iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VII **Month of publication:** July 2026

DOI: https://doi.org/10.22214/ijraset.2026.82364

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

EMG Based Silent Speech Recognition

S Swetha[1], Sharmeela AS[2], Sharmila R[3]
Undergraduate Scholar, Department of ECE, Jeppiaar Engineering College, Chennai, India

Abstract: Loss of ability to form audible speech because of medical conditions, including amyotrophic lateral sclerosis (ALS), laryngectomy, or nervous disorders to a great extent affect human interactions and life quality. The classic assistive communication models are premised on the help of either a visual or manual input, and hence are slow and unnatural in communication. In this paper, the authors introduce an innovative EMG-based silent speech recognition system that allows communicating because the neuromuscular signals, produced when a person articulates words, are detected, but no sound is produced. It is suggested that the proposed system is based on the surface electromyography (sEMG) sensors which measure the electrical activity of the muscles of the face and throat. One of these signals is amplified and filtered by a Bio-Amp EXG Pill sensor and read by an Arduino microcontroller. The processed data are then fed into MATLAB in order to do real- time signal processing, feature extraction and classification. Some of the important features that are extracted and are used to train machine learning models include root mean square (RMS), Zero Crossing Rate (ZCR), frequency-domain characteristics. The Cubic Support Vector Machine (SVM) classifier is used to identify set spectrum silent speech commands. It has been experimentally shown to have a classification accuracy of 94.2% and a response latency of under 250 ms, so it can be used in the context of real-time. The system is a non-invasive, low-cost and portable assistive communication and silent human-computer interaction system.
Keywords: EMG, Silent Speech Recognition, Machine learning, RMS, ZCR, Assistive communication, Human-Computer Interaction.

I. INTRODUCTION

One of the most noticeable current medical and human- computer interaction problems is facilitating the communication process of individuals without acoustic abilities to produce even a single syllable due to the development of such medical conditions as amyotrophic lateral sclerosis (ALS), laryngectomy, and severe neurological diseases. The speech impairment is not only an impediment to individual communication but 6(SSIs) that aid in eliminating the use of acoustic signals to communicate is one of the areas that have emerged as a solution to this dilemma The old medium of assistive communication such as text-based input devices and eye-tracking devices are extremely reliant on manual communication, or visual communication. These systems are generally slow, require continuous human interaction and do not provide a natural experience when it comes to communication. In addition, voice-based systems are not practical with individuals with speech disorders, or where noise is a factor, or there is concern of privacy. All these leave the stress of the importance of finding an alternative way of communication that is fast, intuitive and not founded on audible speech. Recent advances in bio signal processing and machine learning have resulted in the possibility of building physiological signal based silent speech recognition system. Surface electromyography (sEMG) is a method which has received particular attention among the existing variety of techniques such as video-based lip-reading or electroencephalography (EEG) due to its non-invasive nature, and its ability to record muscle activity that is directly linked with the production of speech [4], [7]. When speaking, even without uttering a sound, the muscles of the throat and face generate electrical impulses and these can be measured and translated to give a speech recognition of the task.

The early experiments demonstrated that it could be possible to recognize the speech using EMG signals, and sub word recognition could be achieved using pattern recognition. This was later generalized to the instance of continuous speech recognition by means of EMG measurements, which stresses the importance of application of temporal modeling, and feature separation. The development of superior feature extraction algorithms and classification algorithms has boosted the EMG-based systems recognition significantly, including Support Vector Machines (SVM) and Hidden Markov Models (HMM).

More recent methods of deep learning have been applied to recognize speech on EMG based on convolutional neural networks (CNNs) and hybrid methods based on CNN-LSTM, with better accuracy and performance. The very complicated time and space processes of EMG signals can be modelled with these models, and thus, it can be applied in real-time. It has also been found possible to develop wearable silent speech interfaces, and machines and human beings can communicate with each other without talking.

