Int. J. Advance Soft Compu. Appl, Vol. 18, No. 1, March 2026
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Comparative Study of LSTM, GRU, and BRNN Performance: large-Scale Data Analytics (EMG Signal Classification)

Karim M. Aljebory[1], Yashar M. Jwmah[2], Thabit Sultan Mohammed[3,*], Adnan Saif Al Mamari[4]

[1]Computer Technical Engineering Dept., Al-Qalam University, 36001 Kirkuk, Iraq.
email: karim.eng@alqalam.edu.iq.

[2]Kirkuk Health Directorate, Iraqi Ministry of Health, 36001 Kirkuk, Iraq,
email: agronabaat@gmail.com

[3,*]Computer Technical Engineering Dept., Al-Qalam University, 36001 Kirkuk, Iraq.
email: thabit.sultan@alqalam.edu.iq ; thabit.tsm@gmail.com

[4]Computing science Dept., Modern College of Business and Science, 3 Bawshar St, Muscat 133, Oman.
email: Assa.85@hotmail.com

*Corresponding Author: Thabit Sultan Mohammed

Abstract

Electromyography (EMG) signals play a pivotal role in biomedical applications, such as prosthetic control and human-computer interaction, where advanced classification methods are essential for accurately translating muscle activity. This study evaluates the performance of three neural network architectures: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Bidirectional Recurrent Neural Network (BRNN) for EMG signal classification. The EMG signals were preprocessed using Digital Wavelet Transform (DWT) with Daubechies 2 wavelet to extract time-frequency features. Experiments were conducted on a large-scale training dataset comprising 672 subject recordings across six hand gestures, enabling a robust, data-driven comparison. The highest classification accuracy of LSTM, GRU, and Bidirectional RNN was achieved, corresponding to cD7, with values of 93.04±1.52, 92.72±1.26, and 91.59±0.97, respectively. However, these models exhibited varying degrees of sensitivity to additive noise, particularly at deeper DWT levels. The findings highlight the trade-offs between accuracy and noise tolerance, providing insights for optimizing EMG-based gesture recognition systems in real-world applications. The final analysis confirms that LSTM outperforms the other models for real-time EMG classification, while GRU and BRNN offer a favorable balance between accuracy and computational efficiency. The effective handling of large-scale, high-dimensional EMG data can yield significant performance improvements, particularly for prosthetic and real-time control applications.

Keywords: *Electromyography (EMG), Signal classification, DWT, Deep learning in large-scale data analytics, hand gesture, ANN classifier, (LSTM, GRU, BRNN).*

Received 9November2025; Accepted 1February 2026

1 Introduction

Electromyography (EMG) signals play an important role in biomedical applications, including prosthetic limb control, rehabilitation robotics, and human-computer interaction. Substantial development is required to enhance the intelligence of control methods in these systems [1,2]. Accurate and efficient classification of EMG signals is crucial for translating muscle activity into meaningful commands for assistive devices. Neural Networks (NNs) play a main role in recognition and classification tasks; several neural network architectures have shown effectiveness. Some NN models show their effectiveness in EMG signal classification [3–6].

Convolutional Neural Networks (CNNs) are suited for extracting spatial features from EMG signals by treating the one immersion sequence (time series data) to capture both local and global patterns in muscle activity. They may require larger datasets for training and can be sensitive to variations in signal quality.

Recurrent Neural Networks (RNNs), such as Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), and Bidirectional Recurrent Neural Network (BRNN), are suitable for capturing temporal dependencies in EMG signal data. They can effectively model the sequential nature of muscle activity patterns over time. It is particularly useful for EMG signal classification tasks where both spatial and temporal information are crucial. A Bi-directional Recurrent Neural Network (BRNN) has two distinct recurrent hidden layers, one processes the input sequence forward, and the other processes it backward. The results from hidden layers are collected and input into a prediction-making final layer, and any RNN cell, such as LSTM or GRU, can be used to create the recurrent hidden layers.

The Recurrent Neural Networks (RNNs) are designed for sequential high-dimensional dataset classification; they contain recurrent connections, which maintain a hidden state that evolves [7]. On the other side, CNNs use convolutional layers to capture local data patterns so that they can be applied for feature extraction from spatial data [8,9]. The EMG signals are raw time series, which can be directly processed in one-dimensional layers and sometimes converted into two-dimensional form to capture time-frequency features by Wavelet Transform, Short-Time Fourier Transform, and EMG images. Combining CNNs and RNNs by incorporating convolutional layers to extract spatial features, followed by recurrent layers to capture temporal dynamics, may strengthen the NN performance. The primary challenge is architectural complexity, which demands greater computational resources and careful hyperparameter tuning [10]. In the research presented by Aljebory, et al. [4], a hand movement's surface EMG signal recognition scheme was used with 36 subjects, applying digital wavelet transform (DWT) with 8 levels deep and an NN classifier, in which an accuracy of 89.9% was reached. The selection of the best NN for classification depends on the structure of the network and other factors such as the nature of the signals, the size of the dataset, and computational resources. This paper evaluates the performance of LSTM, GRU, and BRNN in the context of EMG signal classification. The time-frequency features for the sEMG signal are generated by DWT and injected into the NN input layer. By comparing these advanced architectures, we will establish their relative strengths and limitations, contributing to the broader application of NN models in enhancing EMG-based gesture recognition systems. This work aims to identify the most effective NN model for robust, real-time classification in large-scale sequential EMG signal data analysis. The dataset is substantial, comprising 672 subject recordings, and expands significantly when decomposed into 8 channels × 6 gestures × time-series samples × high-dimensional DWT components.

2 Related Work

Neural Networks (NNs) are used for decision-making and classification by processing data through interconnected layers of artificial neurons, learning patterns from training data to categorize new inputs. A vast amount of published literature presents studies and various applications, which are considered in this regard. In [11,12], case studies for the use of neural networks in decision-making and clustering are presented. The learning process enables neural networks to make predictions and decisions in complex tasks like image recognition, signal processing, natural language processing, and recommendation systems. For a certain application, different NN architectures may vary in their performance.

In the study presented by Ali Raza Asif et al. [5], researchers implemented a CNN to decode hand gestures from surface EMG data recorded from 18 subjects. They investigated the effect of hyperparameters on each hand gesture. The key findings are that when the learning rate was set to either 0.0001- 0.001, 80-100 epochs significantly outperformed other considerations. Robust and stable myoelectric control based on the best-performing hand motions achieved. Certain motions consistently performed better throughout the study. Geng et al. [11] proposed an attention-based hybrid CNN-RNN architecture that outperformed standalone CNN and RNN models. They used a new sEMG image feature vector for improved performance. Another study performed by [6] presented a dilated CNN-based approach for EMG classification. It captured both local and global features from EMG signals and achieved 99% accuracy on a dataset of six hand gestures. The researchers in [12] proposed an LSTM-based RNN model for real-time hand gesture classification based on pre-processed EMG signals. The LSTM configuration effectively identified temporal patterns in time series data; they reached significant values of deep BP–LSTM network 92% accuracy, 89% specificity, 91% precision, and 96% F1-score, in the multi-classification of the sEMG signal. Toro-Ossaba et al. [3] implemented an RNN model using LSTM units and dense layers for hand gesture recognition. Their proposed model required only 4 EMG channels to recognize 5 hand gestures. The model achieved an accuracy of up to 87 ± 7% during real-time testing. Deep Learning (DL) techniques have made their mark in EMG-based hand gesture recognition, where DL models learn relevant features directly from raw data.

