complementary time–frequency information from multichannel sEMG signals; (ii) a unified learning strategy that simultaneously captures spatial feature patterns and long-term temporal dependencies of muscle activation, thereby improving recognition performance for highly similar hand gestures compared with existing CNN-LSTM-based approaches; and (iii) a robust and computationally efficient recognition system validated through extensive experiments, demonstrating its practical applicability for real-time sEMG-based human–machine interaction systems.

2. RELATED WORKS

Early research on surface electromyography (sEMG) for robotic hand control primarily applied classical machine learning with engineered features. Bitzer and van der Smagt (2006) demonstrated that support vector machines (SVMs) could discriminate individual finger actions using a cascade structure; no exact accuracy was reported [21]. Khezri and Jahed (2009) employed wavelet thresholding to extract time–frequency features and intelligent classifiers, showing that multi-resolution representations improve separability, though precise performance metrics were not specified [22]. Ju et al. (2012) applied fuzzy Gaussian mixture models to distinguish a wide set of hand postures; reported accuracy was not explicitly provided [23].

Subsequent studies expanded feature extraction and classification strategies. Sapsanis et al. (2013) applied empirical mode decomposition (EMD) to decompose raw sEMG into intrinsic mode components before classification, reporting an accuracy of approximately 91% for basic gesture discrimination [24]. Kim et al. (2013) integrated tensor neural networks to map sEMG and arm posture to continuous grasping force predictions; performance focused on regression error rather than classification accuracy [25]. Li et al. (2015) highlighted tactile feedback integration for robust prosthetic control; improvements were qualitative [26].

Later research examined diverse classifier families and feature strategies. Kakoty and Hazarika (2016) applied radial basis function SVMs to multiple grasp types; exact accuracy was not reported [27]. X.Tang et al. (2012) proposed a multi-channel sEMG-based hand motion recognition system using a cascaded classifier architecture, achieving approximately 98.2% classification accuracy for 11 hand gestures [28]. Nishad et al. (2018) applied Kraskov entropy features with k-nearest neighbors (k-NN), reporting 92% accuracy [29]. Ensemble machine learning approaches emerged subsequently. Peng et al. (2022) combined ensemble extreme learning machines (EELM) with recursive feature elimination on the Ninapro DB5 dataset (52 gestures), achieving 77.9% mean accuracy, outperforming decision trees, single ELM, and random forests [30].

Deep learning has further reduced reliance on manual feature engineering. Convolutional neural networks (CNNs) automatically learn hierarchical spatial–temporal features, while hybrid CNN–LSTM architectures capture both local and sequential patterns. A multi-stream hybrid architecture integrating CNN, temporal convolutional networks (TCNs), and LSTM modules achieved 94.31% accuracy on Ninapro DB1 and 98.96% on Ninapro DB9, effectively modeling complex temporal and spatial patterns (Miah et al., 2024) [31]. Neural architecture search (NAS) has recently been applied to optimize 1D-CNN models; Gehlot et al. (2025) proposed 1D-CNAS using LSHADE optimization on six gestures from ten subjects, achieving 80.29% mean accuracy and 79.77% F1-score, surpassing a baseline 1D-CNN by ~9.57% [17].

Despite these advances, many methods rely on high-density sensors or significant computation. Therefore, recent research emphasizes low-channel, efficient architectures that maintain high accuracy while enabling real-time deployment. In this study, our proposed hybrid 1D-CNN–LSTM framework leverages only two sEMG channels, integrates CNN spatial encoding with LSTM temporal dynamics, and achieves 96.0% accuracy and 0.96 weighted F1-score, balancing performance, robustness, and embedded deployability for prosthetic and robotic hand control.

3. METHODOLOGY

This section presents the proposed methodology for optimizing robot hand gesture recognition based on surface electromyography (sEMG), including the hardware configuration for data acquisition, signal preprocessing, hybrid time–frequency feature extraction, and the design of a deep learning–based classification model. The overall pipeline is designed to balance recognition accuracy and computational efficiency, enabling reliable real-time hand gesture recognition for robotic hand control applications.

3.1. Data Acquisition

Surface electromyography (sEMG) signals were acquired using a customized hardware setup based on a National Instruments (NI) data acquisition (DAQ) system, as illustrated in Fig. 1. The system was interfaced with a personal computer (PC) for real-time signal monitoring and data storage. Two analog input channels (ai0 and ai1) were configured to record muscle activity from selected forearm muscle groups involved in hand and finger movements.