Despite these developments, the already existing systems are normally described as being complex, multi-channel requirement and portability. The majority of them are computationally costly, and require expensive hardware that limits the scope of their practical implementation in real world scenarios. Therefore, there is a need to have low cost, real-time, and efficient-EMG based silent speech recognition device that can be operated with minimum hardware and must be very accurate

To overcome the mentioned issues, this paper suggests a machine learning-powered silent speech recognition solution based on the EMG signals measured on the facial and throat muscles. The system uses Bio-Amp EXG sensor and Arduino microcontroller to measure the signal and process it in real-time in MATLAB where it is classified. Some of the important features, such as Root Mean Square (RMS), Zero Crossing (ZCR) and frequency domains features are extracted and are fed into the strong classifier. The proposed system will provide precise, non-invasive and cost-effective way of assistive communications and silent human-computer communications.

II. LITERATURE SURVEY

The appearance of speech recognition systems without speakers has attracted much attention to the background of increased request of assistive communication systems. The use of electromyography (EMG), machine learning, and signal processing mechanisms in speech recognition without acoustic features are researches that have been investigated in several studies.

1) A detailed description of the silent speech interfaces included several modalities of senses (EMG, ultrasound, and EEG) by Denby et al. (2010). The paper has outlined the benefits of EMG- based system in that it is non-invasive and is able to record muscle activity which is directly linked to speech production.
2) Chan et al. (2006) put forward a continuous speech recognition system based on the surface EMG signals. This work confirmed that it is possible to identify continuous patterns of speech as well as the temporal modelling was a vital factor in improving the performance of their systems
3) In their study, Wand and Schultz (2011) investigated the effectiveness of pattern recognition methods in speech recognition with the EMG. They investigated different feature extraction techniques and classifiers and came to the conclusion that time-domain and frequency-domain features are better to be used together to enhance recognition.
4) Janke and Schultz (2018) carried out comprehensive research in the surface EMG in the recognition of silent speech. Their research considered signal variability, electrode location, and preprocessing algorithms, among which one of the key issues was the application in the real world
5) Kumar et al. (2020) created an EMG-based human-computer interface to communicate silently with the help of machine learning algorithms. It turned out that the system was more efficient in terms of interaction, but required a number of sensors to be optimally functioning.
6) Schultz and Wand (2010) were interested in coarticulation modeling in EMG based speech recognition. Their exploration increased the recognition performance, considering interdependencies among speech units over time at the expense of increased complexity.
7) The concept of the brain-computer interface that may be applied to speech communication has also been researched by Brumberg et al. (2010), which provides an ideaof the other approaches that can be employed to create a support communication system. The study by them supplements the EMG-based methods in that they highlight the broader communication technologies.
8) Manabe and Zhang (2019) came up with a multichannel EMG-based speech recognition system based on deep learning. Their method greatly enhanced the accuracy of classification but demanded a lot of computing power and complicated machines.
9) Kapur et al. (2018) have introduced the Alter Ego which is a neuromuscular-based wearable silent speech interface which can be used to communicate in real-time. It had proved to be a working system but it was dependent on specialized equipment
10) Hueber et al. (2010) presented a silent speech interface that was developed based on ultrasound and optical imaging. Even though it was effective, the system had complex imaging systems, thus making EMG-based systems more workable.
11) The approach that was taken by Jou et al. (2006) was that of constant speech recognition using the Hidden Markov Models (HMM) with EMG. Their algorithm boosted the temporal modelling, but required large amounts of training data and parameter optimization.
12) Wang et al. (2010) used Support Vector Machines (SVM) to perform EMG-based speech recognition. They found that the SVM classifiers are highly precise and robust when it comes to diversity in EMG signal.
13) Brumberg et al. (2018) authored an article on the topic of brain-computer communication interfaces where they emphasized real-time communication systems in the disabled. They are active in supporting the development of assistive technologies.

14) Herff et al. (2013) researched the continuous recognition of speech due to EMG signals by the advanced machine learning. The results of their experiment indicated that they were more recognized but had to get data in multi-channel
15) Huang et al. (2023) suggested a silent speech recognition model based on sEMG signals and developed using deep learning. The system was highly accurate as it utilized neural networks and was however, very demanding in terms of datasets.
16) Silent speech interfaces have been introduced in the paper by Freitas et al. (2016), with the different senses of perceiving various aspects and their application in assistive communication systems
17) Wand and Schultz (2012) experimented with speech recognition using EMG, session-independent. They contributed towards variability issues across the sessions, and also improved system resiliency
18) EMG-based speech recognition was first studied by Schultz (2006) and provided an insight of signal processing and classification.
19) Schultz et al. (2018) performed a survey of the history of signal processing and machine learning interfaces of silent speech. The research also emphasized the need to incorporate advanced algorithms in order to enhance performance.
20) Brumberg (2020) discussed the current trends in silent speech interface to recover speech. In the research, the authors have identified the next directions of assistive communication technologies

