IRO Journal on Sustainable Wireless Systems (ISSN: 2582-3167)
www.irojournals.com/irosws/

EEG based Imagined Speech Classification using CWT and ResNet-50-D Architecture

Immanuel G.[1], Shirly S.[2*]

Division of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, India.

E-mail: [1]immanuelg@karunya.edu.in, [2*]shirly@karunya.edu

Abstract

Imagined speech recognition is considered a novel BCI technique where the human brain can directly communicate with the external world without the need for speech production. Nevertheless, decoding the imagined speech using EEG data is problematic due to the non-stationary properties, signal-to-noise ratio issues, and subject-dependent nature of EEG. The current paper introduces a classification method for imagined speech using EEG signals based on the combination of Continuous Wavelet Transform and the ResNet-50-D deep learning architecture. Firstly, EEG data from the ASU Imagined Speech EEG Dataset are preprocessed to eliminate noise and artifacts. Then, the preprocessed EEG data are converted into the time–frequency representation using CWT with Morlet wavelet function, thus, allowing effective representation of time and frequency components of neural activity. The resulting representations are used as input data for ResNet-50-D. The designed scheme is tested on four different hypothetical speech types, namely vowels, short words, long words, and short-long word combinations. From the experiments, accuracies of classification achieved are 72.4%, 74.0%, 77.5%, and 76.4% correspondingly. The comparative study of the proposed approach with previous works shows the high efficiency of the combination of time-frequency features representation and residual deep learning. The results show that the proposed CWT-ResNet-50-D approach is an efficient and reliable method for decoding imagined speech from EEG and can be further used in BCI systems.

*** Corresponding Author**

DOI: https://doi.org/10.36548/jsws.2026.3.007
Extended version of article received from **International Conference on Smart Innovations in Engineering & Technology (ICSIET 2026)**
Received: 24.05.2026, received in revised form: 30.06.2026, accepted: 18.07.2026, published: 27.07.2026

Keywords: Brain–Computer Interface, Electroencephalography, Imagined Speech Classification, Continuous Wavelet Transform, ResNet-50-D, Deep Learning, Time–Frequency Analysis.

1. Introduction

The Brain–Computer Interface (BCI) is a technological approach that involves direct interactions between the human brain and external devices through interpretation of neural activities without using muscle actions. In many applications of BCI, imagined speech decoding has drawn much interest due to its capacity to offer an alternative pathway of communication for patients suffering from severe motor disorders, locked-in syndrome, and speech disabilities. While Motor imagery BCIs seek to detect actions, imagined speech decoding tries to detect words or phonemes that are silently uttered in the mind. Decoding of EEG signals containing imagined speech is difficult due to its high variability, non-stationarity, and subject-dependence.

Recent developments in signal processing and deep learning have enhanced the performance of EEG-based classification of imagined speech. The conventional machine learning algorithm is based on handcrafted features which are not able to detect the complex nature of time and spectral information in EEG signals. On the contrary, a deep learning approach will be able to extract discriminative features from EEG signals, resulting in better classification performance. However, its performance is contingent upon feature extraction techniques and detection of subtle neural features that can differentiate the imagined speech. Given the information about time and frequency available in EEG signals, time–frequency analysis is an important step in feature extraction. The Continuous Wavelet Transform (CWT) is very appropriate for the analysis of non-stationary EEG signals, as it offers a detailed time-frequency decomposition while maintaining transient properties of the brain. Using CWT in order to transform one-dimensional EEG signals into 2D images allows the use of state-of-the-art image processing deep learning algorithms. Among different wavelets, the Morlet wavelet is commonly used due to its ability to detect localized changes in the spectrum of neural signals [8]. Convolutional neural networks have gained great success in the field of image classification with their automatic feature extraction capabilities. In order to overcome the degradation problem in deep architectures, residual networks utilize skip connections that make

gradient flow more effective. The advanced ResNet-50-D is an improved version of the traditional ResNet architecture and can be applied on CWT-based EEG scalograms [10].

Based on the benefits of this system, the current research introduces an EEG-based imagined speech classification framework which utilizes the combination of Continuous Wavelet Transform (CWT) and ResNet-50-D. Firstly, the EEG data are preprocessed and converted into time-frequency scalograms using the CWT technique. The scalograms are then fed into the ResNet-50-D model to extract features and classify the input data. The current approach is tested on various categories of imagined speech, including vowels, short words, long words, and short-long words.

