IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 15 October 2024, accepted 12 November 2024, date of publication 18 November 2024, date of current version 9 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3501013

RESEARCH ARTICLE

Distinguishing Resting State From Motor Imagery Swallowing Using EEG and Deep Learning Models

SEVGI GÖKÇE ASLAN[1,2] AND BÜLENT YILMAZ[3,4]

[1]Department of Electrical and Computer Engineering, Abdullah Gül University, 38080 Kayseri, Türkiye
[2]Department of Biomedical Engineering Department, İnönü University, 44280 Malatya, Türkiye
[3]GUST Engineering and Applied Innovation Research Center (GEAR), Gulf University for Science and Technology (GUST), Hawally 32093, Kuwait
[4]Department of Electrical and Computer Engineering, Gulf University for Science and Technology (GUST), Hawally 32093, Kuwait

Corresponding author: Sevgi Gökçe Aslan (sevgigokce38@gmail.com)

This work was supported by the Gulf University for Science and Technology, Kuwait, for covering the Article Processing Charges (APC).

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Erciyes University Ethics Committee Clinical Research Ethics Committee, Kayseri, Turkey, under Approval No. 2023/461, on July 12, 2023.

ABSTRACT The primary aim of this study was to assess the classification performance of deep learning models in distinguishing between resting state and motor imagery swallowing, utilizing various preprocessing and data visualization techniques applied to electroencephalography (EEG) data. In this study, we performed experiments using four distinct paradigms such as natural swallowing, induced saliva swallowing, induced water swallowing, and induced tongue protrusion on 30 right-handed individuals (aged 18 to 56). We utilized a 16-channel wearable EEG headset. We thoroughly investigated the impact of different preprocessing methods (Independent Component Analysis, Empirical Mode Decomposition, bandpass filtering) and visualization techniques (spectrograms, scalograms) on the classification performance of multichannel EEG signals. Additionally, we explored the utilization and potential contributions of deep learning models, particularly Convolutional Neural Networks (CNNs), in EEG-based classification processes. The novelty of this study lies in its comprehensive examination of the potential of deep learning models, specifically in distinguishing between resting state and motor imagery swallowing processes, using a diverse combination of EEG signal preprocessing and visualization techniques. The results showed that it was possible to distinguish the resting state from the imagination of swallowing with 89.8% accuracy, especially using continuous wavelet transform (CWT) based scalograms. The findings of this study may provide significant contributions to the development of effective methods for the rehabilitation and treatment of swallowing difficulties based on motor imagery-based brain computer interfaces.

INDEX TERMS Deep learning, EEG, motor imagery, scalogram, spectrogram, swallowing.

I. INTRODUCTION

Swallowing is a fundamental function for daily life activities, and swallowing difficulties can significantly affect the quality of life of individuals [1]. Swallowing difficulties, known as dysphagia, affect a significant portion of the population, particularly among those with neurological disorders or aging-related diseases. Swallowing difficulties, which can arise from various reasons such as neurological disorders, muscle diseases, injuries, or other medical conditions, can lead to challenges in nutrition and fluid intake, resulting in malnutrition and hydration issues [2], [3], [4], [5]. Therefore, developing effective assistance methods for individuals with swallowing difficulties is of critical importance. Particularly, recognizing this condition and implementing appropriate interventions can enhance patients' quality of life and regulate their dietary habits [6].

Current methodologies for detecting and treating dysphagia include a range of invasive and non-invasive techniques. While invasive methods, such as endoscopic [7] and

The associate editor coordinating the review of this manuscript and approving it for publication was Lorenzo Mucchi.

fluoroscopic [8] evaluations, provide detailed anatomical information, they are often uncomfortable for patients and carry certain risks. Non-invasive techniques, such as electroencephalography (EEG) [9], functional magnetic resonance imaging (fMRI) [10], positron emission tomography (PET) [11], near-infrared spectroscopy (NIRS) [12], and magnetoencephalography (MEG) [13], are becoming the preferred choice for detecting dysphagia due to their advantages over invasive techniques in assessing brain activity. Various techniques have been explored for controlling Brain-Computer Interfaces (BCIs) in patient populations, particularly due to the non-invasive nature and low cost of electroencephalography (EEG) and near-infrared spectroscopy (NIRS) [14]. Various approaches exist for the detection and analysis of motor imagery swallowing and motor imagery tongue movement based on electroencephalography (EEG) and near-infrared spectroscopy (NIRS). The comparable and overlapping activation of brain areas for motor imagery swallowing and motor execution swallowing can support the use of motor imagery swallowing in developing practical rehabilitation tools for training stroke-related dysphagia patients. Kober et al. investigated hemodynamic changes in the brain in response to motor execution and motor imagery of swallowing using NIRS [15], [16]. They demonstrated for the first time that motor imagery of swallowing could be successfully used as a mental strategy in a neurofeedback training paradigm. Before and after training, they assessed cortical activation patterns during motor execution and imagery of swallowing [17]. They compared the hemodynamic response observed during the swallowing of water or saliva [18]. Among these, EEG has been widely utilized in the field of brain-computer interfaces (BCIs) due to its ability to capture rapid changes in brain activity with high temporal resolution. Its non-invasive nature and relative cost-efficiency make it a preferred choice for many research and clinical applications. However, EEG signals are often susceptible to noise, which can complicate data interpretation. Despite this limitation, advances in EEG signal processing techniques have improved the reliability of EEG-based assessments, particularly in the context of motor imagery tasks relevant to swallowing rehabilitation. Electroencephalogram (EEG) signals are used to measure brain activity and can be employed to monitor the electrical activity occurring in the brain during swallowing [19]. Swallowing is a complex process managed by the brain and nervous system, resulting in specific patterns and changes in brain activity during this process [20]. EEG signals provide an essential tool for examining these patterns and associating them with swallowing. Specifically, monitoring brain activity related to swallowing through EEG is a crucial step in understanding the neurological basis of swallowing difficulties and developing more effective treatment strategies [21]. Additionally, numerous studies use surface electroencephalography to detect cortical signals related to swallowing. Huckabee et al. found differences in brain activity between tasks by examining healthy individuals performing repetitive finger movements and dry swallows [22]. Nonaka et al. observed earlier brain activity onset and larger amplitudes during voluntary and command dry swallowing tasks [23]. Hiraoka found differences in brain potentials between dry and water swallowing tasks [24]. Cuellar et al. identified sensorimotor control and right-lateralized processing during swallowing tasks [25]. Koganemaru et al. detected changes in brain activity during voluntary swallows with water bolus application [26]. Jestrovic et al. investigated the impact of liquid viscosity on EEG signals during swallowing [27]. Jestrovic et al. explored brain network features during swallowing in different head positions and with various liquid viscosities [28], [29], [30]. Jestrovic et al. analyzed brain network differences during swallowing with distractions and varying water bolus volumes [31].

Brain-Computer Interface (BCI) technology has the potential to rehabilitate the functional capacity of impaired limbs following paralysis, with its efficacy largely dependent on output units such as robots, orthoses, and computers. Motor imagery is employed to better comprehend the impact of BCI therapy on post-stroke rehabilitation [32]. Motor imagery is the process in which individuals mentally visualize movement before physically executing it. Technologies that measure brain activity, such as EEG, can be utilized to observe these mental actions [33]. The differentiation between imagination and resting processes in the context of swallowing disorders allows for a more detailed examination of brain activity related to the swallowing process and a deeper understanding of the neurological underpinnings of swallowing dysfunctions. Moreover, this differentiation enables the development of rehabilitation and therapeutic strategies for individuals with swallowing disorders. Distinguishing imagination from resting processes allows for the assessment of individuals' ability to mentally simulate the swallowing movement and the customization of therapy protocols to enhance this ability when necessary. However, EEG signals are often noisy, making such analysis challenging. Therefore, various preprocessing techniques, such as Independent Component Analysis (ICA), Empirical Mode Decomposition (EMD), and bandpass filtering, are commonly employed to analyze EEG signals related to motor imagery. These techniques are essential for isolating relevant brain signals from noise and artifacts such as eye blinks or muscle activity. In the context of motor imagery swallowing, previous studies have utilized these methods with varying degrees of success, but a comprehensive comparison of their impact on classification performance remains limited. This study builds on existing approaches by systematically comparing the effectiveness of these preprocessing methods in differentiating between motor imagery and resting states. These preprocessing techniques play a significant role in measuring and understanding brain activity associated with motor imagery more accurately [34], [35], [36]. Methods

such as spectrograms and scalograms are commonly used to visualize and understand the time-frequency characteristics of EEG signals. Spectrograms represent the frequency components of the signal over time visually, while scalograms provide a more detailed analysis, visualizing the changes of different frequency components over time. These methods are crucial for reducing the complexity of EEG signals and better understanding the time-frequency characteristics of brain activity [37]. Particularly, time-frequency analysis is considered a valuable tool for examining complex brain activities such as motor imagery [38].

Deep learning is a technology that has made significant advancements in the field of artificial intelligence, providing effective results in various application areas [39]. One of the most widely used models of this technology is Convolutional Neural Networks (CNNs). While traditional machine learning algorithms such as Support Vector Machines (SVMs) and k-Nearest Neighbors (k-NN) have been applied to EEG signal classification [40], CNNs offer the advantage of automatically learning hierarchical features from raw data, reducing the need for manual feature extraction. CNNs have demonstrated exceptional performance, especially in tasks such as image recognition and classification [41]. However, recent studies have shown that CNNs can also be used to analyze time series data such as EEG [42]. This has enabled the expansion and development of EEG-based applications. Specifically, the use of CNNs may offer a more accurate and effective approach to analyze and classify complex brain signals [43].

In this study, we aim to investigate EEG signals related to the swallowing process through four different experimental paradigms. The utilization of various experimental paradigms aims to determine the variability in EEG signals during swallowing, providing insights into various aspects of swallowing-related brain activity. Additionally, we evaluate the impact of preprocessing methods, such as Independent Component Analysis, Empirical Mode Decomposition, and bandpass filtering, on the classification performance of EEG signals. We focus on classification results to determine the effectiveness of these methods in the rehabilitation and treatment of individuals with swallowing difficulties.