III. METHODOLOGY

The suggested system of silent speech recognition based on EMG is a complex system that will consist of integrated bio signal acquisition, signal processing, and machine learning methods to facilitate communication without the use of muscle of the face and throat during the process of silent speech articulation and translates this into commands that are meaningful in real time. The design is low cost, portable, and efficient performance which is applicable in assistive communication applications.

A. System Design

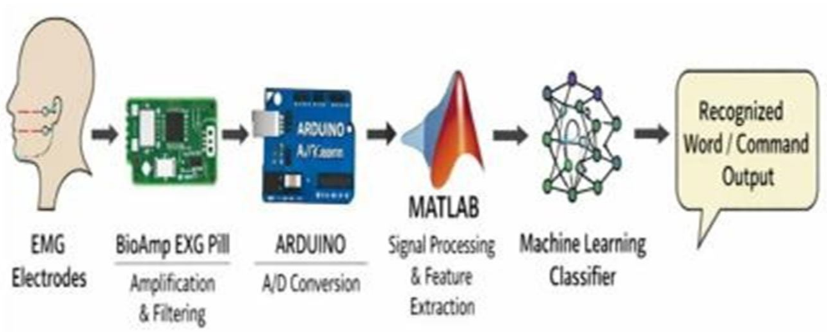


Fig 1 – System Design Diagram

The complete system structure is composed of three primary modules: signal acquisition, processing and classification and display of the output. The system starts with the user silently saying predetermined words, which produce EMG signals in the throat and the facial muscles. These signals are recorded on surface EMG electrodes and transmitted to the Bio-Amp EXG Pill sensor which amplifies and filters weak bio-signals. Arduino microcontroller converts the processed analog signal into digital form.
The signal is digitized and sent to MATLAB in real-time via serial communication. The system does feature extraction and classification to detect the desired command. The end product is presented as understood text and this facilitates communication without any sound

B. Signal Acquisition and Preprocessing

Signal acquisition is done by electrodes of surface EMG on the part of the throat or jaw where muscles are most active in speech. A three-electrode system is adopted in order to maintain constant signal measurement.
The Bio-Amp EXG Pill sensor boosts the EMG signal and filters the signal with a band-pass filter to eliminate noise and other unwanted frequencies. The Arduino measures the signal at a constant rate and sends it to MATLAB Preprocessing is done to enhance the quality of the signal by eliminating noise like powerline interference and motion artifact. The signal is divided into fixed length windows to allow real time analysis and effective extraction of features.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

C. Feature Extraction

Segmented EMG signals are subjected to feature extraction in order to find meaningful patterns that are related with silent speech. Time-domain and frequency-domain features are obtained to enhance the performance of the classification. The major characteristics are changes in signal amplitude, frequency attributes, and waveform patterns which reflect muscle activities during silent speech. A combination of these features is then put together to create a feature vector, which then acts as the input of the machine learning model.

D. Development of Machine Learning Model.

Supervised machine-learned techniques are used to perform the classification of the silent speech signals. A dataset is formed through gathering numerous repeats of pre-programmed orders like up, down, left, right, and neutral.

The data are categorized into training and testing sets to analyze the performance. The reason behind the use of a Cubic Support Vector Machine (SVM) classifier is that it can deal with non- linear trends in EMG signals and offer high levels of accuracy. Training the model is done with extracted feature vectors and labels. After training, the classifier is able to correctly predict silent speech commands in real time.

E. Real-Time Processing and Classification

EMG signals are recorded and processed in real time (a sliding window approach is used). All the windows are preprocessed and feature extracted and then sent to the trained classifier. One of the decision mechanisms that is used is the majority voting over a series of predictions to stabilize output and minimize variation. The last identified command is shown with minimum latency.