The main contributions made by this research work include the design of a Continuous Wavelet Transform (CWT) preprocessing technique to obtain time-frequency representation of EEG signals that can be used to classify the imagined speech. Also, a ResNet-50-D deep neural network model has been used to automatically classify discriminating features of the CWT based scalogram to classify imagined speech. The developed framework is tested on different types of imagined speech and has shown better classification accuracy than many existing techniques.

2. Literature Review

The evolution of Brain Computer Interface (BCI) systems has greatly enhanced imagined speech recognition via electroencephalography (EEG). Speech imagination decoding involves converting the neurological information related to inner speech into useful outputs. It provides a channel of communication for people suffering from severe speech and motor disabilities. Previous works have utilized Riemannian manifold features derived from EEG signals to classify the imagined speech patterns, proving that imagined speech decoding is achievable due to its complexity and non-stationarity [1]. The success of an EEG based BCI system requires efficient signal processing methods that can deal with noise and variability in the signals. Adaptive Signal Decomposition coupled with machine learning has been utilized to enhance EEG feature extraction and classification performance [2]. Additionally, functional connectivity analysis has been applied to model the communication between brain areas using connectivity networks in order to enhance classification of imagined speech [3].

The time-frequency analysis has become the most popular technique in dealing with non-stationary EEG data. The wavelet-based analysis was widely used in extracting discriminative features associated with the speech imagination. Wavelet-augmented phase coherence techniques demonstrated better robustness and classification performances, while Morlet wavelets were proved to be efficient in modeling the time-frequency features of EEG data and transient dynamics of the neural system [4], [8]. The deep learning techniques have attracted increasing attention in decoding the imagined speech due to the capability of the automatic learning of hierarchically structured features. The frameworks of deep feature learning have efficiently performed in extracting the spatiotemporal features from EEG data and outperformed classical machine learning algorithms [5]. In addition, deep neural networks with the CSP channel selection and DWT features have also obtained good classification results [6]. To address the limitation of small EEG datasets, the data augmentation methods based on CSP were developed [7].

In recent times, there has been additional validation on the efficiency of machine learning and deep learning in decoding the signals of imagined speech [11]. CNNs have shown remarkable proficiency in discriminating important features from images in biomedical applications [9]. In addition to that, advancements in residual network architecture have increased feature preservation and optimization stability, and accuracy, making them effective in handling wavelet representation of EEG data [10].

However, due to the issues like the variability of the signal, small dataset size, and the complexity of the neural activity of speech, imagined speech classification continues to be a challenging task. It is found that a combination of efficient time-frequency analysis and effective deep learning models would greatly enhance the performance. This paper takes its inspiration from such studies and makes use of Continuous Wavelet Transform (CWT) for obtaining the time-frequency representation of EEG signals and uses ResNet-50-D architecture for learning discriminatory features for imagined speech classification.

3. Proposed Methodology

3.1 Dataset Description

The EEG data that were employed in the current research were downloaded from the publicly accessible Arizona State University (ASU) Imagined Speech EEG Dataset [12]. The

database comprises EEG data collected from healthy subjects who carried out imagined speech tasks with various speech types, which include vowels, short words, long words, and short-long word types. EEG data were collected from a 64-channel recording setup using the international 10-20 electrode configuration. It is worth noting that the ASU database has gained widespread use in imagined speech decoding studies owing to the variety of speech prompts and applicability for machine/deep learning techniques in EEG-based speech recognition [12].

The data set consists of four speech classes having different degrees of linguistic complexity including vowels, short words, long words, and short-long words. These speech classes have been selected in order to test the capability of the proposed framework for discrimination between simple phonetic units and complex imagined speech patterns. The inclusion of different speech classes makes it possible to evaluate the classification capabilities of the framework in various conditions.

Table 1. Distribution of participants and imagined speech categories in the ASU EEG dataset

Word Classes	Prompts	No. of subjects	Subject's ID
Short words	in, out, up	6	S1, S3, S5, S6, S8 and S12
Long words	Independent, cooperate	6	S2, S3, S6, S7, S9 and S11
Short-long words	In, cooperate	6	S1, S5, S8, S9, S10 and S14
Vowels	a, i, u	8	S4, S5, S8, S9, S11, S12, S13 and S15

Table 1 shows the distribution of the participants among the imaginary speech types that have been employed in the dataset. The vowels type comprises the imaginary words "a," "i," and "u," which are considered basic speech sounds, whose brain activity is fairly simple. The short-word type is made up of the words "in," "out," and "up," while the long-word type has the words "independent" and "cooperate," which need complicated linguistics processing when imagining. The short-long word type comprises the words "in" and "cooperate".