In literature, the DWT features are considered for EMG signal classification, which provides a multi-resolution time-frequency decomposition that matches the transient nonstationary nature. In [13], the authors evaluated Daubechies wavelet families (db1-db6) among other wavelets for EMG signal classification and deduced a comparative performance. William L et al [14] reported that Daubechies-2 (Db2) have compact time support and a robust choice against the uncertainty of the time–frequency region definition.

3 Methodology

The scale and complexity of the EMG dataset necessitate models capable of learning from high-dimensional sequential data. To this end, we evaluate three recurrent architectures. The effectiveness of the proposed system is considered from two aspects: the EMG nature of the signal and the neural network performance in handling the classification task.

3.1 Signal Consideration

The selection of Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and the Bi-directional Recurrent Neural Network (BRNN) for EMG signal classification lies in the distinct way they process and handle the complexities of EMG signals, which are noisy, non-stationary, and a large-scale time-dependent dataset. A closer examination reveals why these models perform well:

EMG Signals are Noisy: Since the sEMG surface electrodes are prone to picking up noise from skin, movement artifacts, and cross-talk between muscles. LSTM, GRU, and BRNN are capable of extracting meaningful features even from noisy sEMG signals.

Time-Varying Nature: EMG data represent muscle activity over time. LSTMs designed for sequential data excel in capturing the dynamics of muscle activations over time, which is crucial for real-time applications like prosthetic control.

Multi-channel Data: EMG signals are often gathered from multiple electrodes, creating a multi-channel time-series dataset. RNNs can process these multi-channel inputs by learning local and global dependencies across these data elements.

3.2 Neural Network Consideration

In the literature, the LSTM, GRU, and BRNN models are often considered as one of the best neural networks for EMG (Electromyography) signal classification. Their ranking is based on how they process and handle the physical nature of EMG signals, which are noisy, non-stationary, and time-dependent [15–18]. A deeper look at why these models perform well in EMG signal classification is stated below;

3.2.1 Long Short-Term Memory Networks (LSTMs)

The LSTM performs well for EMG signals time-series data, where muscle activations unfold over time. LSTMs are designed to capture long-term dependencies in sequential data, making them ideal for detecting patterns across multiple time steps. EMG signals often involve complex, time-varying patterns (like muscle contractions and relaxations), which require the network to remember information over long periods. LSTMs, with their memory cells, can manage long-range dependencies better than standard RNNs. Unlike traditional RNNs, LSTMs mitigate the vanishing gradient problem, allowing them to effectively learn from sequences of varying lengths, which is a common feature in EMG signals. It can handle noisy signals well because it focuses on localized, meaningful features, making them highly effective when EMG data is recorded in non-ideal conditions, such as with surface electrodes. It has fewer parameters than fully connected networks for the same level of complexity, which speeds up training and reduces overfitting on small datasets, which is often the case with EMG signals [3,15,16,19].

3.2.2 Gated Recurrent Unit (GRU)

The GRUs are a simplified version of LSTMs that combine the input and forget gates into a single update gate. This reduction in complexity allows GRUs to be computationally more efficient while still capturing dependencies in sequential data. GRUs have been shown to perform comparably to LSTMs in various tasks, such as natural language processing and time series forecasting, but with fewer parameters, making them faster to train. Their architecture consists of two gates: the update gate and the reset gate, which together manage the information flow similarly to LSTMs but with a more streamlined approach. GRU typically has fewer parameters than LSTM, which may lead to faster convergence and less computational cost, useful for real-time EMG classification where speed is important. It may work as well as LSTM for EMG signal classification in terms of performance, especially if the data doesn't require long-term memory [16–19].

3.2.3 Bidirectional Recurrent Neural Network (BRNN)

The BRNN processes sequences in both directions (forward and backward), so it captures both past and future context. In EMG classification, this is beneficial because muscle signals might contain information from both temporal directions. Since it involves two layers (forward and backward), it requires more parameters than unidirectional RNNs (LSTM and GRU), which could make training more computationally expensive. It is useful for cases where future muscle activity can influence the classification of current states (e.g., in motion prediction or gesture recognition). BRNNs enhance the capabilities of standard RNNs by processing data in both forward and backward directions. This dual processing allows the network to capture context from both past and future states, which is particularly beneficial in tasks where the context is crucial, such as in language translation and sentiment analysis. Bidirectional RNNs can be implemented using either LSTM or GRU cells, thereby inheriting the advantages of these architectures while also gaining the ability to consider the entire sequence of data [20–22].

3.3 Model Architecture and Hyperparameter Tuning

To ensure optimal performance of the LSTM, GRU, and BRNN models, a systematic hyperparameter tuning process is conducted. The architecture of each network is tested and adjusted based on validation performance, with the following parameters explored:

Dropout rate: $\leq$0.5 (to prevent overfitting)
Learning rate: $\geq$ 0.001 (to insure efficient convergence)
Optimizer: Adam, RMSprop (for their adaptive learning capabilities)
Batch size: $\geq$10 (varied to balance gradient estimation stability)
Max Epochs: 30 (with early stopping implemented)

A grid search approach is used to evaluate combinations of these parameters. For clarity the final adopted hyperparameter tuning parameters are summarized in Table 1.

Table 1: Hyperparameter tuning summary for LSTM, GRU, and BRNN models

Hyperparameter	**Search Space**	**Selected (LSTM)**	**Selected (GRU)**	**Selected (BRNN)**
Recurrent Layers	1, 2, 3	2	2	2
Hidden Units	50, 100, 150, 200	100	100	100
Dropout Rate	0.2, 0.3, 0.5	0.3	0.2	0.3

Hyperparameter	Search Space	Selected (LSTM)	Selected (GRU)	Selected (BRNN)
Learning Rate	0.0001, 0.0005, 0.001	0.0005	0.0005	0.0005
Optimizer	Adam, RMSprop	Adam	Adam	Adam
Batch Size	10, 20, 32	10	10	10
Max Epochs	20, 25, 30	25	25	25

The recurrent hidden layer 1 acts as a strictly non-linear filter. It learns to ignore the random electrical noise, and "fire" only on actual muscle activation patterns. Layer 2, on the other hand, performs the actual temporal classification based on the filtered output of layer 1. Adding more layers to increase accuracy, usually fails, due to two reasons: first, overfitting to noise since EMG signals contain a lot of random background noise (from the skin, electronics, or other muscles), and second, Vanishing Gradient where, stacking many layers makes it harder for the gradient to propagate back, making training unstable.

4 EMG Signal Processing and Conditioning

This section concerns the description of the basic processing steps applied to the sEMG signal before testing the NNs' performance. Initially, the raw data sEMG signal is converted to a zero-mean signal, then a rectification step is performed in which the absolute value is applied for further processing of the EMG signal. An Absolute Value Moving Average Filter (AVMAF) was used to reduce the noise level associated with the signal that affects the construction of the envelope shape [23]. Finally, the signal was normalized between a minimum and a maximum value. The data produced from preprocessing is transformed into Digital Wavelet Transform (DWT) components using the eight-level Daubechies 2 algorithm (cD1-cD8). The Daubechies 2 (db2) wavelet transform is suitable for sEMG processing and classification, since sEMG signals typically contain a variety of frequency components, including low-frequency components and high-frequency components [4,24,25] . The DWT components are considered crucial features for the classification step, which makes the final decision by adopting an NN-based classifier [4,24,25]. The basic preprocessing steps in sequence can be summarized by the block diagram shown in Fig.1, and the output of the last step is organized as a data cell array matrix that is suitable for NN input.

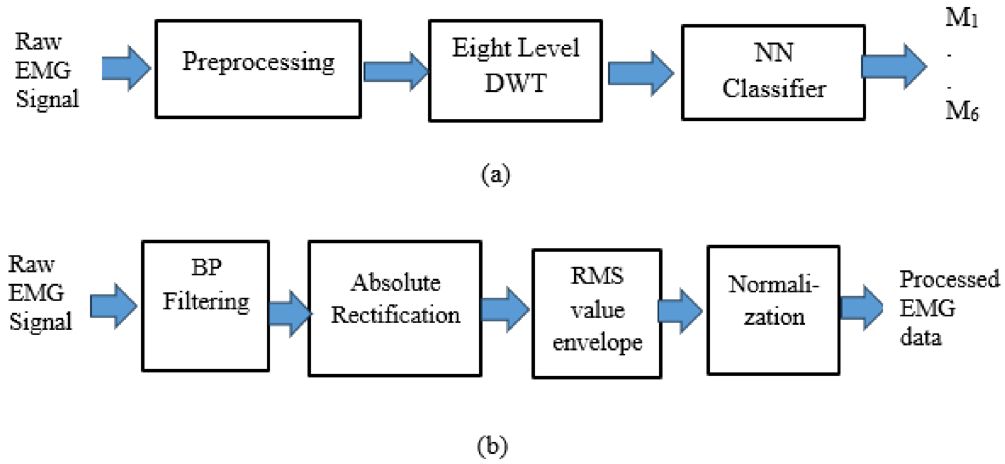


Fig.1 (a) Main Block Diagram, (b) Signal Preprocessing. [Authors' own work]

4.1 Data set source

The sEMG dataset used in this study was acquired from the Physiobank repository at the UCI Machine Learning Repository. It was made publicly available online and used for classification experiments in the research works [4,26]. The data source includes recordings from 36 individuals, originated from the forearm muscles by eight channels that correspond to six distinct motion classes, denoted as: M1- hand at rest, M2- hand clenched in a fist, M3- wrist flexion, M4- wrist extension, M5- radial deviations, M6- ulnar deviations. The sample time-domain raw sEMG data for motion M1 recorded from channel 1 is shown in Fig. 2 below.

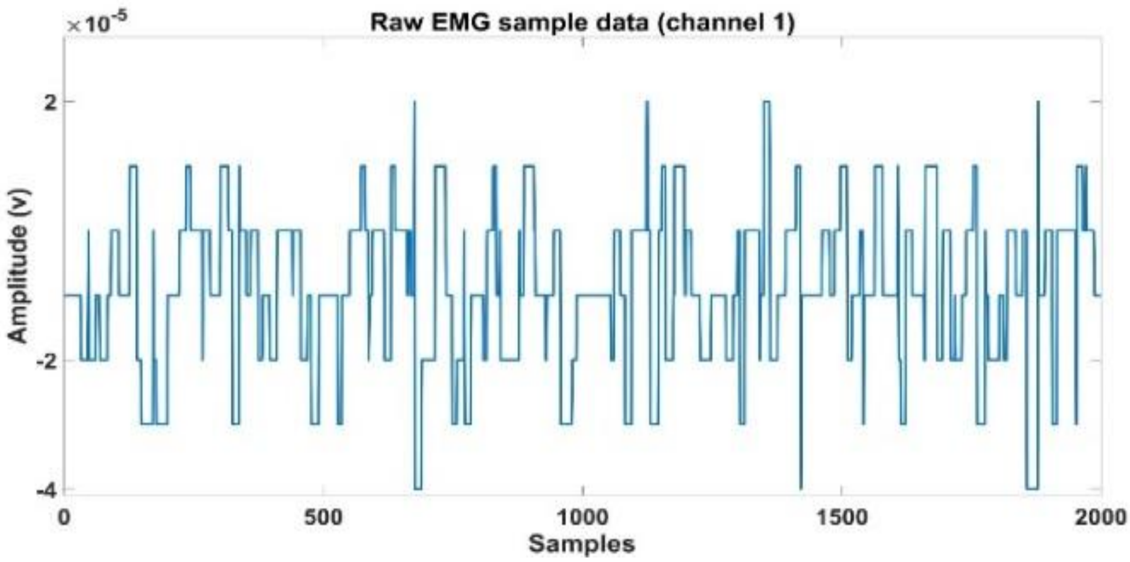


Fig.2 Sample Raw sEMG Signal (Channel 1). [Authors' own work]

4.2 Data Preprocessing

After acquisition, the EMG data were preprocessed for input into the proposed classification model. The preprocessing is essential because the raw EMG signals are random and need to be extracted into valuable information before being used as input data to a classifier. However, research on EMG signal processing showed that the preprocessing step mainly focused on smoothing the signal and removing the noise corrupting it, and leaving the feature extraction for the proposed neural model [27,28]. T The fundamental features of sEMG signals are embedded in their time-domain, analog, and statistical characteristics. So, the root mean square value of each channel (for channel 1 as shown in Fig. 3) is considered.

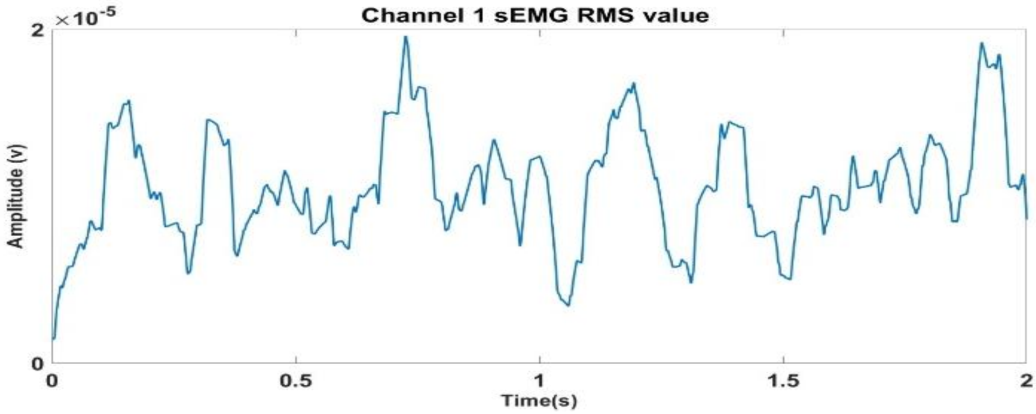


Fig.3 The RMS Value for Channel 1. [Authors' own work]

Meanwhile, the frequency spectrum is fundamental for further processing steps, where the corresponding Fast Fourier Transform (FFT) of the EMG signals is evaluated. Using the evaluated FFT spectra in Fig. 4, the required filtering process can be defined, and the effective filter is selected and implemented.

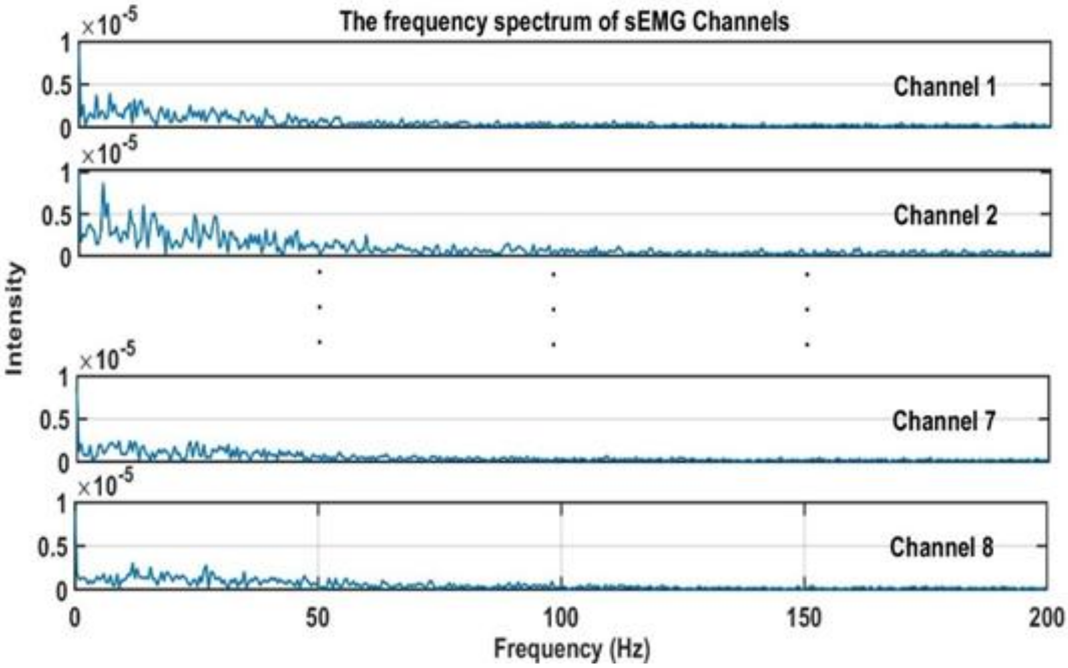


Fig.4 The FFT of sEMG Channels 1, 2, 7, and 8. [Authors' own work]

Based on the above FFT, the band-pass filter is selected and justified in the range 0-100 Hz, which covers the main signal components. The filtered sEMG signal is then injected into the data conditioning and justification process steps, including absolute value rectification and normalization as depicted in Fig. 5 (a) and (b) sequentially.

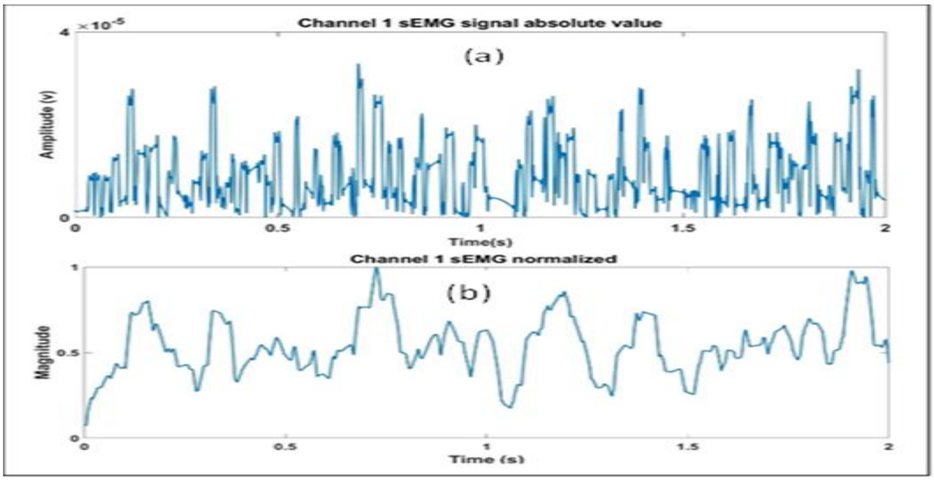


Fig.5 sEMG signal (a) absolute value, (b) normalized. [Authors' own work]

The last step before NN classification, the DWT features extraction operation, is handled by applying the eight-level Daubechies-2 algorithm that generates the detailed eight coefficients defined by the components (cD1-cD8). Figure 6 shows a sample DWT output channel 1 for motion M1. Each level in DWT represents an extracted feature of the sEMG signal, which will be injected as input data into the NN input layer.

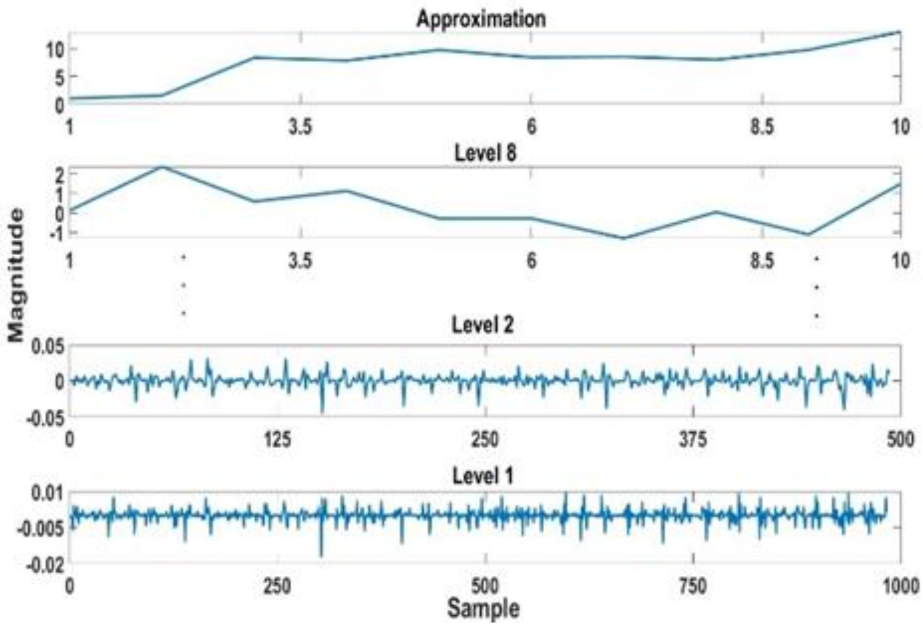


Fig.6 Sample DWT for Channel 1 Signal. [Authors' own work]

The role of DWT decomposition level is critical in shaping neural network performance for all dependent variables, whereas the impact of higher levels significantly affects the EMG signal classification accuracy. Many studies find that moderate DWT levels maintain favorable accuracy in EMG signal classification [29,30].

4.3 Cross-Validation Strategy

In machine learning (ML), generalization usually refers to the ability of an algorithm to be effective across various inputs. Cross-validation (CV) is a commonly used technique in applied ML tasks for evaluating models and testing its performance. It helps to provide an unbiased estimate of model performance, compare and select an appropriate model for the specific predictive modeling problem.

Cross-validation scheme includes few techniques; k-fold cross-validation, where is split into 'k' folds, training on k-1 folds and testing on the remaining one. The process is repeated until each fold serves as the test set. Stratified K-Fold is another strategy caters for imbalanced datasets. While, Leave-One-Subject-Out (LOSO) is more suitable for small data, or time-series splits for temporal data. LOSO Cross-validation scheme is used in EMG research to ensure subject-independent evaluation.

5 Experimental Results and Analysis

The three NN types were individually evaluated by feeding the dataset to their input layers and monitoring the NN performance parameters. The training dataset is organized into a complex 3-D structure representing a substantial volume of data from 672 subjects, 8 channels, and 6 motions, totaling over 2000 instances. Where different sets for validation and testing are used. Processing this large data volume requires efficient feature extraction and model training. Through the preprocessing, each sEMG signal is decomposed by applying the mother wavelet Daubechies-2 DWT to generate the detail coefficients (cD1-cD8) and an Approximation Coefficient (AC). So, the dataset is tabulated and packed in cells organized in [subject (k)* cD(i) * channel (1-8)] matrix as in Fig. 7.

block #1

reading	data cell	case
1	(cD1 * *channels) cell*	1
⋮	⋮	⋮
672	(cD1 * *channels) cell*	6

block #2

reading	data cell	case
1	(cD2 * *channels) cell*	1
⋮	⋮	⋮
672	(cD2 * *channels) cell*	6

⋮ ⋮ ⋮

block #8

reading	data cell	case
1	(cD8 * *channels) cell*	1
⋮	⋮	⋮
672	(cD8 * *channels) cell*	6

Fig.7 The Dataset is Organized in Cell Arrays. [Authors' own work]

The experimentation results of the three NNs, LSTM, GRU, and BRNN model architecture is characterized by: input size = instances of cDs, fully connected, SoftMax activation, and learning rate = 0.0005, Adam optimizer, Max epochs 25, Mini Patch Size 10, Hidden Units 100. A uniform random noise (Additive White Gaussian Noise) defined by MATLAB rand function within specified intervals (-1, +1), (-2, +2), (-3, +3), and (-4, +4) is added to basic test data for testing classification accuracy sensitivity. The individual classification performance for each DWT detail coefficient is statistically illustrated by ANOVA in Table 2.

Table 2: Performance Accuracy of NNs

		% Test accuracy				
DWT level	**NN**	**Basic (M±SD)**	**RND ±1 (M±SD)**	**RND ±2 (M±SD)**	**RND ±3 (M±SD)**	**RND ±4 (M±SD)**
cD1	**LSTM**	83.48 ±1.16	83.23 ±1.12	81.30 ±1.24	81.58 ±1.09	72.61 ±1.41
	GRU	85.35 ±0.89	84.79 ±1.34	84.94 ±1.25	82.85 ±1.14	70.70 ±1.39
	BRNN	80.60 ±1.46	80.1 ±1.33	79.94 ±1.57	79.04 ±1.03	67.68 ±1.26
cD2	**LSTM**	87.90 ±1.32	88.12 ±1.28	87.87 ±1.42	81.69 ±.89	78.55 ±1.51
	GRU	88.58 ±0.89	88.53 ±0.84	81.64 ±1.36	62.06 ±1.51	46.53 ±1.47
	BRNN	88.32 ±1.17	88.22 ±1.25	88.20 ±1.01	73.93 ±1.36	57.53 ±0.98
cD4	**LSTM**	91.07 ±1.26	85.04 ±0.57	56.64 ±1.36	52.11 ±1.27	42.74 ±1.19
	GRU	90.60 ±1.25	86.89 ±1.20	61.73 ±0.99	41.02 ±0.71	29.28 ±0.92
	BRNN	90.27 ±1.29	81.26 ±1.16	38.71 ±1.16	22.12 ±1.24	16.94 ±1.32
cD7	**LSTM**	93.04 ±1.52	48.65 ±1.09	30.29 ±1.49	24.34 ±1.15	19.85 ±1.40
	GRU	92.72 ±1.26	53.97 ±1.18	35.12 ±1.01	22.41 ±1.41	17.15 ±1.26
	BRNN	91.59 ±0.97	47.75 ±1.13	29.28 ±1.20	14.72 ±1.30	12.48 ±1.13
UNI ANOVA	**F score**	8.447	20.085	259.37	460.002	406.684
	P value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
	R square	0.909	0.995	0.997	0.998	0.997

Descriptive statistics for the performance of three neural network architectures (LSTM, GRU, BRNN) across four DWT decomposition levels (cD1, cD2, cD4, cD7) for the basic dependent variable show that; at cD1, GRU achieved the highest mean score (M = 85.36, SD = 0.89), outperforming LSTM (M = 83.48, SD = 1.16) and BRNN (M = 80.60, SD = 1.46), and at cD2, GRU again recorded the highest mean (M = 88.58, SD = 0.89), though differences among models were minimal. At cD4, LSTM surpassed the other models (M = 91.07, SD = 1.26), while GRU and BRNN performed slightly lower (M = 90.60 and M = 90.27, respectively). And at cD7, LSTM maintained superiority (M = 93.04, SD = 1.52), followed by GRU (M = 92.72, SD = 1.26) and BRNN (M = 91.59, SD = 0.97). Across all levels, GRU exhibited the highest overall mean (M = 89.31, SD = 2.94), followed closely by LSTM (M = 88.87, SD = 3.88), while BRNN consistently lagged behind (M = 87.70, SD = 4.47). Eventually, GRU demonstrated the lowest variability, suggesting greater stability in performance compared to LSTM and BRNN. The DWT Level and NN interaction was significant (F score=8.45, p value<0.001), demonstrating that the relative performance of the neural networks varied depending on the decomposition level.

The pairwise comparison in Table 3 revealed that the relative performance of the three recurrent neural architectures varied across wavelet decomposition levels for the basic dependent variable, and statistical significance is evaluated at 0.05 level. At the cD1 level, GRU consistently outperformed both LSTM (mean difference= 1.875, p<0.001) and BRNN (mean difference= 4.752, p<0.001), with statistically significant mean differences indicating its superior effectiveness in capturing features at this resolution (mean difference= 4.752, p<0.001). In contrast, BRNN performed markedly worse than both LSTM and GRU, underscoring its limitations at lower decomposition levels. At intermediate levels (cD2 and cD4), no significant differences were observed among the models, suggesting that architectural choice had little impact on performance when features were extracted at these scales. However, at the cD7 level, LSTM demonstrated a significant advantage over BRNN (mean difference= 1.448, p<0.009), while GRU showed a marginal but weaker edge (mean difference= 1.122, p<0.043). These results highlight that BRNN tended to underperform across contexts, particularly at cD1 and cD7, whereas GRU excelled at the lowest level and LSTM maintained competitive performance at higher levels.

The effect of uniform random noise (Additive White Gaussian Noise) on the classification accuracy of each NN is analyzed and recognized significantly by experimenting with many noise levels. The results in Fig. 8 reveal that;

- GRU excels under clean and low-noise conditions, particularly at shallow (cD1) and deep (cD7) decomposition levels.

- LSTM gains advantage as noise intensity increases, becoming dominant under moderate-high (±3) and high (±4) noise conditions, especially at intermediate decomposition levels (cD2, cD4).

- BRNN consistently underperforms, failing to match the robustness of LSTM or GRU across all conditions.

- Noise intensity interacts with decomposition depth: GRU is optimal in low-noise environments, while LSTM is superior in high-noise contexts.

Table 3: Pairwise comparison

			Basic		RND ±1		RND ±2		RND ±3		RND ±4	
DWT level		**NN**	**Mean Difference**	**P value**	**Mean Difference**	**P value**	**Mean Difference**	**P value**	**Mean Difference**	**P value**	**Mean Difference**	**P value**
cD1	**LSTM**	**GRU**	-1.875	0.001	-1.57	0.002	-3.64	0.000	-1.271	0.020	1.909	0.001
		BRNN	2.877	0.000	3.127	0.000	1.361	0.019	2.536	0.000	4.934	0.000
	GRU	**LSTM**	1.875	0.001	1.57	0.002	3.64	0.000	1.271	0.020	-1.909	0.001
		BRNN	4.752	0.000	4.697	0.000	5.001	0.000	3.807	0.000	3.025	0.000
	BRNN	**LSTM**	-2.877	0.000	-3.127	0.000	-1.361	0.019	-2.536	0.000	-4.934	0.000
		GRU	-4.752	0.000	-4.697	0.000	-5.001	0.000	-3.807	0.000	-3.025	0.000
cD2	**LSTM**	**GRU**	-0.675	0.221	-0.401	0.423	6.225	0.000	19.632	0.000	32.027	0.000
		BRNN	-0.428	0.436	-0.097	0.846	-0.332	0.562	7.76	0.000	21.022	0.000
	GRU	**LSTM**	0.675	0.221	0.401	0.423	-6.225	0.000	-19.632	0.000	-32.027	0.000
		BRNN	0.247	0.653	0.304	0.544	-6.556	0.000	-11.872	0.000	-11.005	0.000
	BRNN	**LSTM**	0.428	0.436	0.097	0.846	0.332	0.562	-7.76	0.000	-21.022	0.000
		GRU	-0.247	0.653	-0.304	0.544	6.556	0.000	11.872	0.000	11.005	0.000
cD4	**LSTM**	**GRU**	0.465	0.397	-1.849	0.000	-5.09	0.000	11.086	0.000	13.467	0.000
		BRNN	0.797	0.148	3.778	0.000	17.933	0.000	29.991	0.000	25.799	0.000
	GRU	**LSTM**	-0.465	0.397	1.849	0.000	5.09	0.000	-11.086	0.000	-13.467	0.000
		BRNN	0.332	0.546	5.627	0.000	23.024	0.000	18.905	0.000	12.333	0.000
	BRNN	**LSTM**	-0.797	0.148	-3.778	0.000	-17.933	0.000	-29.991	0.000	-25.799	0.000
		GRU	-0.332	0.546	-5.627	0.000	-23.024	0.000	-18.905	0.000	-12.333	0.000
cD7	**LSTM**	**GRU**	0.326	0.553	-5.318	0.000	-4.811	0.000	1.935	0.000	2.703	0.000
		BRNN	1.448	0.009	0.892	0.076	1.013	0.078	9.641	0.000	7.375	0.000
	GRU	**LSTM**	-0.326	0.553	5.318	0.000	4.811	0.000	-1.935	0.000	-2.703	0.000
		BRNN	1.122	0.043	6.211	0.000	5.824	0.000	7.705	0.000	4.673	0.000
	BRNN	**LSTM**	-1.448	0.009	-0.892	0.076	-1.013	0.078	-9.641	0.000	-7.375	0.000
		GRU	-1.122	0.043	-6.211	0.000	-5.824	0.000	-7.705	0.000	-4.673	0.000

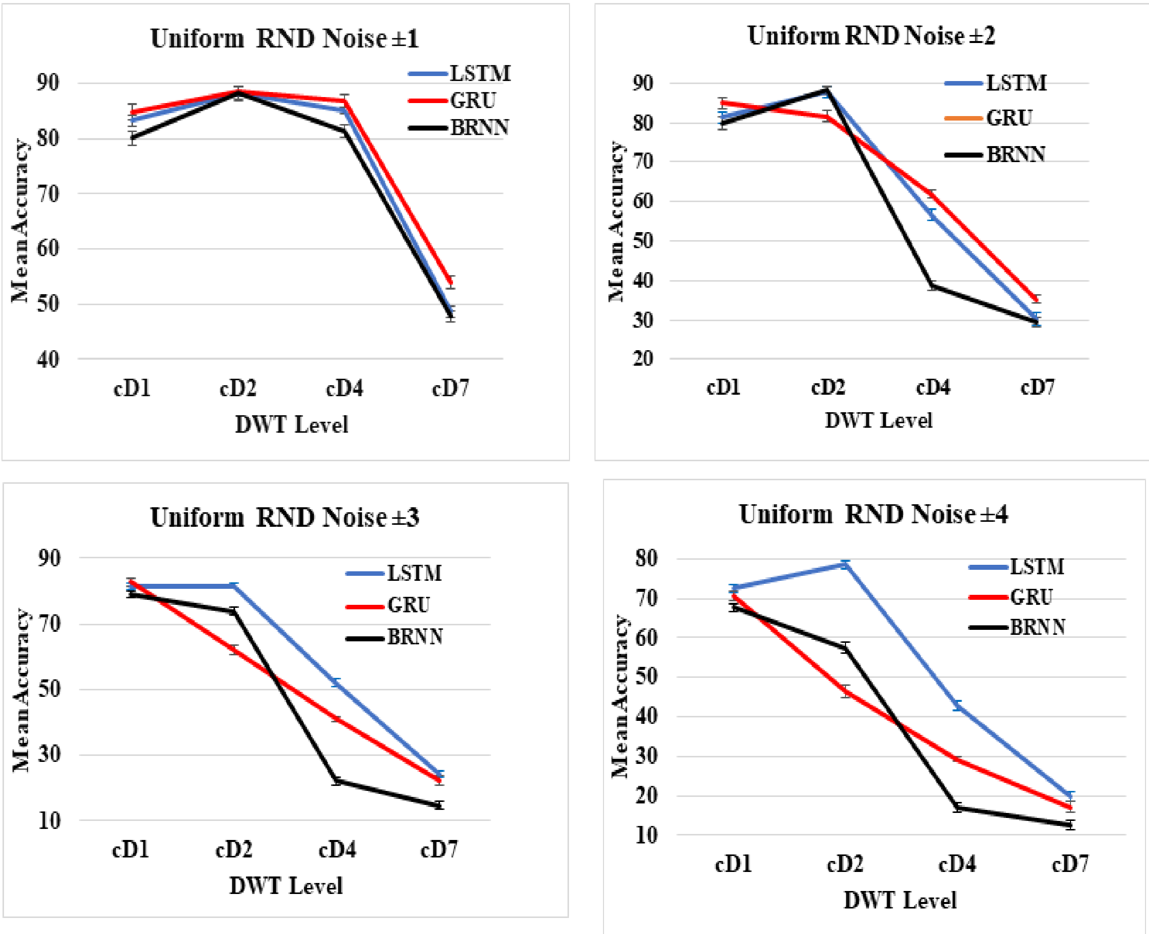


Fig.8 The Dependency of NNs Classification Accuracy on the DWT Levels in the Presence of Noise. [Authors' own work]

The summary findings are listed in Table 4, where in the five experimental conditions, GRU dominates in clean and low-noise settings, while LSTM consistently outperforms under moderate to high noise intensities. BRNN remains the weakest performer across all scenarios. These results demonstrate that noise intensity and decomposition depth jointly moderate neural network performance, reinforcing the necessity of tailoring architecture choice to signal environment in hybrid wavelet–neural frameworks.

Table 4 Condition-specific findings summary

Noise Condition	**Outcomes**	**Interpretation**
Basic (No Noise)	GRU outperforms at cD1; LSTM ≈ GRU at cD7; BRNN weakest.	GRU excels in clean signals, especially shallow decomposition.
Low Noise ±1	GRU superior at cD1, cD4, cD7; LSTM > BRNN.	GRU shows strong robustness under low noise.
Moderate Noise ±2	GRU dominates at cD1, cD4, cD7; LSTM superior at cD2.	GRU generally robust, but LSTM gains advantage at mid-level decomposition.

Moderate Noise ±3	LSTM superior at cD2, cD4, cD7; GRU≈ LSTM at cD1; BRNN weak.	LSTM emerges as most noise-robust at moderate-high intensity.
High Noise ±4	LSTM consistently outperforms GRU and BRNN across all levels.	LSTM dominates under severe noise perturbation.

6 Discussion

Comparing the performance of LSTM, GRU, and BRNN in sEMG classification leads to the following observations: First of all, the three NNs each have unique features and advantages in EMG signal classification, but differ in the way of handling the task, which coincides with the results that are stated in literature as follows;

1. LSTM: Generally, it was designed to address the vanishing gradient problem in traditional RNNs. So, it is effective and excellent in handling sequential data with long-term dependencies by remembering information over extensive timesteps [21,31].
2. GRU: This architecture represents a more streamlined iteration of LSTM, designed to address the same problems (vanishing gradient) but with fewer parameters and simpler gates. It exhibits substantial effectiveness in sequential EMG signal data and may achieve better accuracy [32].
3. BRNN: This model is distinguished by its method of data processing in both forward and backward directions, followed by the concatenation of outputs from both directions prior to being forwarded to the subsequent layer. It has demonstrated commendable accuracy in the classification of sequential EMG signals [31].

This study investigated the efficacy of LSTM, GRU, and BRNN architectures for large-scale EMG signal classification, demonstrating that robust performance can be derived from complex, high-volume datasets. The ability of the three architectures to handle noisy, nonstationary, and time-dependent data, is focused on and explored. Generally, the three NNs demonstrate good performance in the sEMG classification task with slight differences. The experimental results in Table 2 and the graphs in Fig. 8 demonstrate the basic facts embedded in the DWT that the higher levels contain core details of the EMG signal. The accuracy of classification when the basic sEMG dataset for Cd7 is tested by LSTM, GRU, BRNN reached 93.04±1.52, 92.72±1.26, and 91.59±0.97, respectively. In the case of noisy data, the BRNN shows lower performance than GRU and LSTM. This indicates that the EMG signal is corrupted by noise, and the signal-to-noise ratio becomes lower.

It is worth noting that the classification performance of all three networks can be enhanced through systematic architectural tuning. Hyperparameter optimization led to improved generalization, especially under noisy conditions, by balancing model capacity and regularization.

The adopted configurations, mainly the use of two recurrent layers and a moderate dropout rate (0.2–0.3), prevented overfitting while preserving the networks' ability to capture temporal dependencies in EMG data.

7 Conclusions and Future work

This paper presented a comparative study of the performance of three neural network architectures, namely; LSTM, GRU, and BRNN. Several experiments were conducted on a large-scale training dataset comprising 672 subject recordings across six hand gestures, enabling a robust, data-driven comparison. The experimental results and their analysis suggest that the effectiveness of recurrent neural architectures in wavelet-based feature extraction is strongly dependent on the decomposition level. GRU appears particularly well-suited for lower-level detail representation, while BRNN consistently lags, and LSTM provides stable though less dominant performance across levels. BRNNs may be reserved for contexts requiring bidirectional analysis, where incorporating broader contextual information could improve performance. In practice, the optimal choice among these models depends on the noise characteristics, available computational resources, and EMG signal complexity. Future work could explore hybrid architectures or advanced noise reduction techniques to enhance practical application robustness.

As for future work, hybrid architectures that combine the noise robustness of LSTM with the computational efficiency of GRU could be explored to improve real-time performance in noisy environments. Additionally, the integration of advanced signal denoising techniques; such as adaptive filtering or deep generative models, before feature extraction may further enhance classification stability under varying noise conditions. The models can be further generalized for practical prosthetic and human–computer interaction applications by extending the evaluation to larger and more diverse EMG datasets. The Leave-One-Subject-Out (LOSO) Cross-validation scheme should be considered for implementation, where it can provide an unbiased estimate of model performance, compare and select an appropriate model for the specific predictive modeling problem.

References

[1] Desplenter T, Zhou Y, Edmonds BP, Lidka M, Goldman A, Trejos AL. (2020). Rehabilitative and assistive wearable mechatronic upper-limb devices: A review. *Journal of Rehabilitation and Assistive Technologies Engineerin*. 7. doi:10.1177/2055668320917870.

[2] Shi WT, Lyu ZJ, Tang ST, Chia TL, Yang CY. A. (2018). bionic hand controlled by hand gesture recognition based on surface EMG signals: A preliminary study. *Biocybernetics and Biomedical Engineering*, 38(1). doi:10.1016/j.bbe.2017.11.001.

[3] Toro-Ossaba A, Jaramillo-Tigreros J, Tejada JC, Peña A, López-González A, Castanho RA. (2020). LSTM Recurrent Neural Network for Hand Gesture Recognition Using EMG Signals. *Applied Sciences (Switzerland)*, 12(19). doi:10.3390/app12199700.

[4] Aljebory KM, Jwmah YM, Mohammed TS. (2024). Classification of EMG Signals: Using DWT Features and ANN Classifier. *IAENG International Journal of Computer Science*, 51(1), 23–31.

[5] Asif AR, Waris A, Gilani SO, Jamil M, Ashraf H, Shafique M, et al. (2020). Performance evaluation of convolutional neural network for hand gesture recognition using EMG. *Sensors (Switzerland)*, 20(6). doi:10.3390/s20061642.

[6] Dweiri Y, Hajjar Y, Hatahet O. A. (2023). novel neuroevolution model for EMG-based hand gesture classification. *Neural Computing and Applications*, 35(14). doi:10.1007/s00521-023-08253-1.

[7] Malhotra R, Singh P. (2023). Recent advances in deep learning models: a systematic literature review. *Multimedia Tools and Applications*, 82(29). doi:10.1007/s11042-023-15295-z.

[8] LeCun Y, Hinton G, Bengio Y. Deep learning (2015), Y. LeCun, Y. Bengio and G. Hinton. *Nature* 2015; 521.

[9] Yamashita R, Nishio M, Do RKG, Togashi K. (2018). Convolutional neural networks: an overview and application in radiology. Insights into Imaging, 9(4). doi:10.1007/s13244-018-0639-9.

[10] Zihlmann M, Perekrestenko D, Tschannen M. (2017). Convolutional recurrent neural networks for electrocardiogram classification. *Computing in Cardiology*, doi:10.22489/CinC.2017.070-060.