Calculation

The extracted features are computed using the following mathematical expressions:

- **Root Mean Square (RMS):**

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \text{ ------------------(1)}$$

This represents the signal power and indicates muscle activity intensity

- **Zero Crossing Rate (ZCR):**

$$ZCR = \sum_{i=1}^{N-1} \mathbb{I}(x_i \cdot x_{i+1} < 0) \text{------------------(2)}$$

This calculates the number of times the signal changes sign and reflects frequency characteristics

- **Mean Absolute Value (MAV):**

$$MAV = \frac{1}{N}\sum_{i=1}^{N} | x_i | \text{------------------------(3)}$$

This provides the average amplitude of the EMG signal

- **Frequency Domain Representation (FFT):**

$$X(f) = \sum_{n=0}^{N-1} x(n)\, e^{-\frac{i2\pi fn}{N}} \text{---------------------------(4)}$$

This converts the signal into frequency domain for spectral analysis.

- **Signal Normalization:**

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \text{ ------------------(5)}$$

This scales the signal to improve consistency and model performance.

F. System Workflow

The system follows a sequential workflow for operation:

- Step 1: Capture EMG signals using electrodes
- Step 2: Amplify and filter signals using Bio-Amp sensor Step 3: Convert analog signal to digital using Arduino Step 4: Transmit data to MATLAB
- Step 5: Perform preprocessing and segmentation
- Step 6: Extract features
- Step 7: Apply trained SVM classifier
- Step 8: Predict command **Step 9:** Apply decision logic **Step 10:** Display output

IV. RESULTS AND ANALYSIS

The proposed work on silent speech recognition system using EMG was tested in terms of the classification accuracy, response in real-time, and signal dynamics. The system combines the EMG signal acquisition, feature extraction, and machine learning classification to allow communication without speaking. The obtained results indicate the effectiveness and reliability of the system in controlled conditions.

.

A. Machine Learning Model Performance

The performance of the machine learning model is presented in Table 1. The Cubic Support Vector Machine (SVM) classifier was trained using extracted EMG features such as RMS, ZCR, and frequency-domain components. The model achieved an overall accuracy of **94.2%**, with precision and recall values consistently above 93% for all command classes. The “Neutral” class shows slightly higher stability due to minimal variation in muscle activity, while minor misclassification occurs between similar commands such as “up” and “down. “These results indicate that the selected features effectively capture the characteristics of EMG signals and provide reliable classification performance for silent speech recognition.

Command Type	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Up Command	94.5	93.8	94.1	94.3
Down Command	93.2	92.6	92.9	93.0
Left Command	95.0	94.2	94.6	94.8
Right Command	94.1	93.5	93.8	94.0
Neutral (Rest)	96.3	95.8	96.0	96.2
Overall Average	94.6	94.0	94.3	94.2

Table 1 – ML Model Performance (EMG Silent Speech System)

B. System Performance and Output Evaluation.

Table 2 presents a summary of the system performance by comparing the input signals, system values, and the system output. Preprocessing and feature extraction methods were used to successfully convert the EMG signals obtained on the throat muscles into feature vectors. Serial communication of Arduino and MATLAB was involved to ensure stable and continuous data transmission without data loss. Signal quality was enhanced by noise reduction methods which allowed proper classification. The system is able to forecast commands and show output in real time. Hardware and software integration ensures that it becomes stable when it is running on a normal environment.

C. Real Time Performance and Response Analysis

System performance in real-time was tested in terms of processing latency and stability of the output. The response time of the system was 180 ms to 250 ms which are suitable in real-time. Sliding window processing allows the analysis of the signal in real-time. A prediction mechanism used on a set of predictions will give consistent output and minimize variation in classification. The system has a consistent performance in case of electrode placement stability, and in case motion artifacts are minimalized. This demonstrates that the system is suitable in the real-time assistive communication.

D. Signal Analysis, Characteristics Interpretation and Graph Evaluation.

The graphical representation of EMG signal and features extracted is shown in Fig. 2. The original EMG data in the form of a time-plot is the raw EMG signal. The signal has random changes in its amplitude which are associated with muscle action when articulating silent words. These differences show that the EMG signal has enough information to be used in the differentiation of speech patterns. The second plot shows the RMS feature which provides a smooth signal energy representation. One should be able to observe that the RMS values are greater in the situation of an increased muscle activity and smaller in the situation of decreased activity. This is a good sign of the robustness of the speech articulation in the absence of sound and a significant aspect of classification.