The diverse speech prompting and participation of individuals in the different categories forms a strong basis that can be used to measure the performance of EEG-based imagined speech recognition systems. In addition, the use of this dataset makes it possible to test the ability of the models to generalize under various levels of complexity of the speech.

3.2 Proposed Framework

At first, EEG signals are recorded and preprocessed by applying filtering and artifacts removal steps on the collected EEG signals. After preprocessing, the time–frequency representations of the EEG signals are created through Continuous Wavelet Transform (CWT). The scalograms obtained from CWT will be used as image inputs to ResNet-50-D model to perform classification based on spatial feature extraction. The designed framework can discriminate between different classes of imagined speech such as vowels, short words, long words, and short–long words.

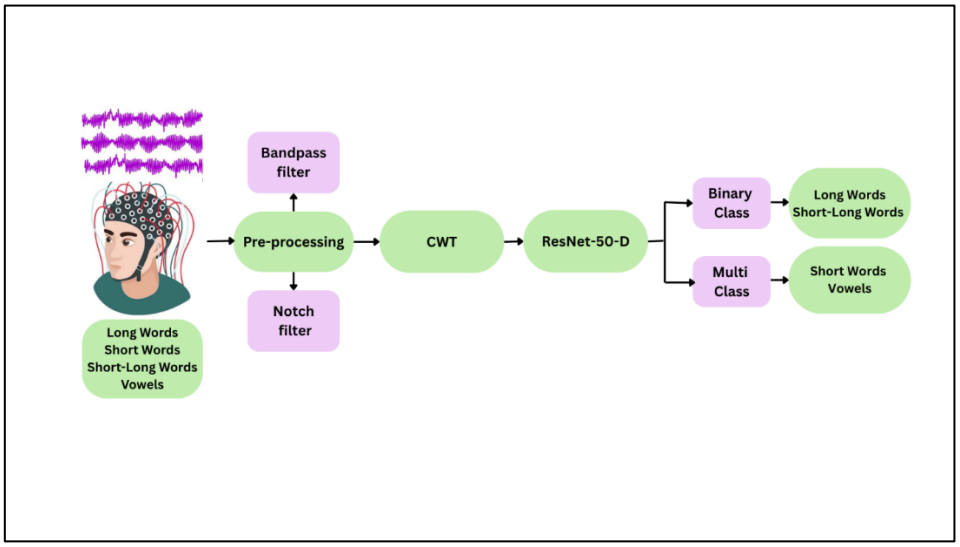

Figure 1. Overall framework of the proposed EEG-based imagined speech classification system

Figure 1 shows an architecture for imaginary speech classification using EEG data. According to the architecture, first, the EEG data is acquired and pre-processed through applying band-pass and notch filters. Then, they are converted into time-frequency data by means of continuous wavelet transform (CWT). Finally, feature extraction and classification are performed through a deep model which is based on ResNet-50-D. The proposed model can be applied to binary as well as multi-class classifications for differentiation of various imagined speeches such as long words, short words, and vowels.

3.3 EEG Preprocessing

EEG readings are intrinsically prone to several types of noise and artifacts, which include eye movement, muscle activity, and electrical noise from power lines. In order for the

reading to be analyzed reliably, an extensive preprocessing step is done before feature extraction. Initially, a band-pass filter is used to retain the frequency that is related to brain activity and eliminate others that do not contribute to the desired result. Next, a notch filter is used to eliminate the electrical noise coming from power lines. Finally, ICA is used to eliminate the artifacts, especially the ones related to eye movement and muscles.

The process of preprocessing is highly important for increasing the quality of EEG signals prior to the stage of feature extraction and classification. The EEG data can be contaminated with different artifacts like eye blinks, muscular activities, and external noise, which may cover up the useful information about the brain state. With the help of filtering, artifact cancellation, and normalization procedures, it is possible to reduce the unnecessary components of the signal and save the essential patterns of brain activity. Such processing increases the signal-to-noise ratio and provides more uniform representation of data.