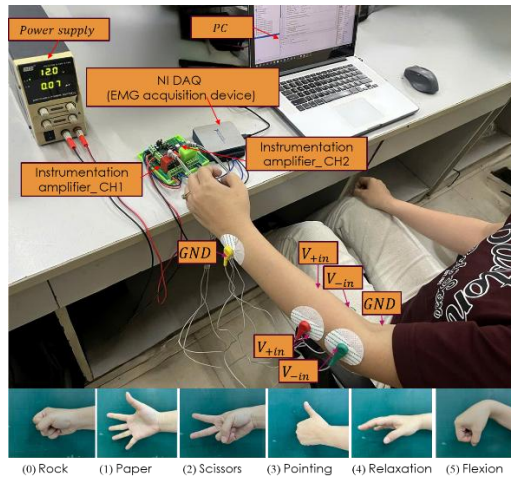


Fig. 1. Experimental hardware setup for two-channel sEMG data acquisition using an NI DAQ system.

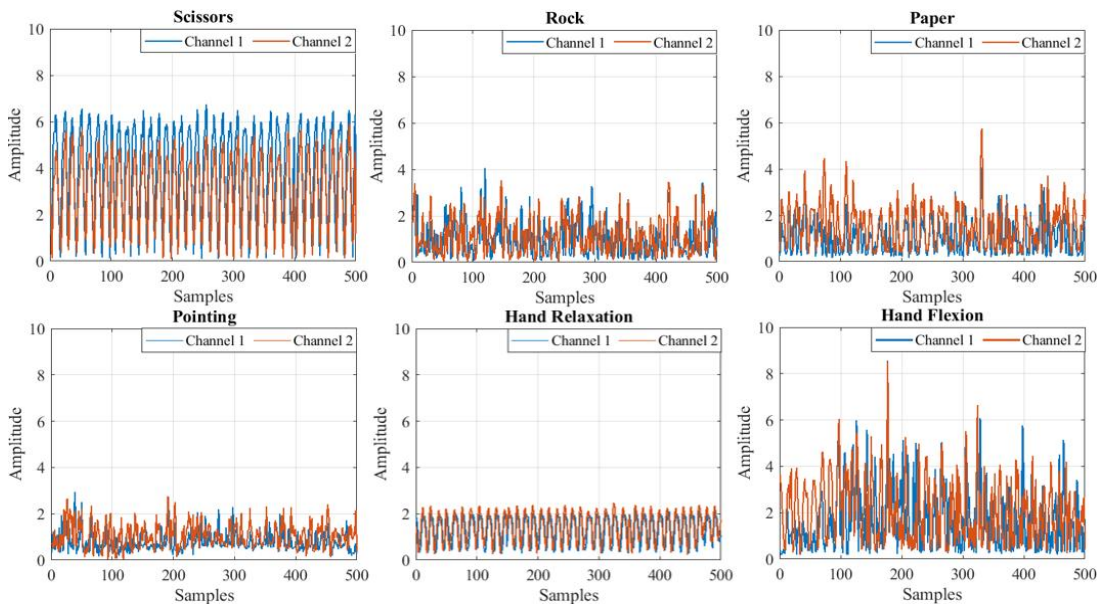


Fig. 2. Representative raw sEMG signals recorded from two channels for six hand gesture

Two-channel sEMG signals were acquired from forearm flexor-extensor muscle groups using bipolar Ag/AgCl electrodes and a common reference electrode. Signals were sampled at 1000 Hz, which sufficiently covers the effective sEMG bandwidth for voluntary contractions while maintaining practical system complexity. Six gestures (Rock, Paper, Scissors, Pointing, Hand Relaxation, and Hand Flexion) were performed, each held for 3 s followed by 3 s of rest, repeated 20 times per class. Representative raw sEMG signals are shown in Fig. 2.classes.

3.2. Signal Preprocessing

Prior to feature extraction, the recorded sEMG signals were preprocessed to enhance signal quality and ensure stable training of the deep learning model. The preprocessing pipeline included the following steps:

- Removal of invalid or saturated samples: Raw sEMG signals were first inspected to discard samples corrupted by acquisition transients, preventing artifacts from affecting subsequent analysis.

- Bandpass filtering: A fourth-order Butterworth filter with cutoff frequencies of 20–450 Hz was applied to suppress low-frequency motion artifacts, baseline drift, and high-frequency electronic noise, preserving the relevant muscle activation components. The discrete-time bandpass filter output $y[n]$ can be expressed as:

$$y[n] = \sum_{k=0}^N b_k x[n-k] - \sum_{k=1}^N a_k y[n-k] \quad (1)$$

Where $x[n]$ and $y[n]$ denote the input and filtered signals, N is the filter order, and a_k, b_k are filter coefficients.

- Notch filtering: A narrowband notch filter at 50 Hz was used to reduce power-line interference.