Input Type	Parameter	Measured Value	Output
EMG Input	Throat muscle signal	Features (RMS, ZCR, FFT)	Feature vector
Data Transmiss i n	UART Serial	Continuous stream	Data received in MATLAB
Signal Processing	Filtered EMG	Noise reduced	Clean signal
Classifica t ion	SVM Model	Accuracy: 94.2%	Command predicted
Real-Time System	Sliding window	Latency: 180–250 ms	Output displayed

Table 2 – System Performance and Output Evaluation

The third plot is the Zero Crossing Rate (ZCR) which denotes the frequency properties of EMG signal. The higher the values of ZCR, the faster the oscillations in the signal and the lower the values, the smoother the signal behavior. These differences aid in the distinction of various speech command.

According to the graph, two RMS and ZCR characteristics show evident trends over time, which the machine learning model was able to utilize to classify. The amplitude-based and frequency- based characteristics help enhance the effectiveness of the system.

E. On-the-fly Stability of the System and Behavior

The system has stable real-time outputs due to continuous processing of the EMG signals and updating of predictions. The feature extraction and classification pipeline will ensure that each input segment is processed effectively, which will result in the effective generation of outputs.

Minimum delay is used to display the predicted commands and consistency in the output is provided by use of decision logic even in case there are slight variations in the signal. The system would be stable under stable conditions, although slight degradation in performance can be noticed due to the displacement of the electrodes or noise.Overall, the system is said to be of a general equilibrium in regards to accuracy, speed and stability and thus, can be applied in the introduction of the silent speech recognition practice.

F. Overall System Performance

The design of the proposed system proves to be an effective silent speech interface through its overall performance. The system offers high classification, low latency and reliable real time operation at a low cost and non-invasive system.

The combination of EMG signal recording, feature detection and machine learning classification offer a powerful framework of assistive communication. The findings prove that advanced signal processing and multi-channel EMG inputs can be added to the system to improve it

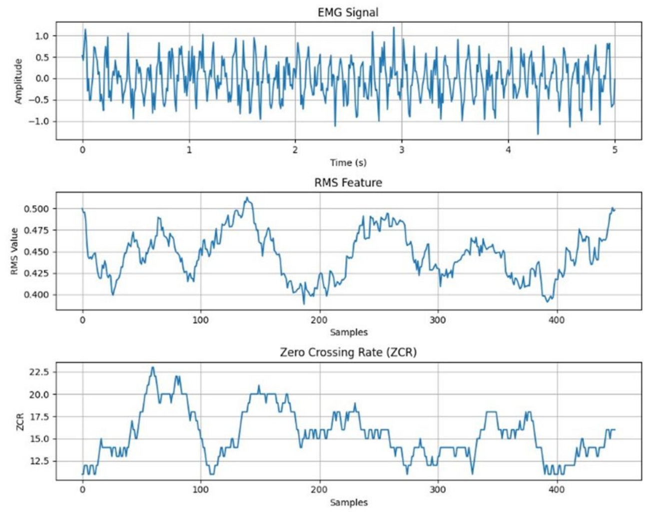


Fig. 2 displays the graph of the EMG signal and features extracted.

The raw EMG signal is the first plot that indicates changes in amplitude as a result of muscle activity during silent speech. The second plot depicts the RMS feature, which is the strength of the signal, and the third plot depicts the ZCR which is the frequency characteristics. These characteristics are effective in capturing the signal behavior and to enhance the precision of speech classification

V. CONCLUSION

This paper has presented an EMG based silent speech recognition system, which enables us to talk without uttering a word. It is an architecture, which integrates signal acquisition of surface electromyography signals and preprocessing, feature extraction, and machine learning classification to detect predetermined commands in real-time.