3.4 Continuous Wavelet Transform

Considering the fact that EEG signals are non-stationary signals, traditional time domain analysis is not sufficient for analyzing transient neural activities. Continuous Wavelet Transform (CWT) is a useful time-frequency analysis tool which decomposes EEG signals in different scales and frequencies at the same time.

The CWT of a signal s(t) is defined as (1):

$$W_{\psi}s(x,y) = \frac{1}{\sqrt{x}} \int_{-\infty}^{\infty} s(t)\psi^{*}\left(\frac{t-y}{x}\right) dt \quad (1)$$

Time-frequency representation of the EEG is generated by the use of the complex Morlet wavelet (2), which serves as the base function because it is commonly used in analyzing the time-frequency features of neuroelectric signals like EEG.

$$\psi(t) = \frac{1}{\sqrt{\pi f_y}}\, e^{2i\pi f_z t}\, e^{-t^2/f_y} \quad (2)$$

Where,

- f_y is the bandwidth of the parameter
- f_z is the wavelet of the center frequency.

The scalogram obtained can then be used to visualize energy changes in the frequency domain over time to aid learning discriminative patterns for imagined speech.

3.5 ResNet-50-D Architecture

The resulting CWT scalogram images are then classified via ResNet-50-D neural network. ResNet-50-D represents a more sophisticated variant of traditional ResNet-50 model with improved design to increase feature extraction and classification efficiency.

In contrast to the traditional ResNet-50 model, ResNet-50-D has made certain changes in its architecture in order to facilitate the better preservation of the features and learning efficiency. The 7 × 7 convolution layer has been replaced by the three 3 × 3 convolution layers in order to extract the fine-grained spatial features efficiently while minimizing information loss. Further, the architecture makes use of residual skip connections which enable efficient backpropagation during training.

The architecture involves four residual stages that include bottleneck blocks with increasing numbers of channels. The global average pooling technique is used for the creation of compact features that are then passed through fully connected layers with the dropout technique.

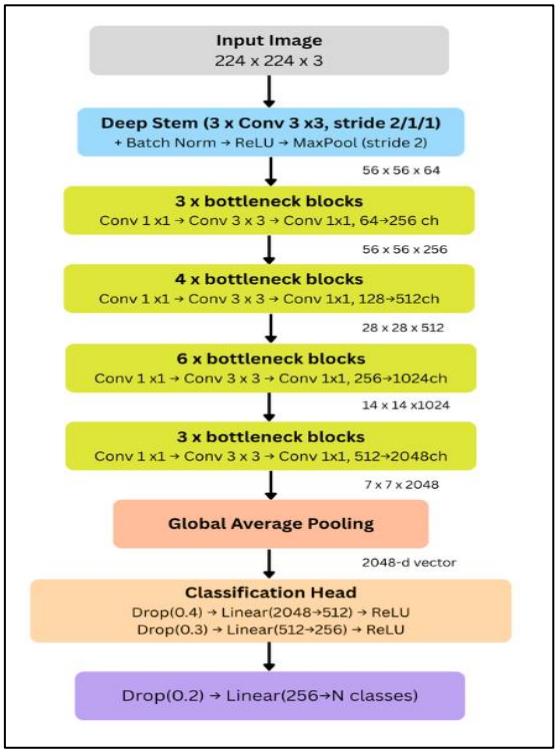

Figure 2. Architecture of the ResNet-50-D network

As can be seen from Figure 2, the "D" variation of ResNet50 has a different stem compared to the regular ResNet50 architecture. The difference is that it consists of three consecutive 3×3 convolutions having 2/1/1 strides instead of one big 7×7 conv. Such an arrangement helps to reduce aliasing and obtain better early feature maps. Then there is batch norm, ReLU and stride-2 MaxPool that downscale the initial 224×224 image to 56×56×64. In stages 2-5 there are stacks of bottleneck blocks. Each bottleneck block has 3 convolutions: 1×1 to compress the number of channels, 3×3 to obtain spatial features and 1×1 to upscale the number of channels. Also, there is residual skip connection across the whole triplet. In stage 2 there are 3 blocks (256 ch), in stage 3 there are 4 blocks (512 ch), in stage 4 there are 6 blocks (1024 ch), in stage 5 there are 3 blocks (2048 ch). By stage 5 spatial resolution is reduced by a few downsampling operations, thus, it is equal to 7×7. This results in a single 2048-dimensional vector representation of the spectrogram obtained through global average pooling on the 7×7×2048 tensor. The proposed head employs three blocks of Dropout, Linear, ReLU (Dropout rates 0.4, 0.3 and 0.2) to progressively reduce the logits to 512, 256 and finally N dimensions. The head is strongly regularized at the top.