Furthermore, we investigate the effectiveness of methods such as spectrograms and scalograms in visualizing and understanding EEG signals' time-frequency characteristics. We analyze these methods to determine which one provides more accurate representations of EEG signals and contributes to a better understanding of swallowing-related brain activity.

Along these lines, the proposed approach in this study involves the classification of EEG signals to differentiate between imagination and resting processes in the context of swallowing using a deep learning model, Convolutional Neural Networks (CNNs). The results of this study showed that CNN model can be utilized to accurately classify brain activity associated with swallowing, offering a novel approach for the rehabilitation and treatment of individuals with swallowing difficulties. The findings of this study may contribute to the development of more effective rehabilitation and treatment strategies to enhance the quality of life for individuals with swallowing difficulties (dysphagia) and may offer a safe, patient-friendly, and affordable way to treat individuals with dysphagia.

II. MATERIALS AND METHODS

A. PARTICIPANTS

The experiments involved 30 right-handed individuals aged 18 to 56, all without diagnosed neurological issues (15 male, 15 female).

B. MATERIALS

During the experiments, we utilized a 16-channel g. Nautilus Research Wearable EEG headset (g.Tec, Schiedlberg, Austria). EEG recordings were acquired following the 10-20 international system of electrode placement, with a sampling frequency of 500 Hz and a range from 0.5 to 200 Hz. This headset facilitated the continuous recording of brain activity of the participants throughout the experiments.

1) EXPERIMENTAL PARADIGMS

Four experimental paradigms were used in the study, and they were approved by the Erciyes University (Kayseri, Turkey) Clinical Research Ethics Committee on July 12, 2023, with approval number 2023/461. Each paradigm is explained in detail below:

a: NATURAL SWALLOWING PARADIGM

The experiment was initiated after providing instructions to the participants. Participants were instructed to execute commands displayed on the computer screen. As illustrated in Figure 1, each participant had a bottle of water positioned in front of them and was instructed to drink only upon prompt, utilizing a pipette without moving. The experiment began with a beep sound, and each trial lasted for 10 seconds. A total of 15 repetitions were conducted. It is worth noting that participants were not given a separate command for imagination during the experiment. In the experiment, the first one second after the beep sound was considered as the rest phase, and one second preceding the initiation of the movement/swallowing phase which started with the ''drink water'' command was considered as the imagination

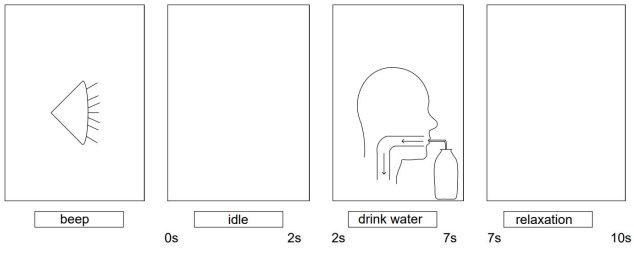

FIGURE 1. Natural swallowing experimental paradigm.

phase. This arrangement enables the analysis of the natural swallowing imagination process.

b: INDUCED SALIVA SWALLOWING PARADIGM

Each trial of the induced saliva swallowing experiment consisted of 9 seconds (see Figure 2). The experiment began with the resting command during which participants were instructed to focus on the circle in the center of the plus sign. In the imagination phase, participants were asked to imagine swallowing their saliva, and finally, in the movement/swallowing phase, participants were instructed to swallow their saliva. During the relaxation phase, participants did not receive any command. Each participant underwent this paradigm 15 times for data collection.

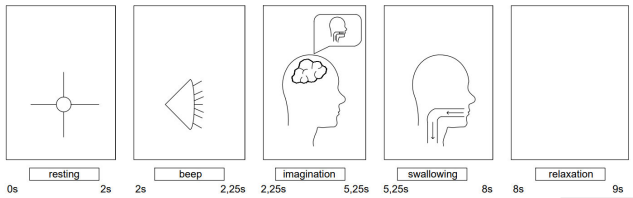

FIGURE 2. Induced saliva swallowing experimental paradigm.

c: INDUCED WATER SWALLOWING PARADIGM

In the induced water swallowing experiment, each participant was presented with a bottle and a pipette. After providing information related to the experiment, as depicted in Figure 3, the experiment started by asking the participant to take a sip of water. Participants were instructed not to swallow the water in their mouths, aiming to observe whether the induced imagination through water would create any differences in the imagination phase. The experiment continued with the resting command. Participants were asked to focus on the circle in the center of the plus sign. During the imagination phase, participants were instructed to imagine drinking the water in their mouths, and finally, in the movement/swallowing phase, participants were asked to swallow the water in their mouths. During the relaxation phase, participants were given a free period without any commands being issued. Each trial lasted 13 seconds, and data was collected from 15 repetitions.

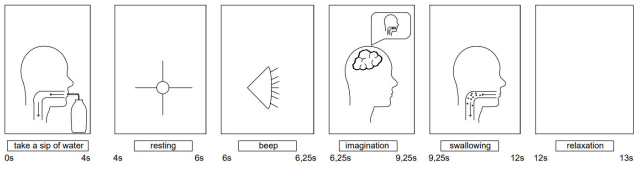

FIGURE 3. Induced water swallowing experimental paradigm.

d: INDUCED TONGUE PROTRUSION PARADIGM

The induced tongue protrusion experiment is designed to emulate the relationship between tongue protrusion and swallowing. Additionally, this paradigm has been employed in various previous studies as a surrogate for swallowing, particularly in the context of dysphagia rehabilitation [44]. As illustrated in Figure 4, the experiment was initiated with the resting phase followed by a beep sound. In the imagination phase, participants imagined protruding their tongue, and then they performed the movement by protruding their tongue. During the relaxation phase, participants were presented a blank screen. The experimental process, which lasted a total of 135 seconds, consists of 15 trials of 9 seconds each.

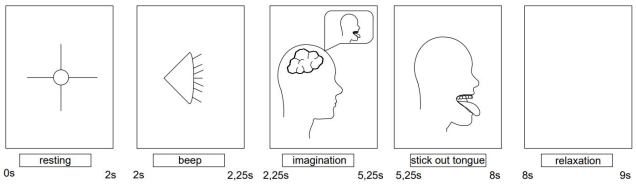

FIGURE 4. Induced tongue protrusion experimental paradigm.

III. DATA ANALYSIS

A. PREPROCESSING

The analysis of EEG signals is one of the main focal points of this study, and we aim to evaluate the importance of preprocessing steps in this analysis. In this section, four methods for the preprocessing of EEG data were examined: Empirical Mode Decomposition (EMD) Method, Independent Component Analysis (ICA), and Bandpass Filtering (BPF, 0.5-40 Hz), and the hybrid model.

To assess the contribution of preprocessing methods in distinguishing rest and imagination of swallowing, initially raw data was employed in the classification process, and later data underwent preprocessing steps including EMD, BPF and the hybrid model before the classification was performed.

1) EMPIRICAL MODE DECOMPOSITION

Empirical Mode Decomposition (EMD) is an effective method for decomposing nonstationary and non-linear data [45]. Using the EMD method to analyze EEG signals, different frequency components contained within the signal have been separated. The aim was to better understand the time-frequency characteristics of the signal. In this study, 10 Intrinsic Mode Functions (IMFs) were obtained from processing EEG signals with the EMD method, and the summation of the first 3 IMFs were selected for further analysis, as depicted in Figure 5.

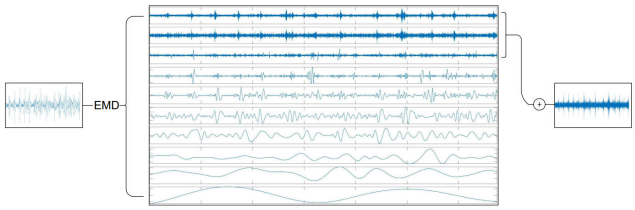

FIGURE 5. Applying EMD to the EEG signal obtained from one channel.

2) ICA FILTERING

In our experiments, we utilized Independent Component Analysis (ICA) filtering, which is commonly used in EEG studies to reduce noise artifacts such as eye blinking, muscle movements, and other physiological interferences. Our goal in applying ICA filtering is to enhance the accuracy of subsequent analysis and interpretation by ensuring that the resulting EEG data is cleaner and more reliable. Through the application of ICA, we seek to enhance the quality of EEG recordings, which is essential for obtaining robust and reproducible results in neurophysiological research.

3) BANDPASS FILTERING

Bandpass Filtering (BPF) serves as a prevalent technique for eliminating undesired components, such as power line noise, and signals beyond the intended frequency range in EEG signals. Given that bandpass filtering within the range of 0.5-40 Hz is widely adopted in EEG studies, we implemented filtering within this specified range for our experiments.

4) HYBRID MODEL: EMD + ICA + BANDPASS FILTERING (0.5-40 HZ)

The hybrid model used for the analysis of EEG signals integrates Empirical Mode Decomposition (EMD), Independent Component Analysis (ICA), and bandpass filtering (BPF). Using the EMD method [46], different frequency components contained within the signal were separated. In this study, only the summation of the first 3 IMFs out of 10 IMFs were utilized. Subsequently, ICA and BPF were applied to the resulting signal using MATLAB and EEGLAB [47], as shown in Figure 6. These steps enabled further analysis of EEG data and facilitated the removal of unwanted components.

B. GENERATION OF SPECTRAL IMAGES

After we performed preprocessing on EEG signals, we were able to work on EEG signals coming from various experimental paradigms that were i) untreated, ii) processed only with Empirical Mode Decomposition (EMD), iii) subjected only to bandpass filtering, and iv) pre-processed with the hybrid model consisting of a cascade of EMD, ICA, and bandpass filter.