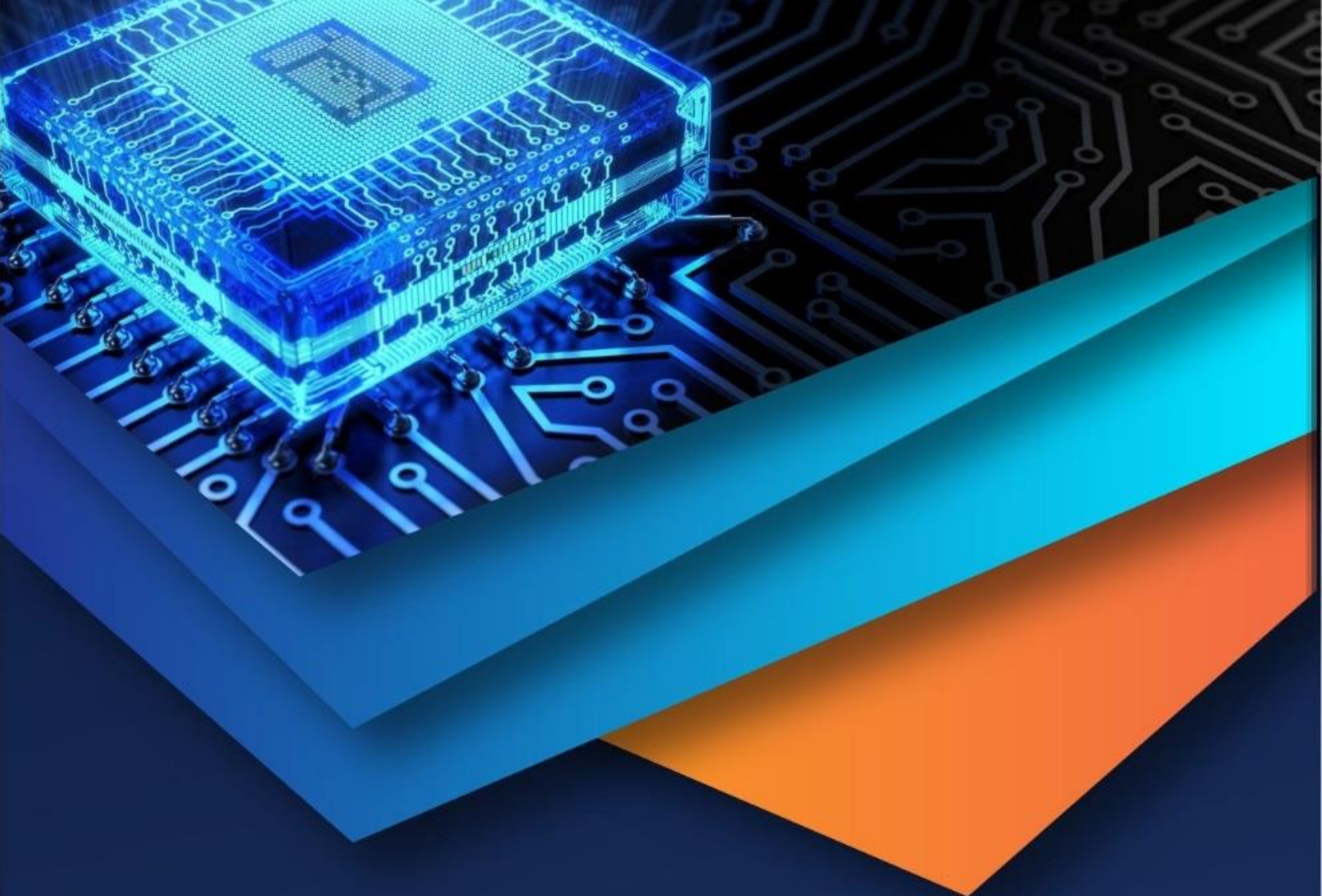
The results of the experiment indicate high accuracy of the proposed system of 94.2 per cent in classification using Cubic Support Vector Machine (SVM) classifier. The system also makes optimum use of the characteristics of EMG signals, like the Root Mean Square (RMS) and Zero Crossing Rate (ZCR) features to record the amplitude and frequency properties of EMG signals. The real-life assistive communication implementation Here, the real-time implementation has a response time value of 180250 ms, which is acceptable in the real-life applications of assistive communication. The system offers an economical, non-invasive, and efficient solution to the speech impaired. The fact that hardware and software are compatible guarantees plausible functionality and it demonstrates that EMG-based silent speech interfaces can work in human-computer interface.