[11] Geng W, Hu Y, Wong Y, Wei W, Du Y, Kankanhalli M. (2018). A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition. *PLoS*, 13(10). doi:10.1371/journal.pone.0206049.

[12] Wang Y, Wu Q, Dey N, Fong S, Ashour A.S. (2020). Deep back propagation–long short-term memory network based upper-limb sEMG signal classification for automated rehabilitation. *Biocybernetics and Biomedical Engineering*, 40(3). doi:10.1016/j.bbe.2020.05.003.

[13] Elsharkawy A.N, Zayed N. (2025). Exploring the Effects of Wavelet Types and Windowing on EMG-based IONM Through Deep Learning Architectures. *Neuroscience Informatics*, 100253.

[14] William L, Dali M, Azevedo Coste C, Guiraud D. (2022). A method based on wavelets to analyze overlapped and dependent M-Waves. *Journal of Electromyography and Kinesiology*, 63. doi:10.1016/j.jelekin.2022.102646.

[15] Aviles M, Alvarez-Alvarado JM, Robles-Ocampo JB, Sevilla-Camacho PY, Rodríguez-Reséndiz J. (2024). Optimizing RNNs for EMG Signal Classification: A Novel Strategy Using Grey Wolf Optimization. *Bioengineering*, 11(1). doi:10.3390/bioengineering11010077.

[16] Deif MA, Solyman AAA, Kamarposhti MA, Band SS, Hammam RE. (2021). A deep bidirectional recurrent neural network for identification of SARS-CoV-2 from viral genome sequences. *Mathematical Biosciences and Engineering*, 18(6). doi:10.3934/mbe.2021440.

[17] Cahuantzi R, Chen X, Güttel S. (2023). A Comparison of LSTM and GRU Networks for Learning Symbolic Sequences. *Lecture Notes in Networks and Systems*. doi:10.1007/978-3-031-37963-5_53.

[18] Sathish C, Mahesh P, Byanigoudar PF. (2024). Comparison of LSTM and GRU Neural Networks' Performance on the IMDB Sentiment Analysis Dataset. *Journal of Nonlinear Analysis and Optimization*, 15(10).

[19] Yang S, Yu X, Zhou Y. (2020). LSTM and GRU Neural Network Performance Comparison Study: Taking Yelp Review Dataset as an Example. *Proceedings - 2020*

International Workshop on Electronic Communication and Artificial Intelligence, IWECAI 2020. doi:10.1109/IWECAI50956.2020.00027.

[20] Da Silva DG, Meneses AA de M. (2023). Comparing Long Short-Term Memory (LSTM) and bidirectional LSTM deep neural networks for power consumption prediction. *Energy Reports*, 10. doi:10.1016/j.egyr.2023.09.175.

[21] Jabbari M, Khushaba RN, Nazarpour K. (2020). EMG-Based Hand Gesture Classification with Long Short-Term Memory Deep Recurrent Neural Networks. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*. doi:10.1109/EMBC44109.2020.9175279.

[22] Mienye ID, Swart TG. (2024). A comprehensive review of deep learning: Architectures, recent advances, and applications. *Information*, 15(12): 755.

[23] Kok CL, Ho CK, Tan FK, Koh YY. (2024). Machine learning-based feature extraction and classification of EMG signals for intuitive prosthetic control. *Applied Sciences*, 14(13): 5784.

[24] Phinyomark A, Nuidod A, Phukpattaranont P, Limsakul C. (2012). Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification. *Elektronika Ir Elektrotechnika*, 122(6). doi:10.5755/j01.eee.122.6.1816.

[25] Puttasakuf T, Sangworasil M, Matsuura T. (2014). Feature extraction of wavelet transform coefficients for sEMG classification. *BMEiCON 2014 - 7th Biomedical Engineering International Conference*. doi:10.1109/BMEiCON.2014.7017435.

[26] Ozkaya U, Coskun O, Comlekci S. (2010). Frequency analysis of EMG signals with Matlab sptool. *9th WSEAS International Conference on Signal Processing, SIP '10*.

[27] Ayachi M, Seddik H. (2022). (2022). Overview of EMG Signal Preprocessing and Classification for Bionic Hand Control. *2022 IEEE Information Technologies and Smart Industrial Systems, ITSIS 2022*. doi:10.1109/ITSIS56166.2022.10118387.

[28] Anand D, Bhateja V, Srivastava A, Tiwari DK. (2018). An approach for the preprocessing of EMG signals using canonical correlation analysis. *Smart Innovation, Systems and Technologies*. doi:10.1007/978-981-10-5547-8_21.

[29] Doulah ABMSU, Fattah SA, Zhu WP, Ahmad MO. (2014). Wavelet domain feature extraction scheme based on dominant motor unit action potential of EMG signal for neuromuscular disease classification. *IEEE Transactions on Biomedical Circuits and Systems*. 8(2). doi:10.1109/TBCAS.2014.2309252.

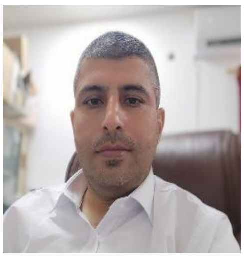
[30] Ahlawat V, Narayan Y, Kumar D. (2021). DWT-Based Hand Movement Identification of EMG Signals Using SVM. *Lecture Notes in Networks and Systems*, doi:10.1007/978-981-33-6546-9_47.

[31] Xiong D, Zhang D, Zhao X, Zhao Y. (2020). Hand Gesture Recognition Using Instant High-density EMG Graph via Deep Learning Method. *Proceedings. Chinese Automation Congress, CAC 2020*. doi:10.1109/CAC51589.2020.9326536.

[32] Samadani A. (2018). Gated Recurrent Neural Networks for EMG-Based Hand Gesture Classification. A Comparative Study. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* 2018; doi:10.1109/EMBC.2018.8512531.

Notes on contributors

Karim M. Aljebory is an Associate Professor and Head of the Computer Technical Engineering Department at Al-Qalam University College, Kirkuk, Iraq. He earned his Ph.D. in Computer Control from Georgian Technical University in 1992. He has held academic and leadership positions at the University of Technology, Baghdad, and Isra University, Jordan. He has published over 40 papers in national and international journals and conferences. Contact: karim.eng@alqalam.edu.iq



Dr.Yashar M. Jwmah, is a Bioengineer at the Kirkuk Health Directorate, Ministry of Health, Kirkuk, Iraq. His teaching and research interests focus on Bioengineering and Biomedical Engineering. He has research works leading to publications in national and international journals in this field of interest.

Thabit Sultan Mohammed is currently an Associate professor at the computer engineering department of Al-Qalam University, Kirkuk, Iraq. He received his PhD. in Computing Systems from Cranfield University, UK. After working at a number of Iraqi Universities, he, joined Al-Zaytoonah University of Jordan, followed by Dhofar university, Oman. He has authored over 40 publications in national and international journals.

Dr. Adnan Saif Said AL Mamari is a lecturer at the Department of Computing and Science at Modern College of business and Science. His main teaching and research interests include Digital Signal Analysis, Physics and Fuzzy Logic. He has published some articles in International Journal of Power Electronics and Drive System (IJPEDS)