3.6 Classification Procedure and Performance Evaluation

The scalograms generated using CWT act as input images for the ResNet-50-D classifier, which is responsible for classifying the imagined speech EEG signals. In the training stage, the network is able to learn hierarchical and discriminative features from the time-frequency representation contained in the scalograms. These features help classify the different imagined speech classes such as vowels, short words, long words, and short-long word combination.

Data classification is done through the last fully connected layer along with a softmax activation function that assigns each sample of the EEG dataset to the best possible speech class. Optimization of the model parameters is done with the help of an appropriate loss function and backpropagation algorithm. In order to assess the performance of the proposed framework, some commonly used performance metrics such as Accuracy, Precision, Recall, and F1-Score have been adopted. Accuracy determines the ratio of correctly classified samples, whereas Precision and Recall determine the correctness of identifying the positive class samples and retrieval of the correct samples, respectively. F1-Score provides a balanced performance measure in terms of Precision and Recall.

4. Results and Discussion

The performance of the suggested CWT–ResNet-50-D method was studied for four types of imagined speech, such as vowels, short words, long words, and combination of short-long words. For performance assessment, the following metrics were used: Accuracy, Precision, Recall, and F1-Score. The use of Continuous Wavelet Transform (CWT) along with ResNet-50-D provided an opportunity to extract discriminating temporal, spectral, and spatial features of EEG signals.

4.1 Classification Performance Analysis

Table 2 illustrates the classification accuracy of the proposed algorithm for all types of imagined speech. The following analysis proves that the proposed model shows good consistency in terms of all assessment criteria. Long words showed the highest classification accuracy of 77.5%, whereas short-long words showed the second highest classification accuracy of 76.4%, short words exhibited 74% classification accuracy, and vowels had 72.4% classification accuracy. These tendencies can be seen in terms of Precision, Recall, and F1-Score as well.

Table 2. Performance evaluation of the proposed CWT–ResNet-50-D framework

TFR Method	Model	Class	Accuracy	F1 Score	Recall	Precision
CWT	ResNet-50-D	Vowels	72.4	72.5	72.4	72.1
		Short Words	74	74.1	74.2	73.9
		Long Words	77.5	77.4	77.4	77.5
		Short-Long Words	76.4	76	77	75

The obtained results show that the long-word classification is the best in terms of all used criteria, which implies that speech imagination patterns consisting of longer words carry more useful neural information that can be efficiently analyzed using time-frequency decomposition techniques and feature learning. Moreover, the high level of correspondence between Accuracy, Precision, Recall and F1-Score proves the consistency and reliability of the proposed classification approach.

The higher efficiency of long-word classification is explained by greater cognitive processes required to perform longer speech imagination task. Longer imagination tasks create

more distinctive EEG patterns, thus, making it possible for the model to learn more discriminative features. On the contrary, vowel classification is less efficient because of low level of neural activity complexity and similarity among different vowels. Yet, the proposed approach exhibits a high level of classification accuracy for all speech categories.

4.2 Performance Metrics Comparison

Classification results were measured in terms of Accuracy, Precision, Recall, and F1-Score for all types of imagined speech classifications. These four metrics offer an overview of the classification capability of the model with regard to correctly classifying imagined speech signals, at the same time maintaining a balance between the two error types. Accuracy reflects the overall accuracy of the classification, Precision determines how reliable the positive predictions are, Recall tests the capability of the model to classify all relevant classes, and F1-Score gives a balanced score by combining Precision and Recall.

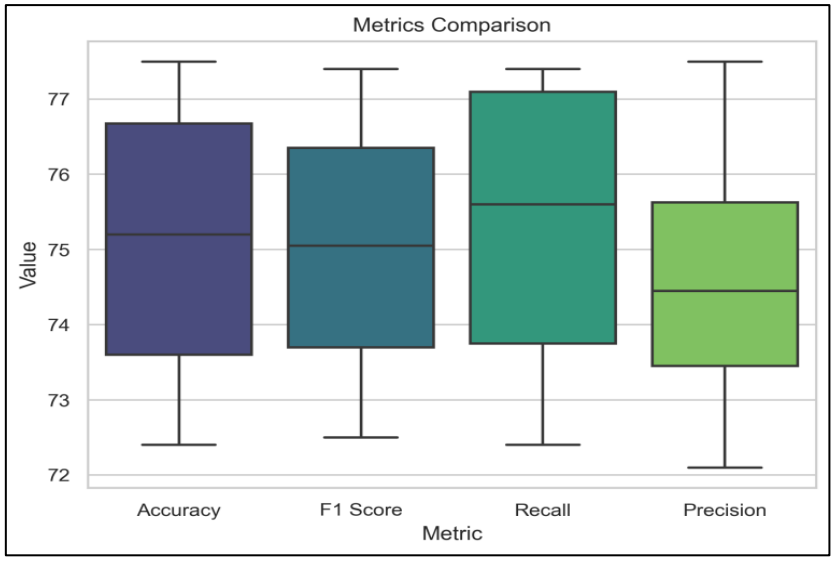

Figure 3. Distribution of performance metrics across imagined speech categories

Figure 3 shows the relative distribution of Accuracy, Precision, Recall, and F1 Score that have been achieved through the proposed methodology. As can be seen from the figure above, Recall has the highest median value with minimum variation. It signifies that the model has an excellent ability to classify imagined speech categories correctly. The distribution is almost similar for Accuracy and F1 Score. However, there is a little bit of variation for Precision.

The balance distribution of metrics shows that the model does not have any major biases in favor of one class. Such behavior becomes especially critical for the imagined speech classification task when class imbalance and inter-subject variability significantly influence classification results. The obtained results prove that the developed CWT-ResNet-50-D framework is capable of capturing informative features of EEG.

4.3 Comparative Analysis with Existing Methods

Further validation of the effectiveness of the proposed CWT-ResNet-50-D framework was carried out through a comparative analysis with state-of-the-art methods of imagined speech classification as documented in literature. Table 3 presents the performance comparison of the proposed CWT-ResNet-50-D framework with several other classification methods on the ASU Imagined Speech EEG Dataset with respect to various types of imagined speech tasks. The selected methods include different kinds of classification approaches such as traditional machine learning, tangent space methods, transfer learning models, and deep learning frameworks. Such a comparative study helps in the assessment of the proposed framework against benchmark models.

Table 3. Comparison of the proposed method with existing approaches

Author	Method	Vowels	Short Words	Long Words	Short-Long Words
Nguyen et al. (2018) [1]	Tangent + RVM	49	50.6	66.31	80.1
	Tangent + ELM	45.19	46.6	-	76.26
Singh et. al (2020) [11]	TS +ANN + Bagging	58.66	60.40	71	71.54
Proposed	CWT + ResNet-50-D	72.4	74	77.5	76.4

In Table 3, the proposed CWT-ResNet-50-D framework is evaluated against existing imagined speech classification systems. In terms of performance, the proposed framework outperforms the other approaches for classifying vowels, short words, and long words. In terms of the short-long word classification task, however, the proposed framework shows comparable performance with other state-of-the-art frameworks.

When compared with Tangent Space approach-based classifiers and traditional machine learning classifiers, it can be seen that there are significant gains in terms of accuracy

when it comes to classification. Specifically, for vowel classification, the proposed classifier provides an accuracy of 72.4%, which is way higher than those achieved by previous algorithms, with less than 60% accuracy rate. Short-word classification, on the other hand, yields an accuracy of 74.0%, which again is better than traditional machine learning methods.

This can be attributed to the following two key reasons. Firstly, Continuous Wavelet Transform is capable of providing accurate time-frequency local information from non-stationary Electroencephalography signals and forming informative scalograms. Secondly, ResNet-50-D can learn hierarchical spatial features using residual learning, allowing the system to differentiate between classes of imagined speech. In doing so, this model ensures the preservation of temporal and spectral properties at the same time.

4.4 Discussion

The empirical results confirm the efficacy of the combination of CWT with ResNet-50-D for imagined speech recognition. The suggested scheme is able to solve critical issues that occur with speech decoding using EEG, which are signal non-stationarity and difficulty in extracting features. Transforming the signals into time-frequency scalograms, the model can take advantage of more information than it could in the time domain representation. In addition, the residual architecture helps in learning features more efficiently. The developed CWT-ResNet-50-D architecture shows reliable and comparable performance, thus being appropriate for the application to future EEG-based imagined speech recognition and brain-computer interface systems. The achieved performance proves the effectiveness of using sophisticated signal processing methods together with deep learning models for the improvement of imagined speech decoding accuracy.

5. Conclusion and Future Work

In this research, an effective imagined speech recognition scheme was developed based on the combination of CWT and ResNet-50-D in order to decode EEG signals that are generated due to internal speech activities. Time-frequency scalogram representation was extracted for the preprocessed EEG signals by means of Continuous Wavelet Transform which allows extracting important information contained within the neural signals but is hardly extractable via the traditional time domain approach. Classification of the extracted time-frequency representations was then performed via the ResNet-50-D architecture which exhibits

strong potential in learning discriminative features from the EEG data. The experimental results, obtained on the ASU Imagined Speech EEG Dataset, confirmed the effectiveness of the proposed approach and showed promising classification accuracy rates, such as 72.4% for vowels, 74.0% for short words, 77.5% for long words and 76.4% for short-long words. This proves that the combination of time-frequency analysis and deep residual learning contributes to the improvement of the performance in imagined speech classification tasks and can be considered as an effective solution for EEG-based brain computer interfaces. In the future, research efforts would be directed towards the development of improved models for robustness via subject-independent learning, feature improvement using attention mechanisms, and transformers-based architectures. The use of the proposed framework on larger and varied sets of EEG data, along with multi-modal physiological information and real-time implementations, could aid in further development of imagined speech decoding systems.

References

[1] Nguyen, Chuong H., George K. Karavas, and Panagiotis Artemiadis. "Inferring Imagined Speech Using EEG Signals: A New Approach Using Riemannian Manifold Features." Journal of neural engineering 2018, vol. 15, no. 1: 016002.

[2] Kamble, Ashwin, Pradnya Ghare, and Vinay Kumar. "Machine-Learning-Enabled Adaptive Signal Decomposition for a Brain-Computer Interface Using EEG." Biomedical Signal Processing and Control 2022, vol 74: 103526.

[3] Mohan, Anand, and R. S. Anand. "Classification of Imagined Speech Signals Using Functional Connectivity Graphs and Machine Learning Models." Brain Topography 2025, vol 38, no. 2: 25.

[4] Mohan, Anand, and R. S. Anand. "Wavelet Augmented Phase Coherence Features for EEG-Based Imagined Speech Classification." IEEE Sensors Letters 2025, vol. 9, no. 8: 1-4

[5] Saha, Pramit, and Sidney Fels. "Hierarchical Deep Feature Learning for Decoding Imagined Speech from EEG." In Proceedings of the AAAI Conference on Artificial Intelligence 2019, vol. 33, no. 01, 10019-10020.

[6] Panachakel, Jerrin Thomas, A. G. Ramakrishnan, and T. V. Ananthapadmanabha. "A Novel Deep Learning Architecture for Decoding Imagined Speech from EEG." arXiv preprint 2020, arXiv:2003.09374.

[7] Panachakel, Jerrin Thomas, Ramakrishnan Angarai Ganesan, and T. V. Ananthapadmanabha. "Common Spatial Pattern-Based Data Augmentation Technique for Decoding Imagined Speech." In 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT) IEEE 2021, 1-5.

[8] Cohen, Michael X. "A Better Way to Define and Describe Morlet Wavelets for Time-Frequency Analysis." NeuroImage 2019, vol 199: 81-86.

[9] Phoommanee, Nonpawith, Peter J. Andrews, and Terence S. Leung. "Deep Learning-Based Phase Recognition in Anterior Nasal Endoscopy for Handheld Devices." In 2025 Research, Invention, and Innovation Congress: Innovative Electricals and Electronics (RI2C), IEEE, 2025, 50-53.

[10] He, Tong, Zhi Zhang, Hang Zhang, Zhongyue Zhang, Junyuan Xie, and Mu Li. "Bag of Tricks for Image Classification with Convolutional Neural Networks." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2019, 558-567.

[11] Singh, Abhiram, and Ashwin Gumaste. "Interpreting Imagined Speech Waves with Machine Learning Techniques." arXiv preprint 2020, arXiv:2010.03360.

[12] Arizona State University (ASU). "Imagined Speech EEG Dataset." Arizona State University, Tempe, AZ, USA.

[13] Available: IEEE DataPort: https://ieee-dataport.org/open-access/imagined-speech-dataset