- Amplitude normalization: Each channel was independently normalized using min-max scaling to map signal values into the $[0, 1]$ range:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

where x is the original signal, x_{min} and x_{max} are the minimum and maximum values of the channel. This step mitigates inter-trial and inter-subject variability and improves numerical stability during gradient-based optimization.

- Segmentation using sliding windows: The normalized signals were partitioned into short-term frames using a sliding window of length $T_w = 200ms$ and a step size $T_s = 100ms$ resulting in 50% overlap between consecutive frames. Each frame X_i is represented as:

$$X_i = [x_i, x_{i+1}, \dots, x_{i+L-1}] \quad (3)$$

Where L is the number of samples per window. Sliding windows act as local receptive fields, capturing short-term patterns in the signal, similar to convolutional operations in deep learning.

3.3. Hybrid Time–Frequency Feature Extraction

For each segmented window, both time-domain and frequency-domain features were extracted to capture complementary characteristics of the sEMG signals.

- Time-domain features (RMS):

The root mean square (RMS) of each frame was computed independently for each channel:

$$RMS = \sqrt{\frac{1}{L} \sum_{n=1}^L x_n^2} \quad (4)$$

Where x_n represents the sample values in a window of length L . RMS provides a robust measure of muscle activation intensity and is computationally efficient.

- Frequency-domain features (FFT):

Spectral features were obtained using the Fast Fourier Transform (FFT):

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi kn/N}, \quad k = 0, 1, \dots, N-1 \quad (5)$$

The magnitude spectrum was computed as:

$$|X[k]| = \sqrt{\Re(X[k])^2 + \Im(X[k])^2} \quad (6)$$

With $N=128$ is the FFT length. Only the first 32 frequency bins per channel were retained to reduce dimensionality. FFT magnitudes from both channels were flattened and concatenated with RMS to form hybrid time–frequency feature vectors for each frame.

3.4. Sequence Construction

To capture short-term gesture dynamics, consecutive frames were grouped into sequences of $F = 6$ frames (approximately 600 ms). Each sequence was labeled according to the class of its final frame, aligning predictions with the current gesture state.

3.5. Model Architecture and Training

The overall architecture of the proposed sEMG-based gesture recognition framework is illustrated in Fig. 3, while a concise layer-wise comparison with a carefully designed baseline model is summarized in Table 1. Two architectures were evaluated to isolate the contribution of hybrid features and temporal modeling.

The baseline model adopts a lightweight yet competitive 1D-CNN operating on RMS features only. It consists of two Conv1D layers with 32 and 64 filters, respectively, followed by max pooling and global average pooling to capture local spatial activation patterns while maintaining low computational complexity. A fully connected SoftMax layer performs final classification.

The proposed model extends this design by jointly exploiting hybrid RMS–FFT features and explicit temporal modeling. As shown in Fig. 3, frame-level feature extraction is performed using Time-Distributed 1D-CNN layers with increased filter capacity (64 and 128 filters), enabling consistent processing across short temporal sequences. L2 regularization is applied to convolutional and dense layers to mitigate overfitting.

To preserve salient activations while reducing feature dimensionality, Time-Distributed global average pooling is employed prior to temporal aggregation. An LSTM layer with 128 hidden units is then introduced to model inter-frame dependencies, which are essential for distinguishing gestures with similar instantaneous sEMG patterns but different temporal dynamics. The classification stage consists of a regularized dense layer followed by a six-unit SoftMax output. Both architectures were trained using the Adam optimizer and sparse categorical cross-entropy loss. Early stopping based on validation performance was applied to ensure stable convergence and fair comparison between models. The detailed architectural differences and hyperparameter settings are reported in Table 1.

Table 1. Layer architecture comparison between the baseline and proposed models

Layer	Layer name	1D-CNN + RMS (baseline)	Hybrid 1D-CNN–LSTM (proposed)
1	1D-CNN	32 filters, kernel 3, ReLU	64 filters, kernel 3, ReLU, $L2 = 1 \times 10^{-4}$
2	1D-CNN_MaxPooling	Pool size 2	Pool size 2 (Time-Distributed)
3	1D-CNN	64 filters, kernel 3, ReLU	128 filters, kernel 3, ReLU, L2
4	Global Avg. Pool	GlobalAveragePooling1D	GlobalAveragePooling1D (Time-Distributed)
5	Temporal modeling	–	LSTM 128 units, dropout = 0.3, recurrent dropout = 0.3
6	Dense (hidden)	256 units, ReLU	128 units, ReLU, $L2 = 1 \times 10^{-4}$
7	Dropout (Regularization)	–	Dropout = 0.4
8	Dense (output)	6 units, SoftMax	6 units, SoftMax

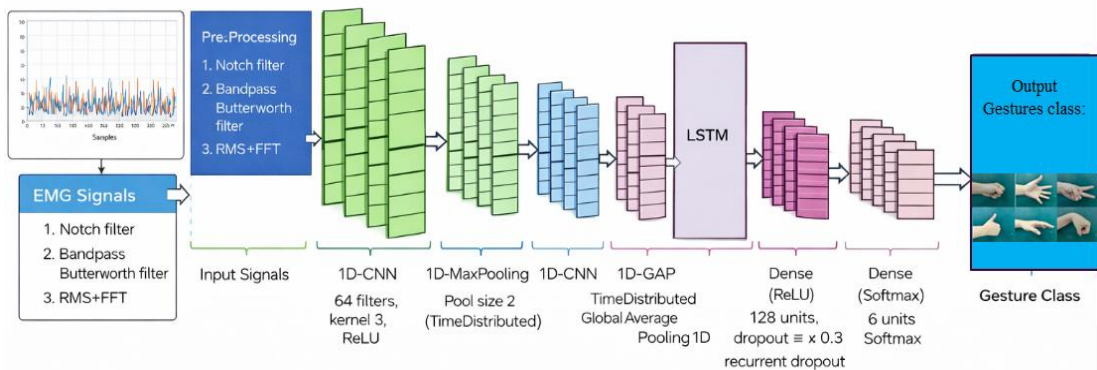


Fig. 3. Architecture of the proposed hybrid 1D-CNN–LSTM framework. RMS and FFT features are extracted at the frame level and processed by Time-Distributed convolutional blocks, followed by global average pooling and LSTM-based temporal modeling for six-class sEMG gesture recognition.

4. EXPERIMENTAL RESULT

This section reports the experimental protocol and the performance of the baseline (1D-CNN+RMS) and proposed (Hybrid 1D-CNN–LSTM) models on the collected same sEMG dataset.

4.1. Evaluation Protocol

After preprocessing and windowing, RMS features (baseline) and hybrid RMS + FFT features (proposed) were extracted as described in Section 3. The hybrid features were further organized into sequences of six frames for the proposed model.

Both models were implemented in TensorFlow/Keras with GPU acceleration. A five-fold stratified cross-validation was used to assess generalization. In each fold, 80% of the sequences were used for training and 20% for validation, preserving class balance. For each architecture, 10 independent runs were performed with different random splits or initializations, and metrics were averaged across runs.

Performance was quantified using overall accuracy, weighted precision, weighted recall, and weighted F1-score. In addition, per-class recognition rates were derived from the confusion matrices.

4.2. Overall Performance

Table 2 summarizes the overall performance of the baseline 1D-CNN+RMS model and the proposed hybrid 1D-CNN+LSTM architecture. Across ten independent runs, the proposed model consistently achieves higher accuracy, precision, recall, and F1-score than the baseline, which is further supported by the training curves and confusion matrices shown in Figs. 4-6. Under the same intra-subject cross-validation protocol, the hybrid model improves mean accuracy by approximately 10 percentage points and increases the weighted F1-score from 0.86 to 0.96, while simultaneously yielding higher precision and recall. These consistent gains across all evaluation metrics indicate that integrating time-domain RMS features with low-frequency Fourier components and explicit temporal modeling via LSTM enhances the separability of sEMG patterns, resulting in a more stable and reliable gesture recognition framework compared with an RMS-based convolutional baseline.

Table 2. Summarizes the average performance of the two architectures over 10 runs.

Model	Accuracy (%)	Precision (weighted)	Recall (weighted)	F1-score (weighted)
1D-CNN + RMS (baseline)	86	0.88	0.86	0.86
Hybrid 1D-CNN + LSTM (proposed)	96	0.97	0.96	0.96

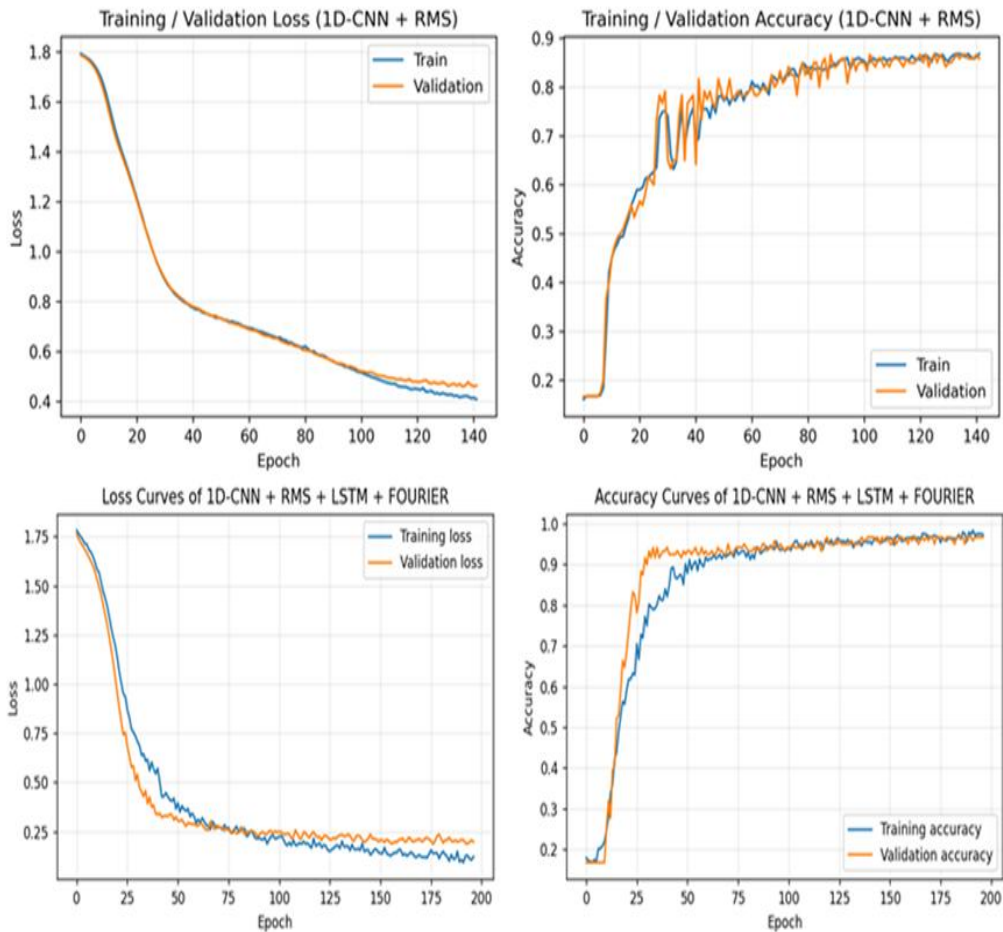


Fig. 4. Accuracy and loss curves comparing baseline 1D-CNN+RMS and hybrid 1D-CNN-LSTM model

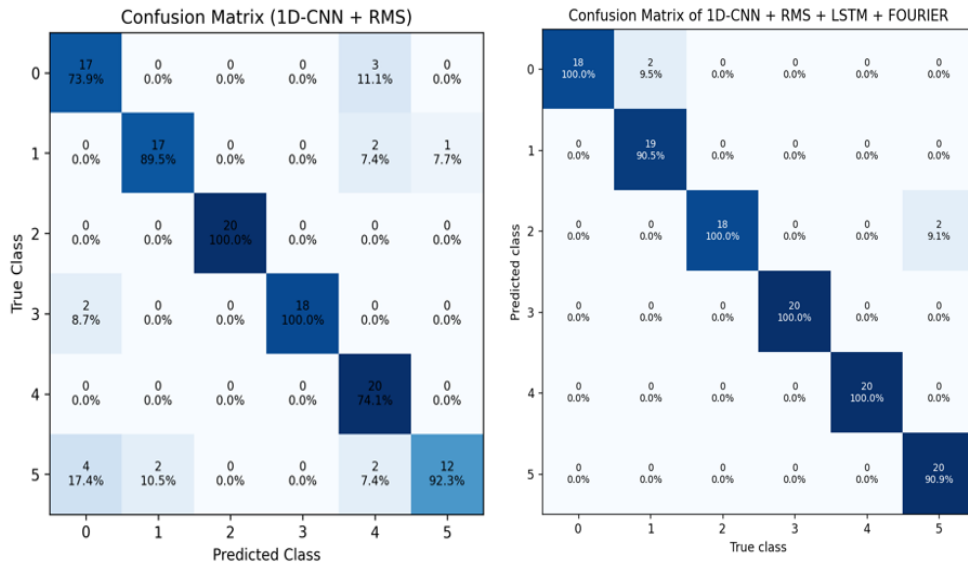


Fig. 5. Confusion Matrix comparing baseline 1D-CNN+RMS and proposed hybrid 1D-CNN-LSTM model

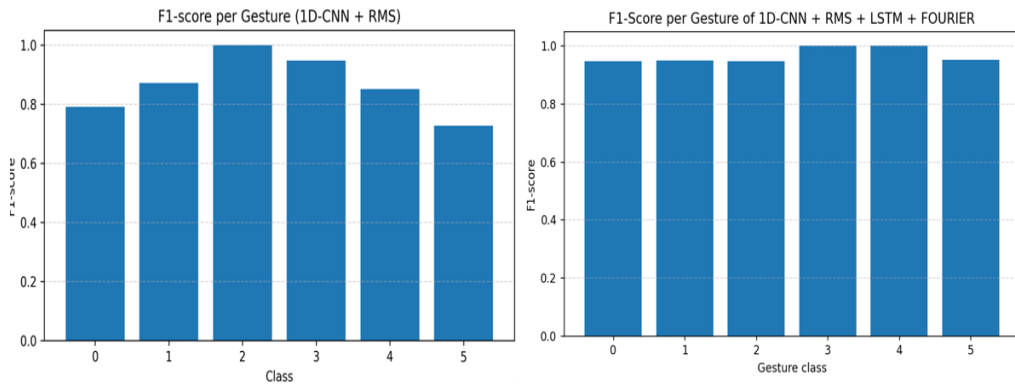


Fig. 6. F1-score comparing baseline 1D-CNN+RMS and proposed hybrid 1D-CNN-LSTM model

4.3. Gesture-Wise Recognition Rates

Gesture-wise performance of the proposed hybrid model was analyzed using the normalized confusion matrix, which exhibits clear diagonal dominance across all six gesture classes. The corresponding gesture-wise recognition rates of both models are summarized in Table 3.

The proposed architecture achieves recognition rates exceeding 90% for all gesture classes under the intra-subject evaluation protocol, indicating stable and consistent class-level generalization. Compared with the RMS-based baseline, notable improvements are observed for Rock, Pointing, and Hand Flexion, where the baseline model shows lower and more variable recognition performance.

Residual misclassifications are mainly observed between gestures with overlapping muscle activation patterns, particularly Paper versus Scissors and Pointing versus Hand Flexion. These confusions are consistent with the similarity of their low-frequency sEMG characteristics and reflect the intrinsic difficulty of discriminating spectrally and temporally similar gestures. Nevertheless, the integration of hybrid time–frequency features with explicit

temporal modeling effectively reduces such ambiguities, resulting in more uniform and reliable gesture-wise performance compared with the baseline architecture.

While the proposed framework is implemented using a 1D-CNN-based architecture combined with LSTM for temporal modeling, the underlying hybrid feature representation and optimization strategy are not restricted to this specific network design. Extending the proposed approach to alternative deep learning architectures to further enhance model flexibility and recognition performance will be explored in future work.

Table 3. Per-class recognition rates of baseline and proposed model

Gesture Class	Gesture Name	1D-CNN + RMS (baseline) Recognition Rate (%)	Hybrid 1D-CNN + LSTM (proposed) Recognition Rate (%)
0	Rock	85	96.2
1	Paper	85	95.8
2	Scissors	100	93.1
3	Pointing	85.0-90.0	95.4
4	Hand Relaxation	100	96.5
5	Hand Flexion	55.0-70.0	92.7

5. CONCLUSION AND FUTURE WORK

This study provides a compact and effective framework for sEMG-based hand gesture recognition that integrates hybrid time–frequency feature representation with explicit temporal modeling. By combining RMS features, low-frequency Fourier components, and an LSTM-based temporal module, the proposed approach consistently outperforms a lightweight RMS-based 1D-CNN baseline under an intra-subject evaluation protocol. Experimental results demonstrate clear improvements in accuracy, weighted precision, recall, and F1-score, together with more uniform gesture-wise recognition performance.

The results indicate that hybrid feature fusion and temporal sequence modeling provide complementary benefits for enhancing the discriminability and robustness of sEMG representations, even when using a limited number of channels. This balance between recognition performance and model complexity supports the suitability of the proposed framework for practical wearable and human–machine interaction systems.

Several limitations remain. The current study focuses on offline analysis and intra-subject evaluation, which may not fully capture inter-subject variability and real-time deployment constraints. Moreover, although the present implementation relies on a 1D-CNN and LSTM architecture, the underlying hybrid representation and optimization strategy are not restricted to this design. Future work will therefore explore cross-subject generalization, continuous gesture recognition, and extensions to alternative deep learning architectures and online learning scenarios.

Acknowledgment: This work belongs to the project grant No: T2025-214 funded by Ho Chi Minh City University of Technology and Education, Vietnam. This work was conducted in the Artificial Intelligence for Automation and Control Laboratory (AIAC Lab D203A) with support from Ho Chi Minh City University of Technology and Engineering, Viet Nam.

REFERENCES

- [1] X. Zhao, J. Huang, J. Zheng, Y. Ma, and H. Tang, "A multimodal-signals-based gesture recognition method for human machine interaction," in *2020 3rd International Conference on Unmanned Systems (ICUS)*, 2020: IEEE, pp. 494-499. DOI: <http://doi.org/10.1109/ICUS50048.2020.9274853>
- [2] Q. Zhang, T. Pi, R. Liu, and C. Xiong, "Simultaneous and proportional estimation of multijoint kinematics from EMG signals for myocontrol of robotic hands," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 4, pp. 1953-1960, 2020. DOI: <http://doi.org/10.1109/TMECH.2020.2999532>
- [3] D. Yang and H. Liu, "An EMG-based deep learning approach for multi-DOF wrist movement decoding," *IEEE Transactions on Industrial Electronics*, vol. 69, no. 7, pp. 7099-7108, 2021. DOI: <http://doi.org/10.1109/TIE.2021.3097666>
- [4] N. Tsujiuchi, K. Takayuki, and M. Yoneda, "Manipulation of a robot by EMG signals using linear multiple regression model," in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, 2004, vol. 2: IEEE, pp. 1991-1996. DOI: <http://doi.org/10.1109/IROS.2004.1389690>
- [5] H.-P. Huang and C.-Y. Chen, "Development of a myoelectric discrimination system for a multi-degree prosthetic hand," in *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, 1999, vol. 3: IEEE, pp. 2392-2397. DOI: <http://doi.org/10.1109/ROBOT.1999.770463>
- [6] C. Xie, Q. Yang, Y. Huang, S. W. Su, T. Xu, and R. Song, "A hybrid arm-hand rehabilitation robot with EMG-based admittance controller," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 15, no. 6, pp. 1332-1342, 2021. DOI: <http://doi.org/10.1109/TBCAS.2021.3130090>
- [7] M. Soto-Florido *et al.*, "Identification of Neuromotor Patterns Associated with Hand Rehabilitation Movements from EMG and EEG Signals," in *2025 International Conference On Rehabilitation Robotics (ICORR)*, 2025: IEEE, pp. 1307-1312. DOI: <http://doi.org/10.1109/ICORR66766.2025.11063054>
- [8] D. K. Singh, A. P. Srivastava, S. Singh, S. Paliwal, A. Singh, and J. Singh, "Machine learning based Development and Evaluation of a Hand Exoskeleton Rehabilitation Robot for Patients with Partial Paralysis," in *2024 2nd International Conference on Disruptive Technologies (ICDT)*, 2024: IEEE, pp. 418-423. DOI: <http://doi.org/10.1109/ICDT61202.2024.10489600>
- [9] N. Sanwlot, K. V. Raj, G. Udupa, and R. Anand, "Cost-Effective EMG Signal Acquisition for Rehabilitation Robotics using Single-Channel Sensors and Machine Learning," in *2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, 2024: IEEE, pp. 1603-1609. DOI: <http://doi.org/10.1109/ICSSAS64001.2024.10761024>
- [10] D. Farina *et al.*, "Man/machine interface based on the discharge timings of spinal motor neurons after targeted muscle reinnervation," *Nature biomedical engineering*, vol. 1, no. 2, p. 0025, 2017. DOI: <http://doi.org/10.1038/s41551-016-0025>
- [11] M. H. Zafar, S. K. R. Moosavi, and F. Sanfilippo, "Federated learning-enhanced edge deep learning model for EMG-based gesture recognition in real-time human-robot interaction," *IEEE Sensors Journal*, 2025. DOI: <http://doi.org/10.1109/JSEN.2025.3529841>
- [12] A. A. Neacsu, G. Cioroiu, A. Radoi, and C. Burileanu, "Automatic EMG-based hand gesture recognition system using time-domain descriptors and fully-connected neural networks," in

- 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), 2019: IEEE, pp. 232-235. DOI: <http://doi.org/10.1109/TSP.2019.8768831>
- [13] N. Trivedi, I. Naik, N. Fadia, D. Karia, and N. Ghatte, "Universal Remote using EMG based Gesture Recognition," in *2021 IEEE Bombay Section Signature Conference (IBSSC)*, 2021: IEEE, pp. 1-5. DOI: <http://doi.org/10.1109/IBSSC53889.2021.9673376>
- [14] P. Kang, S. Jiang, and B. He, "Diffusion-Driven Deep Decoding: Advancing EMG-Based Hand Skill Learning for Environment-Free Human-Robot Interaction," *IEEE/ASME Transactions on Mechatronics*, 2025. DOI: <http://doi.org/10.1109/TMECH.2025.3546753>
- [15] M. Atzori, M. Cognolato, and H. Müller, "Deep learning with convolutional neural networks applied to electromyography data: A resource for the classification of movements for prosthetic hands," *Frontiers in neurorobotics*, vol. 10, p. 9, 2016. DOI: <http://doi.org/10.3389/fnbot.2016.00009>
- [16] K. Bakircioğlu and N. Özkurt, "Classification of EMG signals using convolution neural network," *International Journal of Applied Mathematics Electronics and Computers*, vol. 8, no. 4, pp. 115-119, 2020. DOI: <http://doi.org/10.18100/ijamec.795227>
- [17] N. Gehlot, R. Kumar, S. Hans, and P. Nkomozepi, "1D Convolutional Neural Architecture Search for sEMG Hand Gesture Recognition," *SN Computer Science*, vol. 6, no. 7, p. 789, 2025. DOI: <http://doi.org/10.1007/s42979-025-04324-3>
- [18] M. Coskun, O. Yildirim, Y. Demir, and U. R. Acharya, "Efficient deep neural network model for classification of grasp types using sEMG signals," *Journal of ambient intelligence and humanized computing*, vol. 13, no. 9, pp. 4437-4450, 2022. DOI: <http://doi.org/10.1007/s12652-021-03284-9>
- [19] N. Tsagkas, P. Tsinganos, and A. Skodras, "On the use of deeper CNNs in hand gesture recognition based on sEMG signals," in *2019 10th international conference on information, intelligence, systems and applications (IISA)*, 2019: IEEE, pp. 1-4. DOI: <http://doi.org/10.1109/IISA.2019.8900709>
- [20] S. Briouza, H. Gritli, N. Khraief, S. Belghith, and D. Singh, "A convolutional neural network-based architecture for EMG signal classification," in *2021 International conference on data analytics for business and industry (ICDABI)*, 2021: IEEE, pp. 107-112. DOI: <http://doi.org/10.1109/ICDABI53623.2021.9655876>
- [21] S. Bitzer and P. Van Der Smagt, "Learning EMG control of a robotic hand: towards active prostheses," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, 2006: IEEE, pp. 2819-2823. DOI: <http://doi.org/10.1109/ROBOT.2006.1642128>
- [22] M. Khezri and M. Jahed, "Surface electromyogram signal estimation based on wavelet thresholding technique," in *2008 30th annual international conference of the IEEE engineering in medicine and biology society*, 2008: IEEE, pp. 4752-4755. DOI: <http://doi.org/10.1109/IEMBS.2008.4650275>
- [23] G. Ouyang, X. Zhu, Z. Ju, and H. Liu, "Dynamical characteristics of surface EMG signals of hand grasps via recurrence plot," *IEEE journal of biomedical and health informatics*, vol. 18, no. 1, pp. 257-265, 2013. DOI: <http://doi.org/10.1109/JBHI.2013.2261311>
- [24] C. Sapsanis, G. Georgoulas, A. Tzes, and D. Lymberopoulos, "Improving EMG based classification of basic hand movements using EMD," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013: IEEE, pp. 5754-5757. DOI: <http://doi.org/10.1109/EMBC.2013.6610858>

- [25] S. Kim, J. Kim, M. Kim, S. Kim, and J. Park, "Grasping force estimation by sEMG signals and arm posture: tensor decomposition approach," *Journal of bionic engineering*, vol. 16, no. 3, pp. 455-467, 2019. DOI: <http://doi.org/10.1007/s42235-019-0037-0>
- [26] C. Wu, A. Song, Y. Ling, N. Wang, and L. Tian, "A control strategy with tactile perception feedback for EMG prosthetic hand," *Journal of Sensors*, vol. 2015, no. 1, p. 869175, 2015. DOI: <http://doi.org/10.1155/2015/869175>
- [27] N. M. Kakoty and S. M. Hazarika, "Recognition of grasp types through principal components of DWT based EMG features," in *2011 IEEE International Conference on Rehabilitation Robotics*, 2011: IEEE, pp. 1-6. DOI: <http://doi.org/10.1109/ICORR.2011.5975398>
- [28] X. Tang, Y. Liu, C. Lv, and D. Sun, "Hand Motion Classification Using a Multi-Channel Surface Electromyography Sensor," *Sensors*, vol. 12, no. 2, pp. 1130–1147, 2012. DOI: <http://doi.org/10.3390/s120201130>
- [29] S. Taran and V. Bajaj, "Motor imagery tasks-based EEG signals classification using tunable-Q wavelet transform," *Neural Computing and Applications*, vol. 31, no. 11, pp. 6925-6932, 2019. DOI: <http://doi.org/10.1007/s00521-018-3531-0>
- [30] F. Peng *et al.*, "Gesture recognition by ensemble extreme learning machine based on surface electromyography signals," *Frontiers in human neuroscience*, vol. 16, p. 911204, 2022. DOI: <http://doi.org/10.3389/fnhum.2022.911204>
- [31] J. Shin, A. S. M. Miah, S. Konnai, I. Takahashi, and K. Hirooka, "Hand gesture recognition using sEMG signals with a multi-stream time-varying feature enhancement approach," *Scientific Reports*, vol. 14, no. 1, p. 22061, 2024. DOI: <http://doi.org/10.1038/s41598-024-72996-7>