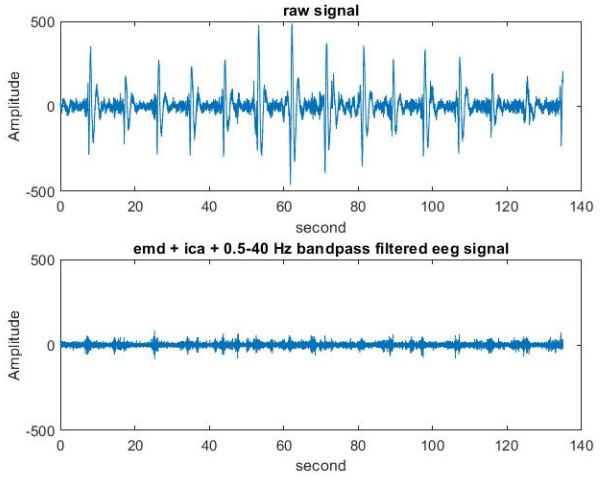

FIGURE 6. Raw and filtered signals from one EEG channel.

Our aim was to extract features from these four different signals. Given that EEG signals contain oscillatory and non-stationary frequency components, we applied time-frequency transformation techniques to obtain more information from the signals. These techniques help us better understand the relationship between the time and frequency characteristics of the signals. Additionally, various methods were employed to transform EEG signals into the time-frequency domain. Initially, the 16-channel EEG signals were combined into a single scalogram, resulting in scalogram images (see Figures 7). Subsequently, Short-Time Fourier Transform (STFT) was applied individually to each channel, followed by obtaining scalogram images using Continuous Wavelet Transform (CWT) with different wavelet types for each channel (see Figures 8). Through the utilization of these diverse transformation techniques, scalogram and spectrogram images were obtained for each experiment. These images were then adjusted to appropriate input dimensions for classification purposes.

In this study, we initially derived black-and-white images from EEG signals, where grayscale values visually represented signal intensities. To enhance the interpretability of these signals, these grayscale images were subsequently converted into color maps, thereby making the visualizations more vivid and facilitating an easier understanding of the signal characteristics. Each pixel in these visualizations corresponds to a specific time-frequency representation, where the height and width of each pixel represent a frequency and time interval, respectively. This pixel-based approach enables a comprehensive analysis of frequency changes over time, allowing for frequency-based comparisons across data from different electrodes. Notably, we did not perform normalization across channels, as we aimed to retain the raw data to reflect natural variations inherent in each electrode due to factors such as electrode placement, biological differences, and measurement conditions. This decision allows for the preservation of natural amplitude differences in the signals, maintaining the biomedical significance of these variations. The signal window size corresponding to each pixel's width was determined by the time-frequency analysis method and sampling rate, optimized to capture frequency changes over time accurately, thus facilitating a clear analysis of both the temporal and frequency components of the signals.

1) IMAGE GENERATION FROM ALL CHANNELS

The EEG data was visualized to create a composite image of all channels. This method allows for the creation of an image using the entire color range of the color map. We visualized EEG signal data as a 2-dimensional color map, transforming signal values into color representations where amplitude is mapped to specific colors, enhancing the

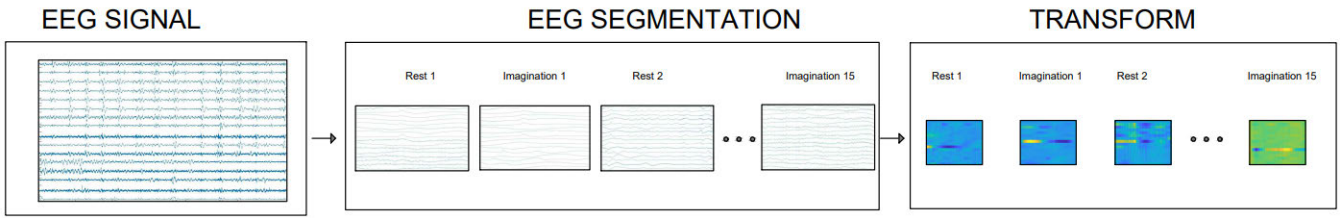

FIGURE 7. Dividing the trials into rest and imagination segments one by one and transforming them into time-channel images.

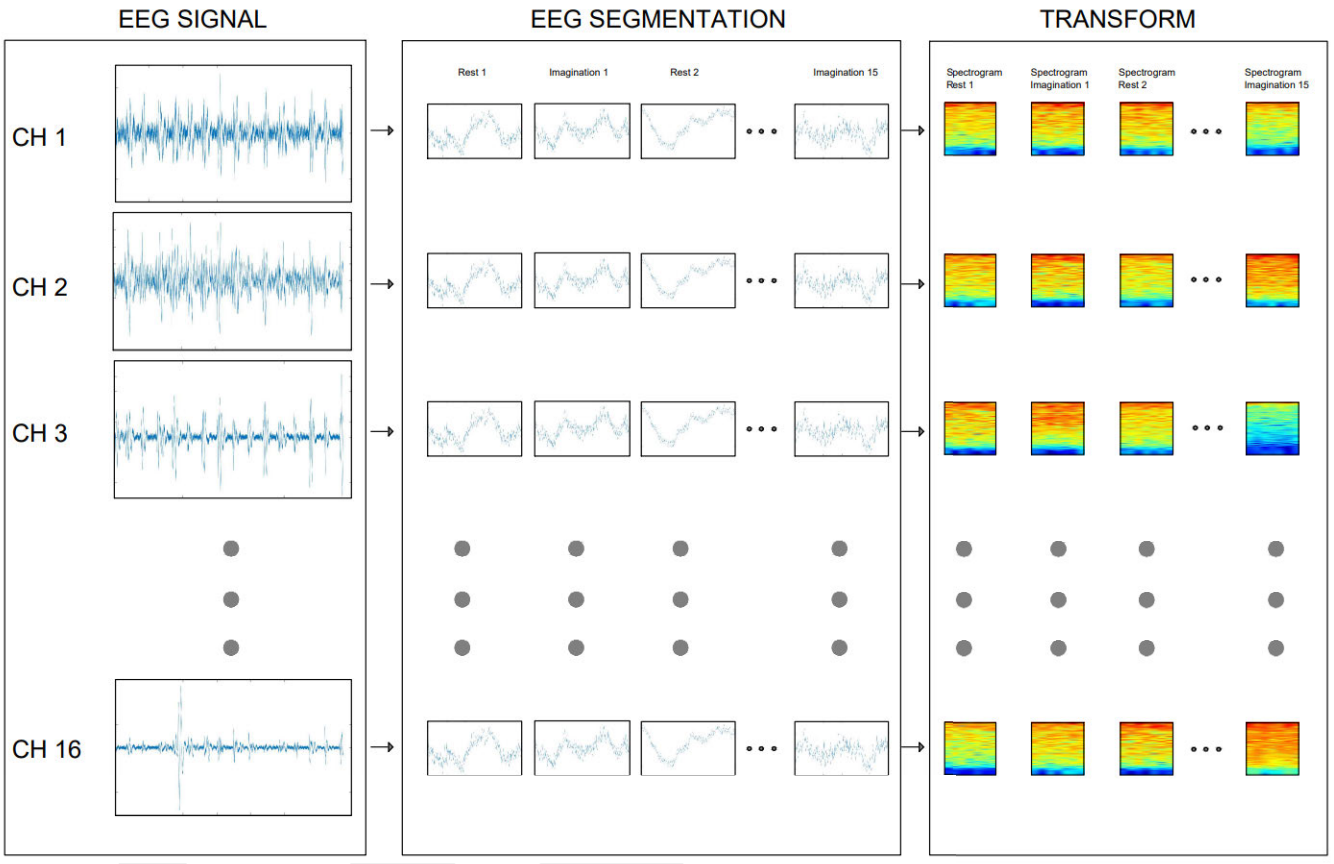

FIGURE 8. Dividing each channel the trials into rest and imagination segments one by one and transforming them into a spectrogram.

understanding of signal strength variations over time and across channels. Each pixel corresponds to an EEG channel (electrode) on the vertical axis and a specific time interval on the horizontal axis, with a 2-second signal window used for visualization; thus, the width of each pixel represents a millisecond change. Notably, normalization across channels was not performed, allowing the signals to retain their original amplitude variations, which enhances visibility of inter-channel differences and provides clearer insights into signal changes. It enables the representation of different trials, each containing rest and imagination processes, across different channels within the same visual. Each element in the signal specifies the color of a pixel in the image.

Each y-axis label represents the electrode positions associated with the EEG (see Figure 7). For example, for a 16-channel EEG recording, the y-axis will display 16 different electrode positions.

The visualization demonstrates how signals from different electrode positions change over time. These variations are made suitable for classification processes, along with other methods used to examine the time-frequency characteristics of the signal.

2) IMAGE GENERATION WITH SHORT-TIME FOURIER TRANSFORM

Short-Time Fourier Transform (STFT) is a transformation method used for analyzing the time-frequency characteristics of a signal. Essentially, it is employed to determine how the frequency components of a signal vary over time [48].

The STFT applies a windowing process to determine the frequency components of the signal at different time intervals and calculates the Fourier transform for each window. The resulting image visualizes the time on the horizontal axis, frequency on the vertical axis, and the signal power as color intensity, creating a 2-dimensional image. This image allows for a detailed visualization of the time and frequency characteristics of the signal (see Figure 7). Particularly, when analyzing frequency components that vary over time, the STFT image provides valuable insights.

3) IMAGE GENERATION WITH CONTINUOUS WAVELET TRANSFORM

Continuous Wavelet Transform (CWT) is another method used to analyze the time-frequency characteristics of a signal [49]. Unlike STFT, which uses fixed-width windows, CWT employs waveforms that can adapt to the local features of the signal in both time and frequency domains.

CWT applies waveforms of varying width and position to analyze the time and frequency characteristics of the signal. This provides a more flexible and detailed analysis compared to STFT. The resulting time-frequency image represents time on the horizontal axis, frequency on the vertical axis, and the magnitude or power of the signal as color intensity. However, the adaptability of waveforms in CWT ensures better localization in both time and frequency domains, depicting the local features of the signal more effectively compared to STFT.

In summary, the CWT time-frequency image provides a detailed representation of the signal's time-varying frequency components and offers valuable insights into local features and behaviors. Additionally, this article compares three different waveforms – Generalized Morse Wavelet, Analytic Morlet (Gabor) Wavelet (AMOR), and Bump Wavelet – to determine the most suitable waveform for feature extraction in motor imagery. The characteristics of these waveforms can influence their effectiveness in determining and analyzing the time-frequency features of motor imagery signals. The comparison results enable the selection of the most appropriate waveform for extracting the best features from motor imagery signals.

C. DEEP LEARNING

Deep learning is a subfield of machine learning developed to identify and analyze complex data patterns. This domain has become a significant tool in various fields such as neuroscience and brain-computer interfaces (BCI) [50].

Deep learning is typically implemented using artificial neural networks with one or more hidden layers. These layers are utilized to represent input data, with each layer taking the outputs of the previous layer as inputs. Consequently, the network learns more complex features and patterns [51].

Different architectures exist within deep learning, including Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) [52], [53] [54]. When working with EEG data, particularly, CNN architecture is often preferred as it can effectively process time-series data and 2D/3D structures.

CNNs can be highly effective in extracting meaningful features and identifying patterns from complex datasets such as EEG data. CNNs offer distinct advantages, especially when EEG data is visualized as image data. CNNs are particularly adept at capturing local features and neighbourhood information robustly. This capability becomes highly beneficial when EEG signals are represented as 2D or 3D structures, such as spectrograms or scalograms. CNNs process features sequentially in a way that helps in identifying significant patterns within EEG data. Additionally, CNNs are optimized for handling fixed-size inputs and outputs, which enhances their efficiency in processing high-dimensional EEG data. Therefore, in this study, CNN architecture was chosen to classify EEG data. The study evaluated the performance of the model on various experiments and different preprocessing methods applied to the data [55].

1) ARCHITECTURE

The dataset has been evaluated using 5-fold cross-validation techniques, where the entire dataset is partitioned into training and validation sets across different folds. The dataset has been divided into train and test sets using 5-fold cross-validation techniques. Subsequently, labels are added to each dataset. During the creation of the CNN architecture, various layers, and parameters were experimented with multiple times. As a result of these experiments, the layers and parameters that yielded the most successful outcome for the CNN model were identified. The CNN model with the most successful performance was constructed. The CNN model is sequenced starting from the input layer. The input layer is defined for images of size 343 × 434 for the first experiment and 256 × 256 with 3 channels (RGB) for subsequent 3 experiments. Convolutional, pooling, normalization, and fully connected layers are defined sequentially (see Figure 9). The output layer of the model is a SoftMax layer for classifying two classes (rest and imagination).

A detailed description of our CNN architecture, comprising 12 layers, is provided in the following Table 1.

- **Input Layer**: Images of size 256 × 256x3 with 'zerocenter' normalization.
- **Convolutional Layers**: Processed with 3 × 3 filters. 'same' padding and [1, 1] stride is used.
- **Batch Normalization Layer**: Batch normalization is applied.
- **ReLU Activation Layer**: ReLU activation function is applied.
- **Max Pooling Layers**: 2×2 max pooling is applied. [2, 2] stride and [0, 0, 0, 0] padding is used.
- **Fully Connected Layers**: Consists of two fully connected layers.
- **SoftMax Layer**: SoftMax activation function is applied.
- **Classification Output**: Classification output is provided using cross entropy loss.

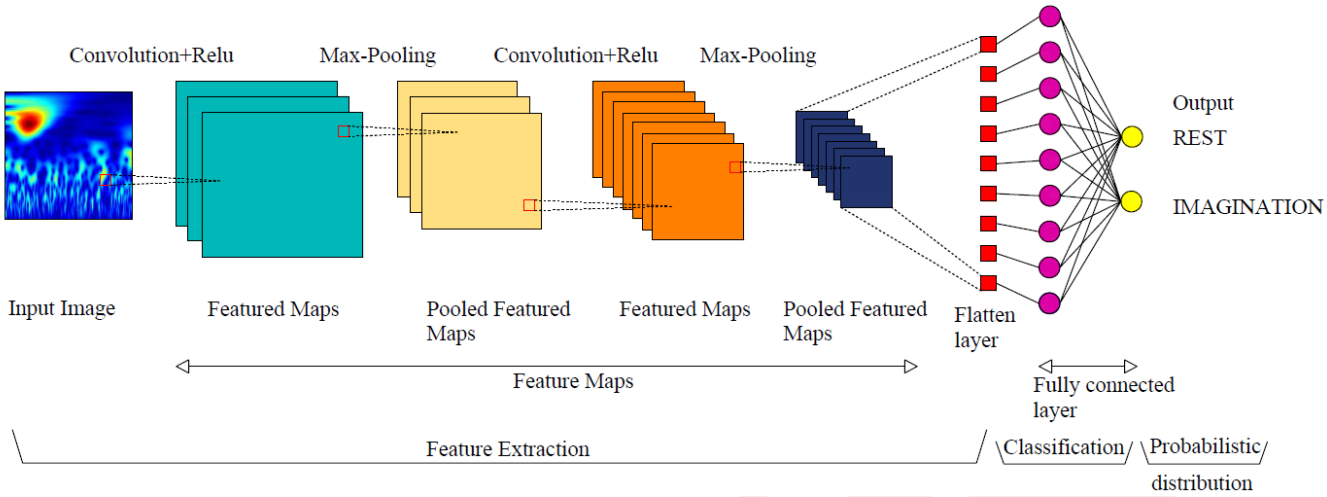

FIGURE 9. CNN architecture.

TABLE 1. Detailed layer-wise description of the CNN architecture.

Layer Name	Input Size	Filter Size	Number of Filters	Stride	Padding	Output Size
Input Layer	256x256x3	-	-	-	-	256x256x3
Conv Layer 1	256x256x3	3x3	16	1x1	same	256x256x16
Batch Norm Layer 1	256x256x16	-	-	-	-	256x256x16
ReLU Activation Layer 1	256x256x16	-	-	-	-	256x256x16
Max Pool Layer 1	256x256x16	2x2	-	2x2	[0,0,0,0]	128x128x16
Conv Layer 2	128x128x16	3x3	32	1x1	same	128x128x32
Batch Norm Layer 2	128x128x32	-	-	-	-	128x128x32
ReLU Activation Layer 2	128x128x32	-	-	-	-	128x128x32
Max Pool Layer 2	128x128x32	2x2	-	2x2	[0,0,0,0]	64x64x32
Fully Connected Layer	64x64x32 (flattened)	-	-	-	-	2 neurons
SoftMax Layer	-	-	-	-	-	2 neurons
Classification Output	-	-	-	-	-	2 neurons

2) CNN MODEL COMPILATION

The CNN model was compiled using the specified training options. Training was performed using the Adam optimization algorithm. The maximum number of epochs for training was set to 10, and the mini-batch size was set to 20. Shuffling was performed at each epoch, and validation was performed every 30 mini-batches. Training stages were not displayed, but training progress was shown graphically.

3) CNN MODEL FIT

The training data was fitted to the CNN model using the specified training options. The weights of the model were updated to adapt to the characteristics of the dataset. After the training process was completed, the model was prepared to evaluate its performance on the test dataset.

4) PERFORMANCE EVALUATION

The trained Convolutional Neural Network (CNN) model was evaluated on the test dataset. The model performed classification using the test data and compared these predictions with the ground truth labels. Based on this comparison, the accuracy, AUC-ROC, and F1 score measures of the model were calculated, and the results were presented as an evaluation.

Accuracy (ACC): This metric indicates the overall correctness of the model by showing the ratio of correctly predicted instances to the total instances. While accuracy is straightforward to understand, it may not be sufficient when dealing with imbalanced datasets.

AUC-ROC (AUC): The AUC-ROC curve is a graphical representation of a model's performance across different classification thresholds.

F1 Score (F1): The F1 Score is the harmonic mean of precision and recall, providing a balance between the two.

By using these evaluation metrics, a more detailed understanding of the model's performance is obtained. Accuracy provides a general overview, while AUC-ROC and F1 Score offer deeper insights into the model's effectiveness, especially in scenarios involving varying decision thresholds. These metrics collectively ensure a robust evaluation of the CNN model, highlighting its strengths and potential areas for improvement. The robustness of the proposed model is demonstrated through several key factors. It generalizes well across multiple paradigms, ensuring applicability in

diverse EEG settings. The model incorporates preprocessing steps to mitigate the effects of noise and artifacts, and it has been tested on raw data to assess performance in real-world applications. Additionally, rigorous cross-validation techniques, including k-fold and leave-one-out strategies, were employed to confirm its stability and accuracy. The CNN architecture further enhances the model's ability to extract features and classify data effectively, even in the presence of variability and noise.

IV. RESULTS

In this study, we analyzed electroencephalography (EEG) datasets derived from four distinct experimental paradigms pertaining to both swallowing and the imagination of swallowing. Each dataset encompasses four varieties of EEG data: raw data, data refined through Empirical Mode Decomposition (EMD) or a bandpass filter, and data pre-processed via a sequential application of EMD, Independent Component Analysis (ICA), and bandpass filtering.

Each type of data underwent transformations into images through various techniques. The first method involved converting EEG data from all channels into a single image. The second method used Short-Time Fourier Transform (STFT) to produce spectrograms of the EEG data. Lastly, Continuous Wavelet Transform (CWT) scalograms were generated using various wavelet types. Figure 10 illustrates the diverse EEG data types (raw, ica, emd, bpf, emd+ica+bpf) of the same individual within the same trial, along with different Continuous Wavelet Transform (CWT) visualization techniques representing resting and imagination states. Different visualization methods and data types yield distinct visual representations within the same dataset. Furthermore, it is observed that the resting and imagination images within the same trial are also different.

The obtained results were used to evaluate the classification performance of different EEG data types under each visualization method. This evaluation aimed to determine which preprocessing method(s) and visualization technique(s) provided the best classification results.

In our study, we employed 5-fold Cross-Validation to ensure the reliability and consistency of our model's performance on different data subsets. We divided our dataset into 5 parts to ensure equal representation of each class in the folds. The model was trained on 4 folds and evaluated on the remaining fold, with performance measured using specific metrics for each test fold. This process was repeated by rotating the folds, with each fold used exactly once for testing. Finally, we calculated the average model performance across all iterations. This approach helped guard against overfitting and produced a more generalizable and dependable model.

Tables 2 to 5 demonstrate the classification performance of EEG data types and various visualization methods. The classification performance was assessed using three different metrics. It was observed that the accuracy, AUC, and F1 score rates decreased for hybrid models and data with a bandpass filter applied. However, results obtained from experiments with Raw data, pre-processed data with Independent Component Analysis (ICA), and pre-processed data with Empirical Mode Decomposition (EMD) showed good results for accuracy, AUC, and F1 score metrics. Also, low accuracy rates were obtained when all channels merged into a single image in the channel stack.

Comparison of spectrograms obtained using Short-Time Fourier Transform (STFT) and scalograms obtained using Continuous Wavelet Transform (CWT) revealed that CWT images outperformed STFT images. The type of success varied among different CWT scalograms based on different preprocessing methods. Despite variations across experiments, the 'bump' type generally yielded the most superior results.

Results from experiments with raw data, data pre-processed with Independent Component Analysis (ICA), and data pre-processed with Empirical Mode Decomposition (EMD) exhibited variations across experiments. While data pre-processed with EMD performed better in the natural swallowing, induced saliva swallowing, and tongue protrusion experiments, raw data results were superior in the remaining experimental paradigm.

In conclusion, the findings of this study demonstrate the effectiveness of different preprocessing methods and visualization techniques in the classification of rest and imagination of swallowing and analysis of EEG signals. These findings may contribute to the development of EEG-based medical rehabilitation and treatment strategies in future studies.

V. DISCUSSION

Brain-computer interface (BCI) technologies, especially methods developed using motor imagery (MI), have made significant advancements in recent years. Motor imagery allows individuals to alter their brain activities by imagining specific movements without physically performing them. These changes are measured through methods such as electroencephalography (EEG) and are transmitted to and analyzed by computer systems. In this way, physically disabled individuals can control computers, robotic devices, or prosthetic limbs solely with the power of their minds. The development of these technologies holds great potential for both medical applications and everyday use [56].

However, there are few approaches aimed at detecting motor imagery swallowing for the rehabilitation of dysphagia patients. Yang et al. conducted comparative studies to detect motor imagery swallowing and imagery tongue movement for post-stroke dysphagia rehabilitation [44], [57], [58], [59]. These studies investigated the following hypotheses: (1) motor imagery swallowing and imagery tongue movement can be distinguished from the background resting state for use in post-stroke dysphagia rehabilitation, (2) a simple and relevant model for detecting motor imagery swallowing can be developed using imagery tongue movement EEG signals, (3) the accuracy of these hypotheses can be tested by determining classification accuracies across different sessions and modes, and (4) classification accuracies can be

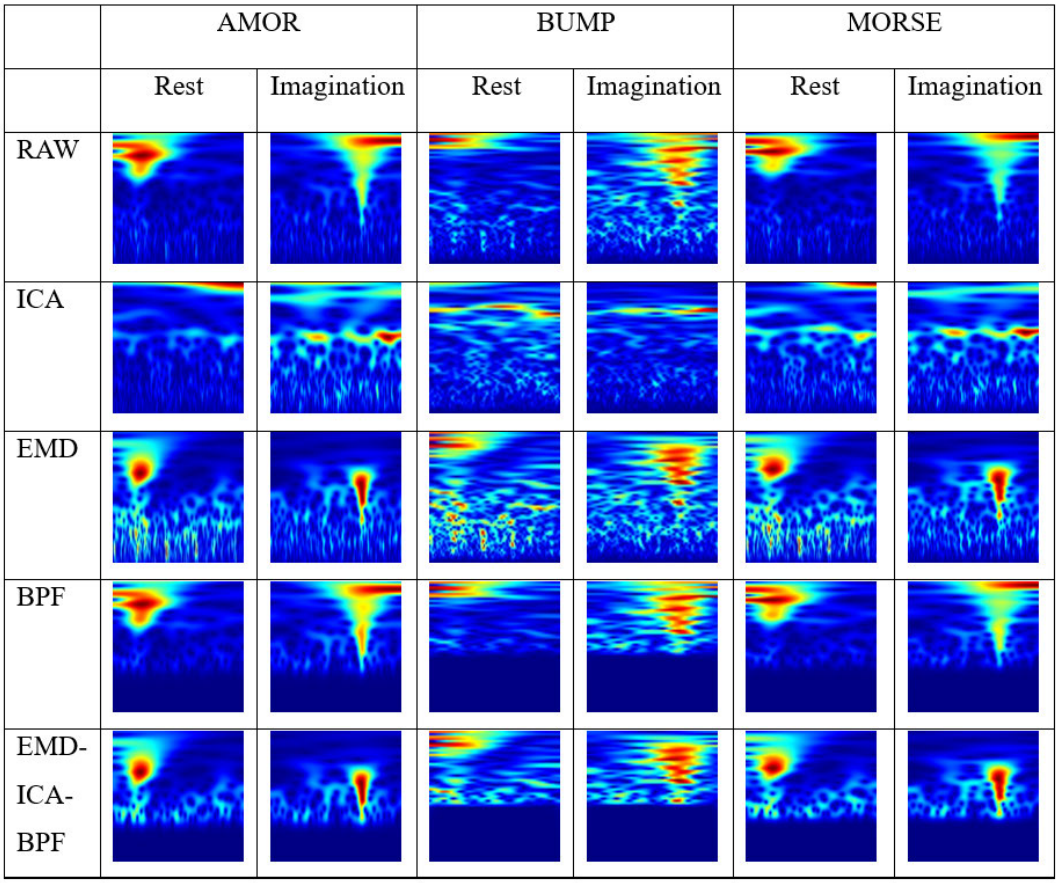

FIGURE 10. Diverse EEG data types and CWT visualization techniques.

TABLE 2. Natural swallowing dataset classification results.

Method	CHANNEL STACK	STFT-SPECTROGRAM	CWT SCALOGRAM		
			AMOR	BUMP	MORSE
RAW	ACC 0.572 AUC 0.598 F1 0.533	ACC 0.723 AUC 0.798 F1 0.73	ACC 0.839 AUC 0.924 F1 0.841	ACC 0.867 AUC 0.945 **F1 0.87**	ACC 0.856 AUC 0.936 F1 0.858
ICA	ACC 0.524 AUC 0.554 F1 0.556	ACC 0.719 AUC 0.8 F1 0.73	ACC 0.828 AUC 0.921 F1 0.837	ACC 0.867 AUC 0.943 F1 0.869	ACC 0.851 AUC 0.933 F1 0.843
EMD	ACC 0.551 AUC 0.571 F1 0.545	ACC 0.754 AUC 0.837 F1 0.755	ACC 0.855 AUC 0.94 F1 0.854	**ACC 0.873** **AUC 0.947** F1 0.852	ACC 0.864 AUC 0.941 F1 0.862
0.5-40 Hz Filter	ACC 0.568 AUC 0.587 F1 0.565	ACC 0.572 AUC 0.603 F1 0.571	ACC 0.7 AUC 0.771 F1 0.705	ACC 0.703 AUC 0.776 F1 0.701	ACC 0.707 AUC 0.776 F1 0.706
ICA 0.5-40 Hz EMD Filter	ACC 0.532 AUC 0.568 F1 0.563	ACC 0.584 AUC 0.62 F1 0.58	ACC 0.711 AUC 0.78 F1 0.709	ACC 0.693 AUC 0.766 F1 0.707	ACC 0.707 AUC 0.775 F1 0.715

determined in a sample of ten healthy volunteers and one chronic stroke patient. Additionally, swallowing and tongue movements share some common brain activation areas, and tongue movements are an integral part of the swallowing process [58], [59]. Yang et al. conducted a study with six healthy subjects using EEG to classify motor imagery tasks with a Support Vector Machine (SVM), achieving a cross-validation accuracy of 69.96% [57]. Yang et al. continued the study with six healthy subjects, achieving session-to-session accuracies of 74.29% and 72.12% using

TABLE 3. Induced saliva swallowing dataset classification results.

Method	CHANNEL STACK	STFT-SPECTROGRAM	CWT SCALOGRAM		
			AMOR	BUMP	MORSE
RAW	ACC 0.733 AUC 0.77 F1 0.716	ACC 0.806 AUC 0.887 F1 0.803	ACC 0.856 AUC 0.93 F1 0.855	ACC 0.854 **AUC 0.952** F1 0.866	ACC 0.855 AUC 0.936 F1 0.857
ICA	ACC 0.697 AUC 0.773 F1 0.707	ACC 0.797 AUC 0.887 F1 0.797	ACC 0.84 AUC 0.92 F1 0.839	ACC 0.863 AUC 0.949 F1 0.863	ACC 0.855 AUC 0.934 F1 0.856
EMD	ACC 0.569 AUC 0.608 F1 0.499	ACC 0.786 AUC 0.879 F1 0.776	ACC 0.859 AUC 0.937 F1 0.858	**ACC 0.872** **AUC 0.952** **F1 0.87**	ACC 0.858 AUC 0.943 F1 0.86
0.5-40 Hz Filter	ACC 0.714 AUC 0.786 F1 0.698	ACC 0.684 AUC 0.762 F1 0.676	ACC 0.748 AUC 0.824 F1 0.738	ACC 0.733 AUC 0.812 F1 0.731	ACC 0.747 AUC 0.822 F1 0.744
ICA 0.5-40 Hz EMD Filter	ACC 0.567 AUC 0.606 F1 0.577	ACC 0.603 AUC 0.651 F1 0.62	ACC 0.746 AUC 0.82 F1 0.748	ACC 0.735 AUC 0.809 F1 0.747	ACC 0.744 AUC 0.818 F1 0.747

TABLE 4. Induced water swallowing dataset classification results.

Method	CHANNEL STACK	STFT-SPECTROGRAM	CWT SCALOGRAM		
			AMOR	BUMP	MORSE
RAW	ACC 0.636 AUC 0.694 F1 0.637	ACC 0.81 AUC 0.893 F1 0.811	ACC 0.86 AUC 0.937 F1 0.861	**ACC 0.898** **AUC 0.967** **F1 0.902**	ACC 0.881 AUC 0.953 F1 0.882
ICA	ACC 0.662 AUC 0.721 F1 0.692	ACC 0.794 AUC 0.89 F1 0.774	ACC 0.847 AUC 0.934 F1 0.847	ACC 0.884 AUC 0.964 F1 0.887	ACC 0.86 AUC 0.947 F1 0.856
EMD	ACC 0.566 AUC 0.612 F1 0.63	ACC 0.819 AUC 0.904 F1 0.826	ACC 0.87 AUC 0.952 F1 0.87	**ACC 0.898** **AUC 0.967** F1 0.90	ACC 0.889 AUC 0.96 F1 0.893
0.5-40 Hz Filter	ACC 0.671 AUC 0.728 F1 0.662	ACC 0.70 AUC 0.773 F1 0.722	ACC 0.761 AUC 0.843 F1 0.765	ACC 0.755 AUC 0.839 F1 0.745	ACC 0.753 AUC 0.837 F1 0.756
ICA 0.5-40 Hz EMD Filter	ACC 0.651 AUC 0.721 F1 0.609	ACC 0.635 AUC 0.696 F1 0.65	ACC 0.759 AUC 0.833 F1 0.764	ACC 0.754 AUC 0.831 F1 0.757	ACC 0.764 AUC 0.838 F1 0.764

SVM classifiers and Dual-Tree Complex Wavelet Transform (DTCWT) features [58], [59]. Yang et al. expanded the study to include healthy and dysphagia patients, achieving cross-validation accuracies of 70.89% motor imagery swallowing and 73.79% imagery tongue movement for different motor imagery tasks [59]. Yang et al. analyzed motor imagery correlations using EEG and spectral power analysis, showing significant correlations between different motor imagery tasks in a sample of ten healthy subjects and one stroke patient [44], [60]. Considering these studies, it becomes evident that research on motor imagery of swallowing is relatively scarce and insufficient.

Based on the existing studies, we decided to design four different experimental paradigms. We adopted a careful approach in selecting the experimental paradigms for our study. Taking into account the gaps and needs in the literature, we investigated the relationship between swallowing and motor imagery. Additionally, we aimed to contribute to the development of rehabilitation tools by focusing on the specific needs of dysphagia patients. We also examined contemporary preprocessing techniques and imaging methods. Preprocessing techniques are critical for preparing data for cleaning, organizing, and analysis. Methods such as Empirical Mode Decomposition (EMD) and Independent Component Analysis (ICA) are commonly used to separate the components of EEG signals and filter out unwanted noise [61], [62]. Additionally, simple filtering techniques are utilized to isolate specific frequency intervals in the signal [63]. Visualization techniques are crucial for understanding the complexity and temporal changes of EEG data. For example, spectral analysis techniques like Short-Time Fourier Transform (STFT) and Continuous Wavelet Transform (CWT) are used to visualize the temporal distribution of different frequency components [64], [65]. These visualization techniques aid in better understanding the patterns and changing dynamics of brain activities.

TABLE 5. Induced tongue protruding dataset classification results.

Method	CHANNEL STACK	STFT-SPECTROGRAM	CWT SCALOGRAM		
			AMOR	BUMP	MORSE
RAW	ACC 0.764 AUC 0.832 F1 0.765	ACC 0.794 AUC 0.889 F1 0.794	ACC 0.854 AUC 0.935 F1 0.861	**ACC 0.881** **AUC 0.956** **F1 0.878**	ACC 0.861 AUC 0.943 F1 0.859
ICA	ACC 0.723 AUC 0.797 F1 0.728	ACC 0.794 AUC 0.88 F1 0.802	ACC 0.847 AUC 0.927 F1 0.847	ACC 0.874 AUC 0.951 F1 0.868	ACC 0.854 AUC 0.937 F1 0.86
EMD	ACC 0.592 AUC 0.628 F1 0.617	ACC 0.790 AUC 0.883 F1 0.801	ACC 0.86 AUC 0.942 F1 0.861	ACC 0.873 **AUC 0.956** F1 0.864	ACC 0.874 AUC 0.949 F1 0.876
0.5-40 Hz Filter	ACC 0.708 AUC 0.808 F1 0.713	ACC 0.685 AUC 0.758 F1 0.705	ACC 0.746 AUC 0.833 F1 0.764	ACC 0.741 AUC 0.821 F1 0.751	ACC 0.751 AUC 0.833 F1 0.758
ICA 0.5-40 Hz EMD Filter	ACC 0.572 AUC 0.608 F1 0.546	ACC 0.620 AUC 0.671 F1 0.633	ACC 0.741 AUC 0.820 F1 0.738	ACC 0.733 AUC 0.819 F1 0.742	ACC 0.742 AUC 0.819 F1 0.74

In conclusion, we leveraged advancements in technology to enhance deep learning techniques and examined the topic from different perspectives using various experimental paradigms. It also demonstrates the effectiveness of various preprocessing methods and visualization techniques in classifying resting and swallowing imagery, as well as in analyzing EEG signals. In evaluating the performance of our model, we opted to utilize three complementary metrics: AUC (Area Under the Curve), F1 score, and accuracy, to provide a comprehensive assessment of its effectiveness across different dimensions. AUC serves as a robust measure of the model's ability to distinguish between the two classes across various decision thresholds, making it particularly useful for understanding performance in the presence of class imbalance, which can occur in real-world scenarios. It captures the overall discriminatory power of the model, offering insights into its performance beyond a single threshold. The F1 score, on the other hand, harmonizes precision and recall, allowing us to evaluate the model's effectiveness in minimizing both false positives and false negatives—critical in applications where misclassifications can have significant consequences, such as in medical diagnoses or safety-critical systems. Lastly, accuracy provides an intuitive metric for assessing the overall correctness of the model's predictions. Given our balanced dataset, accuracy reliably indicates how well the model performs in classifying instances. By employing AUC, F1 score, and accuracy together, we gain a holistic understanding of our model's performance, enabling us to identify strengths and weaknesses in its predictive capabilities. This multi-faceted approach not only enhances our confidence in the model's effectiveness but also guides potential areas for future improvement.

In our current study, performance was evaluated within each experiment, ensuring that results were consistent within their specific contexts. Notably, we did not perform cross-experiment analyses, such as utilizing data from water swallowing to train and classify other swallowing imagery. This independent analysis approach allows for a clearer understanding of each experimental setup's effectiveness and minimizes potential confounding factors that might arise from integrating data across different contexts. These approaches can help provide a comprehensive and in-depth understanding of the subject. Identifying the limitations and constraints of the study is essential. These limitations include the sample size and diversity. In our study, we gathered data from 30 right-handed individuals, which is a sample size often utilized in EEG research. However, we recognize that this limited sample size may restrict the generalizability and validity of our findings. The limited sample size of the study may prevent the results from being more generalizable. While we view this study as a preliminary investigation, we recognize the limitations imposed by our restricted sample size. Moving forward, we plan to expand our participant group to include a larger and more diverse population. Additionally, we aim to conduct further studies to evaluate the performance of our model across various demographic groups. Additionally, some challenges may arise during data collection when using EEG and brain imaging techniques, which can affect the quality of the obtained data and make the interpretation of the results more difficult. Moreover, standardizing the experimental conditions used in the study is crucial. Non-standard conditions can affect the consistency of the results and reduce comparability. The absence of individuals with dysphagia in our study is also a limitation.

VI. CONCLUSION AND FUTURE WORK

The primary aim of this study was to assess the classification performance of deep learning models in distinguishing between resting state and motor imagery swallowing, utilizing various preprocessing and data visualization techniques applied to multichannel electroencephalography (EEG) data.

Firstly, the impact of preprocessing techniques on classification accuracy was examined. It was observed that the

use of methods like Empirical Mode Decomposition (EMD) and Independent Component Analysis (ICA) significantly influenced the filtering of EEG data. However, in some experiments, these techniques or their combination did not show the expected performance. Particularly, it was noted that bandpass filtering and hybrid models combining EMD, ICA, and bandpass filtering had lower accuracy rates compared to other methods in certain cases. These results emphasize the importance of selecting preprocessing methods considering the characteristics of EEG data and the classification tasks.

Secondly, the effects of different visualization techniques on the classification performance of EEG data were examined. Comparing visualization methods such as visualizing all channels together, individual Short-Time Fourier Transform (STFT) spectrograms, and Continuous Wavelet Transform (CWT) scalograms, it was shown that CWT images better capture the time-frequency characteristics. Specifically, it was generally found that "bump" type CWT scalograms outperform other types and effectively represent complex brain activities.

The effectiveness of these methods is critical in various fields such as the recognition of EEG-based neurological disorders, rehabilitation, and the development of brain-computer interface technologies. Particularly, the accurate identification of complex brain activities such as motor imagery and swallowing is vital for early rehabilitation and treatment of diseases.

Additionally, classification performance varies across different experimental paradigms, underscoring the importance of considering different scenarios and contexts in EEG data analysis. While EMD-applied data perform better in certain experiments, raw data perform better in others, highlighting the need for customized approaches depending on specific tasks and datasets.

In conclusion, this study evaluates the impact of different preprocessing and visualization techniques on the classification performance of EEG data. It offers 4 distinct experimental paradigms for analyzing swallowing events, along with a CNN model that produced superior results in testing existing experiments. Future research can further explore the optimization of preprocessing methods and visualization techniques to enhance the classification performance of EEG data.This study examines the potential of motor imagery (MI) in swallowing rehabilitation. Although the neuroplastic effects of motor imagery on swallowing have gained increasing attention in the literature, research in this area remains limited. Studies investigating the neural correlates of swallowing movements using motor imagery and classification approaches based on electroencephalogram (EEG) technologies indicate that further research is needed in this field. In this context, our study aims to contribute to the development of motor imagery classification systems, enhancing the effectiveness of MI in the treatment of swallowing disorders. Consequently, it can be concluded that the use of motor imagery holds significant potential in swallowing rehabilitation, and further innovative studies are needed in this area. These findings provide valuable insights for the development of EEG-based rehabilitation and therapeutic approaches and may lead to the creation of a device specifically designed to effectively interface with the small and complex muscles involved in swallowing events.

For future studies, we aim to work with dysphagia patients to better understand and improve the effectiveness of rehabilitation strategies tailored to this population. Additionally, by examining the effects of different motor imagery tasks on brain activity, we can gain a deeper understanding of the relationship between swallowing and tongue movements. The increased use of deep learning techniques can make data analysis processes more efficient and yield more precise results. Finally, by investigating the clinical application potential of our findings, we can evaluate how dysphagia patients can benefit in real-world conditions.

REFERENCES

[1] A. M. Bhutada, T. M. Davis, and K. L. Garand, "Electrophysiological measures of swallowing functions: A systematic review," *Dysphagia*, vol. 37, no. 6, pp. 1633–1650, Dec. 2022, doi: 10.1007/s00455-022-10426-4.

[2] M. Panebianco, R. Marchese-Ragona, S. Masiero, and D. A. Restivo, "Dysphagia in neurological diseases: A literature review," *Neurological Sci.*, vol. 41, no. 11, pp. 3067–3073, Nov. 2020, doi: 10.1007/s10072-020-04495-2.

[3] N. N. Ansari, M. Tarameshlu, and L. Ghelichi, "Dysphagia in multiple sclerosis patients: Diagnostic and evaluation strategies," *Degenerative Neurolo. Neuromuscular Disease*, vol. 10, pp. 15–28, Mar. 2020, doi: 10.2147/dnnd.s198659.

[4] J. Iruthayarajah, A. McIntyre, M. Mirkowski, P. Welch-West, E. Loh, and R. Teasell, "Risk factors for dysphagia after a spinal cord injury: A systematic review and meta-analysis," *Spinal Cord*, vol. 56, no. 12, pp. 1116–1123, Dec. 2018, doi: 10.1038/s41393-018-0170-3.

[5] J. E. Prasse and G. E. Kikano, "An overview of pediatric dysphagia," *Clin. Pediatrics*, vol. 48, no. 3, pp. 247–251, Apr. 2009, doi: 10.1177/0009922808327323.

[6] M. Inoue, "Dysphagia rehabilitation in Japan," *J. Nutritional Sci. Vitaminol.*, vol. 61, pp. S72–S73, 2015, doi: 10.3177/jnsv.61.s72.

[7] A. Vogel, I. Claus, S. Ahring, D. Gruber, A. Haghikia, U. Frank, R. Dziewas, G. Ebersbach, F. Gandor, and T. Warnecke, "Endoscopic characteristics of dysphagia in multiple system atrophy compared to Parkinson's disease," *Movement Disorders*, vol. 37, no. 3, pp. 535–544, Mar. 2022, doi: 10.1002/mds.28854.

[8] T. T. N. Nguyen, T. T. Dam, C. C. Le, D. T. Do, and D. L. Vu, "Evaluation of oropharyngeal dysphagia using the video-fluoroscopic swallowing study in stroke patients," *Vietnamese J. Radiol. Nucl. Med.*, no. 3, pp. 17–21, Dec. 2023, doi: 10.55046/vjrnm.03.1068.2023.

[9] C. D. Binnie and P. F. Prior, "Electroencephalography.," *J. Neurol., Neurosurgery Psychiatry*, vol. 57, no. 11, pp. 1308–1319, Nov. 1994, doi: 10.1136/jnnp.57.11.1308.

[10] D. Attwell and S. B. Laughlin, "An energy budget for signaling in the grey matter of the brain," *J. Cerebral Blood Flow Metabolism*, vol. 21, no. 10, pp. 1133–1145, Oct. 2001, doi: 10.1097/00004647-200110000-00001.

[11] G. Muehllehner and J. S. Karp, "Positron emission tomography," *Phys. Med. Biol.*, vol. 51, no. 13, pp. R117–R137, Jul. 2006, doi: 10.1088/0031-9155/51/13/r08.

[12] C.-H. Han, K.-R. Müller, and H.-J. Hwang, "Enhanced performance of a brain switch by simultaneous use of EEG and NIRS data for asynchronous brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 10, pp. 2102–2112, Oct. 2020, doi: 10.1109/TNSRE.2020.3017167.

[13] M. Proudfoot, M. W. Woolrich, A. C. Nobre, and M. R. Turner, "Magnetoencephalography," *Practical Neurol.*, vol. 14, no. 5, pp. 336–343, Oct. 2014, doi: 10.1136/practneurol-2013-000768.

[14] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain-computer interfaces in the completely locked-in state and chronic stroke," *Prog. Brain Res.*, vol. 228, pp. 131–161, Jan. 2016, doi: 10.1016/bs.pbr.2016.04.019.

[15] S. E. Kober and G. Wood, "Changes in hemodynamic signals accompanying motor imagery and motor execution of swallowing: A near-infrared spectroscopy study," *NeuroImage*, vol. 93, pp. 1–10, Jun. 2014, doi: 10.1016/j.neuroimage.2014.02.019.

[16] S. E. Kober, G. Bauernfeind, C. Woller, M. Sampl, P. Grieshofer, C. Neuper, and G. Wood, "Hemodynamic signal changes accompanying execution and imagery of swallowing in patients with dysphagia: A multiple single-case near-infrared spectroscopy study," *Frontiers Neurol.*, vol. 6, pp. 1–10, Jul. 2015, doi: 10.3389/fneur.2015.00151.

[17] S. E. Kober, B. Gressenberger, J. Kurzmann, C. Neuper, and G. Wood, "Voluntary modulation of hemodynamic responses in swallowing related motor areas: A near-infrared spectroscopy-based neurofeedback study," *PLoS ONE*, vol. 10, no. 11, Nov. 2015, Art. no. e0143314, doi: 10.1371/journal.pone.0143314.

[18] S. E. Kober, "Hemodynamic signal changes during saliva and water swallowing: A near-infrared spectroscopy study," *J. Biomed. Opt.*, vol. 23, no. 1, p. 1, Jan. 2018, doi: 10.1117/1.jbo.23.1.015009.

[19] I. Jestrović, J. L. Coyle, and E. Sejdić, "Decoding human swallowing via electroencephalography: A state-of-the-art review," *J. Neural Eng.*, vol. 12, no. 5, Oct. 2015, Art. no. 051001, doi: 10.1088/1741-2560/12/5/051001.

[20] G. A. Malandraki, B. P. Sutton, A. L. Perlman, D. C. Karampinos, and C. Conway, "Neural activation of swallowing and swallowing-related tasks in healthy young adults: An attempt to separate the components of deglutition," *Human Brain Mapping*, vol. 30, no. 10, pp. 3209–3226, Oct. 2009, doi: 10.1002/hbm.20743.

[21] H. Yang, K. K. Ang, C. Wang, K. S. Phua, and C. Guan, "Neural and cortical analysis of swallowing and detection of motor imagery of swallow for dysphagia rehabilitation a review," *Prog. Brain Res.*, vol. 228, pp. 185–219, Jan. 2016, doi: 10.1016/bs.pbr.2016.03.014.

[22] M. L. Huckabee, L. Deecke, M. P. Cannito, H. J. Gould, and W. Mayr, "Cortical control mechanisms in volitional swallowing: The bereitschaftspotential," *Brain Topography*, vol. 16, no. 1, pp. 3–17, 2003, doi: 10.1023/A:1025671914949.

[23] T. Nonaka, M. Yoshida, T. Yamaguchi, A. Uchida, H. Ohba, S. Oka, and I. Nakajima, "Contingent negative variations associated with command swallowing in humans," *Clin. Neurophysiol.*, vol. 120, no. 10, pp. 1845–1851, Oct. 2009, doi: 10.1016/j.clinph.2009.06.029.

[24] K. Hiraoka, "Movement-related cortical potentials associated with saliva and water bolus swallowing," *Dysphagia*, vol. 19, no. 3, pp. 155–159, Aug. 2004, doi: 10.1007/s00455-004-0002-9.

[25] M. Cuellar, A. W. Harkrider, D. Jenson, D. Thornton, A. Bowers, and T. Saltuklaroglu, "Time-frequency analysis of the EEG mu rhythm as a measure of sensorimotor integration in the later stages of swallowing," *Clin. Neurophysiol.*, vol. 127, no. 7, pp. 2625–2635, Jul. 2016, doi: 10.1016/j.clinph.2016.04.027.

[26] S. Koganemaru, F. Mizuno, T. Takahashi, Y. Takemura, H. Irisawa, M. Matsuhashi, T. Mima, T. Mizushima, and K. Kansaku, "Event-related desynchronization and corticomuscular coherence observed during volitional swallow by electroencephalography recordings in humans," *Frontiers Human Neurosci.*, vol. 15, pp. 1–9, Nov. 2021, doi: 10.3389/fnhum.2021.643454.

[27] I. Jestrović, J. L. Coyle, and E. Sejdić, "The effects of increased fluid viscosity on stationary characteristics of EEG signal in healthy adults," *Brain Res.*, vol. 1589, pp. 45–53, Nov. 2014, doi: 10.1016/j.brainres.2014.09.035.

[28] I. Jestrović, J. L. Coyle, and E. Sejdić, "Characterizing functional connectivity patterns during saliva swallows in different head positions," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, pp. 1–11, Dec. 2015, doi: 10.1186/s12984-015-0049-x.

[29] I. Jestrović, J. L. Coyle, S. Perera, and E. Sejdić, "Functional connectivity patterns of normal human swallowing: Difference among various viscosity swallows in normal and chin-tuck head positions," *Brain Res.*, vol. 1652, pp. 158–169, Dec. 2016, doi: 10.1016/j.brainres.2016.09.041.

[30] I. Jestrović, J. L. Coyle, and E. Sejdić, "Differences in brain networks during consecutive swallows detected using an optimized vertex–frequency algorithm," *Neuroscience*, vol. 344, pp. 113–123, Mar. 2017, doi: 10.1016/j.neuroscience.2016.11.047.

[31] I. Jestrović, J. L. Coyle, S. Perera, and E. Sejdić, "Influence of attention and bolus volume on brain organization during swallowing," *Brain Struct. Function*, vol. 223, no. 2, pp. 955–964, Mar. 2018, doi: 10.1007/s00429-017-1535-7.

[32] J. Dinarès-Ferran, R. Ortner, C. Guger, and J. Solé-Casals, "A new method to generate artificial frames using the empirical mode decomposition for an EEG-based motor imagery BCI," *Frontiers Neurosci.*, vol. 12, pp. 1–9, May 2018, doi: 10.3389/fnins.2018.00308.

[33] A. S. C. Caldas, W. K. Coelho, R. F. G. Ribeiro, D. A. D. Cunha, and H. J. D. Silva, "Motor imagery and swallowing: A systematic literature review," *Revista CEFAC*, vol. 20, no. 2, pp. 247–257, Apr. 2018, doi: 10.1590/1982-0216201820214317.

[34] M. Naeem, C. Brunner, R. Leeb, B. Graimann, and G. Pfurtscheller, "Seperability of four-class motor imagery data using independent components analysis," *J. Neural Eng.*, vol. 3, no. 3, pp. 208–216, Sep. 2006, doi: 10.1088/1741-2560/3/3/003.

[35] P. Gaur, R. B. Pachori, H. Wang, and G. Prasad, "An empirical mode decomposition based filtering method for classification of motor-imagery EEG signals for enhancing brain–computer interface," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2015, pp. 1–7, doi: 10.1109/IJCNN.2015.7280754.

[36] T. D. Pham, "Classification of motor-imagery tasks using a large EEG dataset by fusing classifiers learning on wavelet-scattering features," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 1097–1107, 2023, doi: 10.1109/TNSRE.2023.3241241.

[37] A. Yildiz, H. Zan, and S. Said, "Classification and analysis of epileptic EEG recordings using convolutional neural network and class activation mapping," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102720, doi: 10.1016/j.bspc.2021.102720.

[38] S. Chaudhary, S. Taran, V. Bajaj, and A. Sengur, "Convolutional neural network based approach towards motor imagery tasks EEG signals classification," *IEEE Sensors J.*, vol. 19, no. 12, pp. 4494–4500, Jun. 2019, doi: 10.1109/JSEN.2019.2899645.

[39] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, Jun. 2019, Art. no. 031001, doi: 10.1088/1741-2552/ab0ab5.

[40] A. Sharmila, S. Madan, and K. Srivastava, "Epilepsy detection using DWT based Hurst exponent and SVM, K-NN classifiers," *Serbian J. Experim. Clin. Res.*, vol. 19, no. 4, pp. 311–319, Dec. 2018, doi: 10.1515/sjecr-2017-0043.

[41] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, pp. 1–74, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[42] W.-L. Mao, H. I. K. Fathurrahman, Y. Lee, and T. W. Chang, "EEG dataset classification using CNN method," *J. Phys., Conf. Ser.*, vol. 1456, no. 1, Jan. 2020, Art. no. 012017, doi: 10.1088/1742-6596/1456/1/012017.

[43] M. Rashid, N. Sulaiman, A. P. P. A. Majeed, R. M. Musa, A. F. A. Nasir, B. S. Bari, and S. Khatun, "Current status, challenges, and possible solutions of EEG-based brain–computer interface: A comprehensive review," *Frontiers Neurorobotics*, vol. 14, p. 25, Jun. 2020, doi: 10.3389/fnbot.2020.00025.

[44] H. Yang, C. Guan, C. C. Wang, K. K. Ang, K. S. Phua, S. S. Chok, C. K. Y. Tang, and K. S. G. Chua, "On the correlations of motor imagery of swallow with motor imagery of tongue movements and actual swallow," in *5th Adv. Cogn. Neurodynamics (V) Proc. Int. erence on Cogn. Neurodynam*, Singapore. Cham, Switzerland: Springer, Jan. 2016, pp. 397–404, doi: 10.1007/978-981-10-0207-6.

[45] D.-X. Zhang, X.-P. Wu, and X.-j. Guo, "The EEG signal preprocessing based on empirical mode decomposition," in *Proc. 2nd Int. Conf. Bioinf. Biomed. Eng.*, May 2008, pp. 2131–2134, doi: 10.1109/ICBBE.2008.862.

[46] S. Sadeghi and A. Maleki, "The empirical mode decomposition-decision tree method to recognize the steady-state visual evoked potentials with wide frequency range," *J. Med. Signals Sensors*, vol. 8, no. 4, pp. 225–230, Dec. 2018, doi: 10.4103/jmss.JMSS_20_18.

[47] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004, doi: 10.1016/j.jneumeth.2003.10.009.

[48] B. Mandhouj, M. A. Cherni, and M. Sayadi, "An automated classification of EEG signals based on spectrogram and CNN for epilepsy diagnosis," *Anal. Integr. Circuits Signal Process.*, vol. 108, no. 1, pp. 101–110, Jul. 2021, doi: 10.1007/s10470-021-01805-2.

[49] Ö. Türk and M. S. Özerdem, "Epilepsy detection by using scalogram based convolutional neural network from EEG signals," *Brain Sci.*, vol. 9, no. 5, p. 115, May 2019, doi: 10.3390/brainsci9050115.

[50] Z. Khademi, F. Ebrahimi, and H. M. Kordy, "A review of critical challenges in MI-BCI: From conventional to deep learning methods," *J. Neurosci. Methods*, vol. 383, Jan. 2023, Art. no. 109736, doi: 10.1016/j.jneumeth.2022.109736.

[51] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electron. Markets*, vol. 31, no. 3, pp. 685–695, Sep. 2021, doi: 10.1007/s12525-021-00475-2.

[52] Y. Mohan, S. S. Chee, D. K. P. Xin, and L. P. Foong, "Artificial neural network for classification of depressive and normal in EEG," in *Proc. IEEE EMBS Conf. Biomed. Eng. Sci. (IECBES)*, Dec. 2016, pp. 286–290, doi: 10.1109/IECBES.2016.7843459.

[53] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013, doi: 10.1088/1741-2552/aace8c.

[54] N. Güler, E. Übeyli, and I. Güler, "Recurrent neural networks employing Lyapunov exponents for EEG signals classification," *Expert Syst. Appl.*, vol. 29, no. 3, pp. 506–514, Oct. 2005, doi: 10.1016/j.eswa.2005.04.011.

[55] J. J. Bird, D. R. Faria, L. J. Manso, P. P. S. Ayrosa, and A. Ekárt, "A study on CNN image classification of EEG signals represented in 2D and 3D," *J. Neural Eng.*, vol. 18, no. 2, Apr. 2021, Art. no. 026005, doi: 10.1088/1741-2552/abda0c.

[56] J. Zhang and M. Wang, "A survey on robots controlled by motor imagery brain–computer interfaces," *Cognit. Robot.*, vol. 1, pp. 12–24, Jan. 2021, doi: 10.1016/j.cogr.2021.02.001.

[57] H. Yang, C. Guan, K. K. Ang, C. C. Wang, K. S. Phua, and J. Yu, "Dynamic initiation and dual-tree complex wavelet feature-based classification of motor imagery of swallow EEG signals," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jun. 2012, pp. 1–6, doi: 10.1109/IJCNN.2012.6252603.

[58] H. Yang, C. Guan, K. K. Ang, C. Wang, K. S. Phua, C. T. K. Yin, and L. Zhou, "Feature consistency-based model adaptation in session-to-session classification: A study using motor imagery of swallow EEG signals," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 429–432, doi: 10.1109/EMBC.2013.6609528.

[59] H. Yang, C. Guan, K. S. G. Chua, S. S. Chok, C. C. Wang, P. K. Soon, C. K. Y. Tang, and K. K. Ang, "Detection of motor imagery of swallow EEG signals based on the dual-tree complex wavelet transform and adaptive model selection," *J. Neural Eng.*, vol. 11, no. 3, Jun. 2014, Art. no. 035016, doi: 10.1088/1741-2560/11/3/035016.

[60] H. Yang, C. Guan, C. C. Wang, and K. K. Ang, "Detection of motor imagery of brisk walking from electroencephalogram," *J. Neurosci. Methods*, vol. 244, pp. 33–44, Apr. 2015, doi: 10.1016/j.jneumeth.2014.05.007.

[61] A.-O. Boudraa and J.-C. Cexus, "EMD-based signal filtering," *IEEE Trans. Instrum. Meas.*, vol. 56, no. 6, pp. 2196–2202, Dec. 2007, doi: 10.1109/TIM.2007.907967.

[62] L. Sun, Y. Liu, and P. J. Beadle, "Independent component analysis of EEG signals," in *Proc. IEEE Int. Workshop VLSI Design Video Technol.*, May 2005, pp. 219–222, doi: 10.1109/IWVDVT.2005.1504590.

[63] R. Nath Bairagi, M. Maniruzzaman, S. Pervin, and A. Sarker, "Epileptic seizure identification in EEG signals using DWT, ANN and sequential window algorithm," *Soft Comput. Lett.*, vol. 3, Dec. 2021, Art. no. 100026, doi: 10.1016/j.socl.2021.100026.

[64] P. Peng, Y. Song, L. Yang, and H. Wei, "Seizure prediction in EEG signals using STFT and domain adaptation," *Frontiers Neurosci.*, vol. 15, Jan. 2022, Art. no. 825434, doi: 10.3389/fnins.2021.825434.

[65] P. Kant, S. H. Laskar, J. Hazarika, and R. Mahamune, "CWT based transfer learning for motor imagery classification for brain computer interfaces," *J. Neurosci. Methods*, vol. 345, Nov. 2020, Art. no. 108886, doi: 10.1016/j.jneumeth.2020.108886.

SEVGI GÖKÇE ASLAN received the dual B.Sc. degrees in biomedical engineering and electrical and electronics engineering from Erciyes University, Kayseri, Türkiye, in 2017 and 2020, respectively. She is currently pursuing the integrated Ph.D. degree in electrical and computer engineering with Abdullah Gül University, Kayseri. Since 2020, she has been a Research Assistant with Inonu University, Malatya, Türkiye. Her current research interests include machine learning, deep learning, brain–computer interfaces, biomedical signal processing, and biomedical engineering.

BÜLENT YILMAZ received the B.Sc. and M.Sc. degrees in electrical-electronics engineering from Middle East Technical University, Ankara, Türkiye, in 1997 and 1999, respectively, and the Ph.D. degree from the Bioengineering Department, The University of Utah, Salt Lake, UT, USA. He is currently a Professor with the Electrical Engineering Department, Gulf University for Science and Technology, Hawally, Kuwait. His current research interests include biomedical signal and image processing applications on brain–computer interfaces and cancer and disease assessment.