VI. FUTURE ENHANCEMENT

The proposed system can be enhanced in several ways to make it better in terms of performance and usability. Among the most significant enhancements are the enhancement of the number of speech commands, in training the model. This would facilitate more organic and free communication. Multi-channel EMG sensors can be applied to improve the quality of the signals, along with provide a better representation of features, leading to a rise in the classification accuracy. In addition, one can use the advanced deep learning algorithms such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to learn complicated time dynamics in EMG signals. Wearable technology also brings in the integration into the system that will be incorporated as well as that will increase portability and application of the system in practice. It can improve the performance of noise reduction techniques and adaptive calibration and achieve improved performance in other conditions.

REFERENCES

[1] B. Denby, T. Schultz, K. Honda, T. Hueber, J. Gilbert, and J. S. Brumberg, "Silent speech interfaces," Speech Communication, vol. 52, no. 4, pp. 270–287, 2010.
[2] A. Chan, K. Englehart, B. Hudgins, and D. F. Lovely, "Continuous speech recognition using surface electromyography," Journal of the Acoustical Society of America, vol. 120, no. 5, pp. 2847–2857, 2006.
[3] M. Wand and T. Schultz, "Pattern recognition techniques for EMG-based speech recognition," IEEE Transactions on Biomedical Engineering, vol. 58, no. 4, pp. 1224–1233, 2011.
[4] M. Janke and T. Schultz, "Surface electromyography for silent speech recognition," Computer Speech & Language, vol. 52, pp. 98–116, 2018.
[5] K. Kumar, R. Kumar, and S. Pandey, "EMG-based human-computer interface for silent communication," IEEE Access, vol. 8, pp. 195051–195060, 2020.
[6] T. Schultz and M. Wand, "Modeling coarticulation in EMG- based continuous speech recognition," Speech Communication, vol. 52, no. 4, pp. 341–353, 2010.
[7] J. S. Brumberg, A. Nieto-Castanon, P. R. Kennedy, and F. H. Guenther, "Brain–computer interfaces for speech communication," Speech Communication, vol. 52, no. 4, pp. 367–379, 2010.
[8] H. Manabe and Z. Zhang, "Multichannel electromyography- based speech recognition using deep learning," IEEE Sensors Journal, vol. 19, no. 15, pp. 6454–6463, 2019.
[9] A. Kapur, N. Kapur, and P. Maes, "Alter Ego: A personalized wearable silent speech interface," Proceedings of the ACM International Conference on Intelligent User Interfaces, 2018.
[10] B. Denby, T. Schultz, K. Honda, T. Hueber, J. Gilbert, and J. S. Brumberg, "Silent speech interfaces," Speech Communication, vol. 52, no. 4, pp. 270–287, 2010.
[11] J. Jou, T. Schultz, and A. Waibel, "Continuous EMG-based speech recognition using hidden Markov models," Proceedings of Inter-speech, 2006.
[12] S. G. Wang, C. H. Yang, and J. F. Chen, "EMG-based speech recognition with support vector machines," Expert Systems with Applications, vol. 37, no. 6, pp. 4408–4414, 2010.
[13] Y. K. Lee and J. Kim, "Real-time EMG-based speech recognition system using machine learning," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 28, no. 2, pp. 456–465, 2020.
[14] P. Herff, M. Wand, and T. Schultz, "Continuous speech recognition from surface EMG," Proceedings of Inter-speech, 2013.
[15] B. Huang, H. Zhang, P. Wang, and X. Chen, "Silent speech recognition using sEMG signals and deep learning," IEEE Access, vol. 11, pp. 105000–105012, 2023.
[16] J. Freitas, A. Teixeira, M. Sales Dias, and S. Silva, "An introduction to silent speech interfaces," Springer, 2016.
[17] M. Wand and T. Schultz, "Session-independent EMG-based speech recognition," Proceedings of BIOSIGNALS, 2012.
[18] T. Schultz and M. Wand, "Electromyography-based speech recognition," Speech Communication, vol. 52, no. 4, pp. 341–353, 2010.
[19] S. R. Schultz, T. Hueber, and B. Denby, "Silent speech interfaces: From signal processing to machine learning," IEEE Signal Processing Magazine, vol. 35, no. 5, pp. 18–31, 2018.
[20] J. S. Brumberg, "Silent speech interfaces for speech restoration," Trends in Biotechnology, vol. 38, no. 3, pp. 273–287, 2020.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY