Leveraging Deep Neural Networks and Surface Electromyography for Real-Time Airwriting Gesture Recognition

By

Atiqa Saeed

(Registration No: 00000401207)

Department of Biomedical Engineering

School of Mechanical & Manufacturing Engineering

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2024)

Leveraging Deep Neural Networks and Surface Electromyography for Real-Time Airwriting Gesture Recognition

By

Atiqa Saeed

(Registration No: 00000401207)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Master of Science in
Biomedical Engineering

Supervisor: Dr. Asim Waris

School of Mechanical and Manufacturing Engineering

National University of Sciences & Technology (NUST)

Islamabad, Pakistan

(2024)

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis written by **Regn No. 00000401207 Atiqa Saeed** of **School of Mechanical & Manufacturing Engineering (SMME)** has been vetted by undersigned, found complete in all respects as per NUST Statues/Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS/MPhil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis titled. **Leveraging Deep Neural Networks and Surface Electromyography for Real-Time Airwriting Gesture Recognition**

Signature: ____________

Name (Supervisor): Dr Muhammad Asim Waris

Date: 19/12/2024

Signature (HOD): ____________

Date: 20/12/2024

Signature (DEAN): ____________

Date: 20.12.2024

Form TH-4

National University of Sciences & Technology (NUST)

MASTER'S THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: Atiqa Saeed (00000401207)

Titled: Leveraging Deep Neural Networks and Surface Electromyography for Real-Time Airwriting Gesture Recognition be accepted in partial fulfillment of the requirements for the award of MS in Biomedical Engineering degree.

Examination Committee Members

1.	Name: Aneeqa Noor	Signature:
2.	Name: Ahmed Fuwad	Signature:
3.	Name: Muhammad Asim Waris	Signature:
Supervisor: Muhammad Asim Waris	Signature: Date: 10 - Dec - 2024	

	10 - Dec - 2024
Head of Department	Date

COUNTERSINGED

10 - Dec - 2024	
Date	Dean/Principal

CERTIFICATE OF APPROVAL

This is to certify that the research work presented in this thesis, entitled "Leveraging Deep neural networks and surface electromyography for real-time airwriting gesture recognition" was conducted by Mr./Ms Atiqa Saeed under the supervision of Dr Asim Waris. No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the Department of Biomedical Engineering and Sciences in partial fulfillment of the requirements for the degree of Master of Science in Field of Biomedical Engineering. Department of ………… National University of Sciences and Technology, Islamabad.

Student Name: Atiqa Saeed Signature:

Supervisor Name: Dr Muhammad Asim Waris Signature:

Name of Dean/HOD: Dr Muhammad Asim Waris Signature:

AUTHOR'S DECLARATION

I Atiqa Saeed hereby state that my MS thesis titled "Leveraging Deep neural networks and surface electromyography for real-time airwriting gesture recognition" is my own work and has not been submitted previously by me for taking any degree from National University of Sciences and Technology, Islamabad or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: Atiqa Saeed

Date: 10-12-2024

PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled "……Leveraging Deep neural networks and surface electromyography for real-time airwriting gesture recognition" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and National University of Sciences and Technology (NUST), Islamabad towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the University reserves the rights to withdraw/revoke my MS degree and that HEC and NUST, Islamabad has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Atiqa Saeed

Student Signature: ____________________________

Name: Atiqa Saeed

To my parents, whose endless love and sacrifices have been my guiding light,

To my husband, whose unwavering support and belief in me kept me motivated,

To my siblings, for their constant encouragement,

To my supervisor, for their invaluable guidance,

and to myself, for staying determined and resilient throughout this journey.

ACKNOWLEDGEMENTS

I express my deepest gratitude to Allah Subhanahu Wa Ta'ala, the Most Merciful and Most Compassionate, His endless blessings and guidance have enabled me to achieve this milestone. Without His divine grace, wisdom, and strength, I would not have been able to reach this milestone.

I am deeply indebted to my parents *Saeed Akhtar Nadeem* & *Abida Kousar*, whose unconditional love, prayers, and sacrifices have been the foundation of my success. To my husband, *Abdul Majid*, thank you for your unwavering support, and belief in me throughout this journey.

I am also grateful to my siblings *Amina Saeed, Aatifa Saeed, Sibga Saeed,* and *Fatima Saeed* for their continuous encouragement and for always being there for me. I sincerely thank you *Bhai Umair* and *Bhai Awais* for your constant encouragement,

My sincere thanks go to my supervisor, *Dr. Asim Waris*, whose expertise, mentorship, and constructive feedback were instrumental in shaping this thesis. I would like to thank the members of the GEC *Dr. Ahmad Fuwad, Dr. Aneeqa Noor* and *Dr. Daniyal Mahmood* for their insightful suggestions and contributions.

I am also deeply thankful to my friends *Asma, Ayeza, Nimra, Omna, Rabia,* and my confidant roommate, *Mina Khalid,* for their understanding, support, and moral encouragement throughout this process. Special thanks to my lab mates *Nida* and *Nazli* for their support, collaboration, and shared knowledge.

Finally, I extend my gratitude to everyone who has contributed, directly or indirectly, to the completion of this work.

TABLE OF CONTENTS

LIST OF TABLES

Page No.

LIST OF FIGURES

Page No.

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

CNN	Convolutional Neural Networks
DNN	Deep Neural Networks
RNN	Recurrent Neural Networks
HCI	Human Computer Interaction
VR	Virtual Reality
AR	Augmented Reality
KNN	K Nearest Neighbors
SVM	Support Vector Machine
sEMG	Surface Electromyography
LDA	Linear Discriminant Analysis
PCA	Principle Component Analysis
RFECV	Recursive Feature Elimination with Cross Validation
LSTM	Long Short-Term Memory
STFT	Short Time Fourier Transform

ABSTRACT

Airwriting enables users to write letters or characters in free space using hand or finger movements, with potential applications in human-computer interaction, virtual reality, augmented reality and development of assistive technologies. Despite advancements in gesture recognition technology, dynamic airwriting faces challenges with accuracy and often lacks real-time capabilities, limiting its application in non-verbal communication and rehabilitation devices. The primary objective of this research is to develop a novel real-time deep learning-based framework for airwriting recognition using surface electromyography (sEMG). This study presents a technique for real-time identification of uppercase English language alphabets written in free space by analyzing the electrical activity of forearm muscles involved in writing letters. The proposed framework involves sEMG data collection from 16 right handed healthy subjects with no neuromuscular or motor impairments, signal preprocessing, feature extraction, classification using Convolution neural network (CNN), Deep neural network (DNN) and Recurrent Neural Network (RNN).The best performing model was implemented in real-time and it was evaluated using performance metrics such as accuracy, precision, recall, F1 score, Confusion metrics and latency. Results show that 1 Dimensional (1D) CNN outperforms other models ($p<0.05$) and achieved an offline test accuracy of 89.81 ±0.87% and an average real-time test accuracy of 73.71 ±8.46% across subjects. The individual model of each subject performed even better, with an accuracy of 90.01 ±2.85% on offline testing of data and 75.45 ±1.53 % in real-time alphabet prediction. Thus, this work highlights the potential of deep learning models for real-time airwriting detection and provides foundations for sEMG-based airwriting applications in healthcare and telemedicine.

Keywords: ***:*** Electromyography (EMG), Airwriting, Deep learning, Convolution neural network (CNN), Human Computer Interaction

CHAPTER 1: INTRODUCTION

1.1 Brief overview of electromyography (EMG)

Electromyography (EMG) is a technique used to measure the electrical activity of skeletal muscles. It detects the electrical potentials generated by muscle cells, when these cells are electrically or neurologically stimulated [1]. It can be recorded using two main types: surface electromyography (sEMG) and intramuscular EMG. sEMG does not require penetration of the skin and involves placing electrodes directly on the skin surface to record activity of muscles and tendons from underneath the skin. Intramuscular EMG on the other hand utilizes thin needle electrodes that are inserted into the muscles and is mainly applied clinically and gives more localized information. EMG is suitable for application in clinical practices and research studies. In the clinical settings, it is used to detect diseases like carpal tunnel syndrome, muscular dystrophy and motor neuron diseases among others. It has been applied in a variety of research fields, including kinesiology, sports science, human-computer interaction[2], rehabilitation, and prosthesis control [3]. It offers insightful information on how muscles operate.

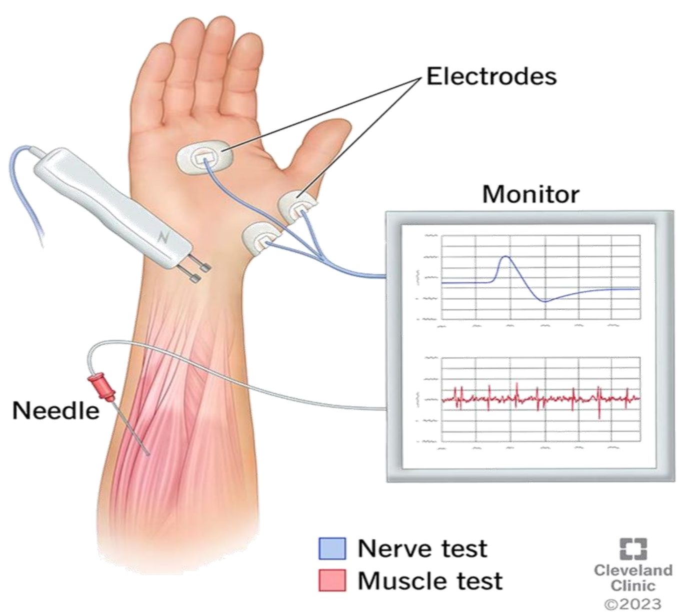


Figure 1.1: Electromyography (EMG) Technique for Monitoring Muscle and Nerve Activity

In general, EMG is a general and effective technique for analyzing the muscle contraction pattern, for diagnosing various neuromuscular disorders and for investigating new paradigms of operating devices based on muscle signals. Figure 1.1 illustrates an EMG test where electrodes are placed on the skin surface to record electrical activity from the muscles and nerves in the arm. The needle represents intramuscular EMG, used for more precise measurement of muscle activity, while surface electrodes capture signals from the skin. The monitor displays nerve (blue) and muscle (red) electrical signals, used to evaluate neuromuscular function.

1.2 Concept of airwriting

Airwriting is the process of writing words or characters in the free space using hand or finger movements [4]. This technology holds great promise for hands-free and touchless interaction since it allows users to input text without making physical contact with a surface. Airwriting recognition algorithms attempt to interpret and translate these motions into digital text for use in augmented reality (AR), virtual reality (VR), and assistive technologies for people with impairments [5]. Figure 1.2 shows the framework of airwriting recognition using sEMG.

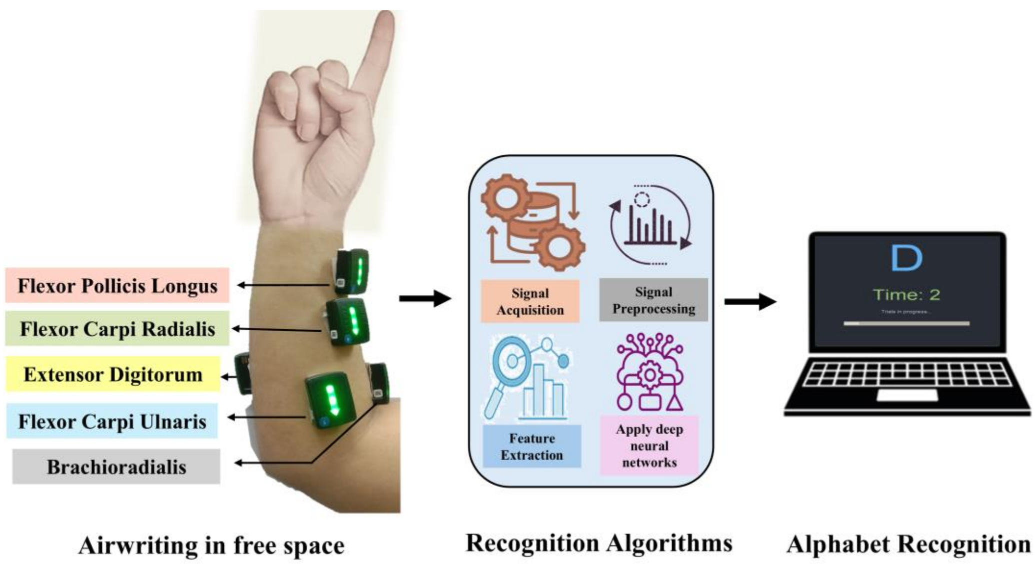


Figure 1.2: Framework of airwriting recognition using surface Electromyography

1.3 Limitations in current airwriting recognition systems

The first and main problem of airwriting recognition is the differentiation of complicated and changing hand gestures in real-time data. In contrast to conventional writing that is done on a surface, which limits hand movements and serves as a reference point, airwriting does not have a spatial and temporal reference for the analysis of hand gestures. The movements involved in airwriting gestures may differ from one person to another depending on the speed at which the person writes, the size of the hand, and the amount of energy the person uses while airwriting. These variations and the problem of maintaining accurate character classification still prove to be quite challenging to catch. One of the primary challenges in airwriting recognition is accurately analyzing the complex and fluctuating sEMG signals generated during the airwriting process. These signals are influenced by numerous factors, including muscle fatigue, electrode placement, and individual variations in anatomy of muscles. Character recognition in existing airwriting recognition systems is generally inaccurate due to substantial variability in sEMG signals. Furthermore, a lot of the existing techniques rely on manually created features and conventional machine learning techniques, which might not adequately capture the complex patterns in sEMG data. sEMG-based airwriting recognition systems

1.4 Problem Statement

Despite the advancement in gesture recognition, most of the work has shifted towards working with static gestures with little or no work done on dynamic gestures such as airwriting. Some of the few systems, which try to detect airwriting gestures, often do not possess the real-time characteristic, which is essential in most real-life applications including the HCI. This gap in real time recognition hinders the growth of assistive technologies and nonverbal communication mechanisms that are essential to people with speech impediments or those with prosthetics. The lack of accurate and real time systems presents an important challenge to the further development of air writing as a viable and efficient method of text input. Hence, deep learning in combination with surface electromyography (sEMG) signals has the potential to eliminate these drawbacks and allow accurate real-time identification of airwriting and improve the usability of such systems in multiple contexts.

1.5 Aims and Objectives

The main objective of this research is to develop a deep learning-based framework for airwriting recognition using sEMG signals. The work in this paper uses Neural Networks (DNNs) for the realization of automatic learning and extraction from sEMG signals to develop meaningful features that are expected to boost the recognition accuracy of airwriting. Advanced signal processing techniques and a robust deep learning model would be used to address the problem identified by developing a framework with the potential to handle and reduce the sEMG signal variability, hence improving overall airwriting recognition system performance. The following are some of the ways that the proposed framework can address the identified problems and limitations. First, sEMG sensors are used allowing the system to record the electrical muscles' activity, which are engaged in the movements characterizing airwriting. This is much more accurate and detailed than other movements such as motion sensors for instance accelerometers or gyroscopes. By using sEMG it eliminates the need for spatial tracking and the adverse impacts of noise thus enhancing the accuracy in recognizing the characters. Second, the framework uses deep neural networks (DNNs) that can make predictions about non-linear and high-dimensional features, such as the fluctuations in muscle signals during airwriting. DNNs as opposed to other conventional algorithms, does not require feature engineering from raw data since it independently extracts features. This is for a purpose to improve the generalization across the different users; ensuring that variations in the writing speed, style and muscle mass are well addressed by the model. Training the network by exposing it with different data sets can also explain the variation between persons and enhance the general classification ability of the system. In addition, the framework is a real-time one and enables characters to be recognized as soon as the airwriting gesture required has been performed. This is attained through the signal preprocessing I have described and by attaining a short inference time in the deep learning model. The real time aspect of the system makes it very useful for such applications as interactive applications that involve virtual reality, augmented reality and even assistive technologies that require instantaneous feedback.

1.6 Research related to surface electromyography-based systems and airwriting recognition

The latest studies have also tried to expand the methods based on sEMG to dynamics, including air writing. For instance, Maier-Hein et al. in 2012 described an airwriting system in which inertial sensors fused with sEMG the characters written in the air[6]. While their system worked well it did not perform well in real-time and failed to stay consistent with users. Likewise, Li et al. (2019) worked on airwriting recognition using surface electromyography (sEMG) employing time domain features with k-nearest neighbor (KNN) classification [7]. Although there was an enhancement in the system, there was still an issue of consistency of writing speed and muscle fatigue as it reduced the rate of recognition. Gesture and airwriting recognition using sEMG-based interfaces has been targeted for research for a long time. The conventional procedure often includes feature extraction, usually followed by classification using machine learning algorithms like support vector machine (SVM) [8] and k-nearest neighbors (KNN) [9]. However, the methods did, in fact, call for high domain knowledge towards feature engineering and do not serve as general solutions for different kinds of users and conditions.

Recent advances in deep learning have brought closer the possibility of outdoing these limitations. Neural networks with CNNs have been found useful for the classification of sEMG signals and outperformed traditional methods. The automatic learning capability in CNNs has been strongly proven by different studies for learning relevant features from raw sEMG signals [10], [11], [12], hence reducing the need for manual feature extraction. However, there lies a further gap to be explored in CNN architectures, and new steps toward their optimization are absolutely needed to enhance robustness against signal variability. More advanced techniques known as 'deep learning' have in the recent past been incorporated in enhancing airwriting recognition systems. In a study [13], the authors used the CNN for classification of the airwriting characters based on the sEMG signals acquired from the hand. To enhance the spatial and temporal characteristics of muscle activities, the CNNs were incorporated in the model. However, the use of real-time processing and computing encountered limitations because of the computational issues concerning deep learning and its performances varied across subjects. In the same regard, another limitation witnessed is the need for massive training data to get desirable accuracy levels.

1.7 Importance of airwriting recognition by sEMG and DNNs

The further potential from this study resides in the contributions of airwriting recognition by utilizing surface electromyography (sEMG) and deep neural networks (DNNs). As airwriting recognition will have vast utility in low-proper HCI, assistive technologies, and AR, a system that comprises high recognition rate real-time airwriting recognition will highly boost uses of such technologies. This research also fills the existing knowledge gap which has been a subject of debate, especially due to the shortcomings of current systems in recognizing dynamic gestures in real time, that is, in the development of smart and efficient user interfaces. The above theoretical framework is especially significant when it comes to developing new assistive technologies, which could enrich the lives of people with speech or physical disabilities. Because airwriting allows real-time and non-verbal interaction, it can be useful for those who have mobility or speaking difficulties in navigating technology. Moreover, the prosthetic devices' users might benefit from the more sensitive and accurate signal especially when interacting with the environment, thus meaning an increase of their daily usability of the system.

The value of this research is a potential breakthrough in airwriting recognition and its applications. This contribution brought to light a deep learning-based framework to handle critical limitations in current airwriting recognition systems that are sensitive to signal variability and rely on manual feature extraction. The devised strategy has the end goal to enhance the robustness and accuracy of airwriting recognition to guarantee real-world application reliability. Improved airwriting recognition will create, through improved human-computer interaction, the ability for the use of digital devices to be conducted seamlessly and without physical contact. This will be especially important in environments such as AR and VR. It could be of help to the disabled for intuitive and efficient ways of communication, therefore granting greater independence and a higher quality of life.

The methods developed in this research could also apply to other areas of sEMG signal analysis such as monitoring muscle activity in rehabilitation programs or controlling prosthetic devices with better precision. It contributes to the development of novel input methods beyond traditional keyboards and touchscreens, assisting in new paradigms of user interaction with technology. This research develops new capabilities for sEMG-based applications by addressing most of the limitations and enhancing the performance of airwriting recognition systems. To this end, this work provides a solution that helps to overcome several of the difficulties with previous airwriting systems.

First, it improves accuracy compared to previous methods such as surface electromyography (sEMG) and deep neural networks (DNNs) capture detailed patterns of contraction and relaxation of the muscles in airwriting gestures. Due to the ability of deep learning models to handle large number of features in dataset, models derived from this approach provide better classification results as compared to conventional machine learning algorithms. Second, the proposed framework is aimed to be a real-time recognition schema that is possible by the signal preprocessing pipeline's optimization and the use of efficient deep learning architectures. This real-time capability is clearly superior to many other systems that gather data offline, resulting in a suitability for interactive systems used in augmented reality (AR), virtual reality (VR), and human-computer interaction (HCI). Third, the flexibility of the proposed deep learning models allows for the system to be able to deal with different user's writing patterns and speeds, as well as the different muscular inputs required to carry out the writing. This makes it possible to improve reliability and accuracy for various users of the result. Finally, in comparison with other frameworks, ours considers both opportunities for real-time performance during the experiment and the possibility of using it in practice, therefore its usability was higher. The system is designed to be implemented in wearable gadgets, AR terminals and support applications so that the result is as natural airwriting as possible for the user.

Thus, the present work presents a proposed framework as a solution to the major issues associated with current air writing recognition systems, which include efficiency, speed, flexibility, and ease of use. The proposed approach utilizes the surface electromyography (sEMG) signal and deep neural networks to analyze airwriting gestures at a muscle level that increases the accuracy of classification. Real-time performance and the ability to learn from one user and perform well with others makes it possible for the system to be applied in many practical applications such as human computer interaction, augmented reality and in assistive technologies. Overall, this thesis advances the topic of airwriting recognition and helps to outline a steadier and more manageable course of development for the proposed type of interface, which will be beneficial for users with different disabilities.

CHAPTER 2: LITERATURE REVIEW

Airwriting recognition includes the identification of characters or digits traced in the air with the hands, by the individual involved. I conducted a systematic search across established databases (e.g., PubMed, ACM Digital Library, IEEE, Xplore,) using related keywords such as "airwriting," "sEMG," "gesture recognition," "machine learning," and "deep learning."

Table 2.1: Overview of the current state of research in real-time airwriting recognition using sEMG and deep neural networks

#	Title	Key Findings	Methodology	Limitations
1	SurfMyoAir: A Surface Electromyography-Based Framework for Airwriting Recognition [14]	Achieved 78.50% accuracy in user-dependent and 62.19% in user-independent airwriting recognition using short-time Fourier transform and deep learning.	EMG signal recording, feature extraction, deep learning for classification	Limited vocabulary (English uppercase alphabets)
2	Hand gesture classification using time–frequency images and transfer learning based on CNN [15]	Demonstrated feasibility of real-time airwriting recognition with 86.2% accuracy using online EMG processing and CNNs.	Online EMG signal processing, CNN classification	High computational cost, limited dataset
3	SCLAiR: Supervised Contrastive Learning for User and Device Independent	Achieved high user and device-independent airwriting recognition (78.50% user-	Supervised contrastive learning with CNNs, two-stage classification	Requires labeled data, computationally expensive

	Airwriting Recognition [16]	dependent, 62.19% user-independent) using supervised contrastive learning with raw motion sensor data.		
4	ImAiR: Airwriting Recognition Framework Using Image Representation of IMU Signals [17]	Achieved high accuracy in airwriting recognition (89.20%) using 2D images generated from IMU data and CNN classification.	CNN classification with 2D IMU image representation	Limited user and device independence, 2D representation might lose information
5	Deep Learning Based Air-Writing Recognition with the Choice of Proper Interpolation Technique[18]	Proposed a hybrid model using multi-channel EMG and RNNs to achieve 91.3% accuracy in alphabet recognition and 76.4% in word recognition.	Multi-channel EMG recording, RNN for gesture decoding	Complex model requires substantial training data
6	Multi-Stroke handwriting character recognition based on sEMG using CNNs [19]	Utilized transfer learning from pre-trained CNNs for EMG signals, reaching 87.1% accuracy in character recognition.	Transfer learning with pre-trained CNNs, EMG feature extraction	Potential lack of domain-specific adaptation
7	EMG-Based Airwriting on Virtual Keyboards:	Evaluated different feature extraction methods for EMG	EMG feature extraction comparison, virtual	Limited focus on letter shape recognition

	A Comparative Study of Feature Extraction Techniques [20]	signals in virtual keyboard airwriting, achieving up to 92.5% accuracy with Mel-Frequency Cepstral Coefficients.	keyboard application	
8	On the Use of Temporal and Spectral Central Moments of Forearm Surface EMG for Finger Gesture Classification [21]	Developed a system using wrist-worn EMG for handwriting, demonstrating its viability for users with limited hand mobility.	Analyzed sEMG signals from three forearm sensors positioned in an armband configuration.	Relatively small dataset (20 participants).Limited gesture vocabulary (seven gestures). Not evaluated for real-time applications

2.1 Methodological Approaches

2.1.1 Surface Electromyography

sEMG might be able to capture the electrical signal that represents muscle contractions, thus offering a non-invasive gesture recognition technique. It is indeed more appropriate for dynamic gestures like airwriting because conventional methods fail when reference points are fixed. From existing literature, it can be determined that sEMG can successfully recognize air-written characters by combining activity patterns of their muscles with different gestures[22]. sEMG has found extensive application in gesture recognition, rehabilitation, prosthetics, and human-computer interaction due to its non-invasive nature. sEMG records electrical activity generated by skeletal muscles during movement, and its sensitivity to recognize subtle hand and wrist movements makes it preferable for the recognition of airwriting applications. Studies proved that sEMG could catch complex muscle signals corresponding to fine motor tasks, such as writing, further supporting the use of the gesture-based systems [23]. sEMG yields high-resolution information about muscle activity. Hence, the specific contractions related to writing tasks can be decoded that cannot be caught by vision-based systems. It has identified how muscle activities that track writing motions can be captured using sEMG electrodes on the forearm [24], which means that special activation patterns in forearm muscles are prompted by specific alphabetic gestures and allow for accurate characterization of air written characters.

2.1.2 Deep Learning Techniques

Convolutional Neural Networks (CNNs) also find applications in gesture recognition since it can learn spatial hierarchies of features[25]. Experimentations illustrate that CNNs outperform in high accuracy of airwriting gesture recognition on using data from cameras or sensors[4]. Long Short-Term Memory (LSTM) networks along with CNNs are also studied to consider the temporal dependencies in the data to improve performance. DNNs, particularly CNNs and RNNs, have already been largely applied for the analysis of sEMG data in gesture and airwriting recognition[26].

CNNs are good at capturing spatial dependencies in the sEMG signals, while LSTMs are skilled at modelling the temporal aspects. Unlike the traditional methods of machine learning, which rely on hand-crafted features, DNNs can autonomously extract hierarchical representations from raw sEMG data. CNN-LSTM hybrid model was successfully applied to time-series sEMG data for recognizing dynamic gestures[27]. In airwriting, Liu et al. (2020) demonstrated that CNNs learn local features such as muscle activation intensity and LSTMs learn the temporal sequence of movements corresponding to different alphabets[28].

2.1.3 Recognition of Real-Time Airwriting Characters

Low latency sEMG signal processing and accurate classification are also required for real-time airwriting recognition[29]. In fact, the processing speed of real-time sEMG-based systems should not be at the cost of accuracy; thus, DNNs are apt as they can classify complex patterns while keeping latency low. Real-time feedback is highly essential for an airwriting system, particularly in the context of certain applications, such as HCI and rehabilitation [30]. Instant recognition is needed to ensure seamless interaction between users and devices. Optimized architectures of neural networks and efficient signal preprocessing techniques, such as low pass filtering and denoising, are important. Yang et al. showed a system achieved real time by reducing the computational complexity while keeping high accuracy with lightweight CNNs[31].

Table 2.2: Basic key points of Airwriting Recognition System

Modality	Approach	Key points
Custom glove with sensors[32]	Hidden Markov Models + Language Model	This method uses a specially designed glove with sensors to capture hand movements, followed by Hidden Markov Models for character recognition and a language model to improve accuracy.
Wifi remote[33]	Hidden Markov Model	This approach relies on a Wi-Fi remote control as the input device, with Hidden Markov Models for gesture recognition.
Wrist-worn inertial sensor[34]	Convolutional Neural Network	This method employs a wrist-worn sensor to measure hand motion and utilizes a Convolutional Neural Network for gesture classification.
Wrist-worn inertial sensor[16]	Supervised Contrastive Loss	This approach features a wrist-worn sensor and leverages a supervised contrastive loss function for improved gesture recognition accuracy.
Wrist-worn inertial sensor[17]	Image representation of signals	This method converts the sensor data into images and then uses image-based recognition techniques, such as CNNs, for gesture classification.
Computer vision[35]	Spatio-temporal residual architecture	This method uses computer vision techniques with specific neural network architecture (spatio-temporal residual) to recognize gestures based on video recordings.
Computer vision[36]	Region-based CNN	This approach relies on computer vision and Region-based Convolutional Neural Networks to identify gestures from video data.

2.2 Theoretical Framework

2.2.1 Temporal Trajectory Analysis

This approach bases its interest on tracking the trajectory of hand movement across the three-dimensional space while a user writes in air[37]. It argues that unless the traces of such were known, it would be impossible to identify characters since such traces differ significantly compared to static writing styles. The theory emphasizes the fact that in the design of such recognition systems, there is a high demand for the ability of algorithms to change according to the varied writing styles and the speed of different users[38].

2.2.2 Multi-Loss Minimization Framework

Some recent studies develop frameworks that minimize several loss functions simultaneously while training the model[39]. Such a framework is targeted for enhancing feature embedding and classification accuracy in sEMG-based airwriting recognition systems[40]. Optimization of several aspects of the model simultaneously has helped researchers achieve some better performance metrics compared to traditional single-loss models[41].

2.3 Harmonizing Approach

2.3.1 Integration of Various Technologies

Combining sEMG with other sensing technologies, such as radar or inertial sensors, has been proposed to enrich the accuracy of recognition and the user experience[42]. An alternative possibility for the use of radar is to detect hand movements without physical contact, which provides an intuitive interface for users[43]. The harmonization of various methods allows for better strategies for data capture and processing and leads to better real-time performance.

2.3.2 User-cantered design

Most of the papers focus on designing systems in ways that accommodate individual user differences in writing styles and preferences[44]. The aim of such an approach to the user is both to improve accuracy and to please and involve the user with the technology. In ensuring effective airwriting systems, it is critical to align the sEMG acquisition with deep learning models. The real-time airwriting systems require the right balance between signal processing, neural network computation, and user comfort[45]. Poor integration may produce high latencies or errors in recognition while thereby reducing user experience. Optimization of preprocessing steps and selection of lightweight DNN architectures, such as shallow CNNs, ensures that the airwriting system remains responsive while at the same time providing high classification accuracy[46].

2.4 User Variability and Generalization

One of the most significant problems with airwriting recognition, however, is its user dependency in terms of writing as data augmentation techniques enhance generalizability across different users[47]. Writing styles can differ dramatically between people, and an individual even writes letters differently under different contexts. Generalization across users ensures that the system can be applied to a wide range of individuals. Variability has largely been approached with data augmentation and transfer learning. Transfer learning is used to port models that were trained from one class of users to the target unseen ones and argued that if it were adapted well[48], it could greatly improve performance in sEMG-based air-writing recognition. I also investigated citations of key papers and explored related conference proceedings. Studies were included if they were focusing on airwriting recognition using sEMG signals. Quantitative results for character or gesture recognition accuracy are reported

Table 2.3: Inclusion/Exclusion Criteria and Literature Search

Sr	Database	Inclusion Criteria	Exclusion Criteria
1	PubMed	Focused on airwriting recognition using sEMG signals. Reported quantitative results for character or gesture recognition accuracy. Employed a clear methodology with details on data acquisition, feature extraction, model selection, and evaluation techniques. Published	Relied solely on modalities other than sEMG (e.g., vision, motion tracking). Used simulated or synthetic data without validation on real user data. Did not report quantitative results for airwriting recognition. Focused on static gesture recognition.

		within the last five years.	
2	IEEE Xplore	Focused on machine learning or deep learning approaches for airwriting recognition using sEMG .Reported results on real-time or online recognition systems (preferred).	Relied solely on traditional signal processing techniques Focused on offline analysis or pre-recorded data.
3	Research gate	Relevant to airwriting recognition research (e.g., studies on similar tasks like handwriting recognition	Not focused on sEMG or motor control aspects Primarily review papers or theoretical work.
4	Google Scholar	Broader search to identify potential studies not indexed in other databases. Focused on the development of user-friendly and accessible airwriting interfaces.	Duplicate entries from previous searches. Primarily engineering-oriented studies without strong evaluation of recognition accuracy.

CHAPTER 3: METHODOLOGY

The surface electromyography (sEMG) signals used in the proposed system for airwriting identification involve multiple crucial steps, including data preparation, signal preprocessing, feature extraction using a CNN, and classification. These steps are described in Figure 3.1.

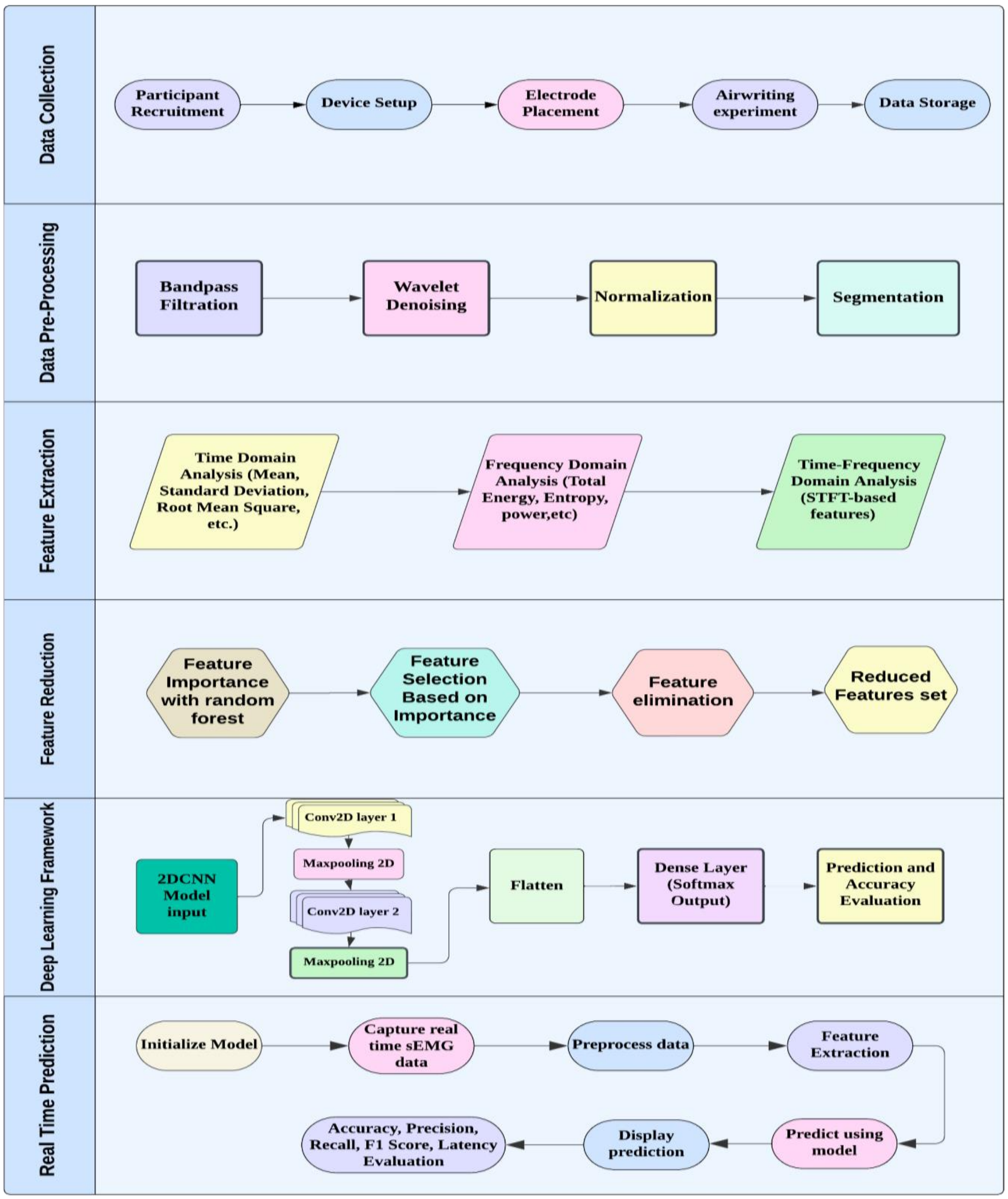


Figure 3.1: Process involve in developing and implementing Real-Time airwriting recognition system

The system's overall performance depends on each of these phases. The details of the methodology adopted for developing and implementing this proposed system will be discussed in the succeeding chapter. This research aims to design an alphabet recognition system using real-time sEMG and deep learning models. The chapter deals with the steps taken from data acquisition, preprocessing, and feature extraction up to the training, testing, and validation phases of the machine learning models. Figure 3.2 shows a block diagram of proposed framework. A thorough explanation of the process is provided below.

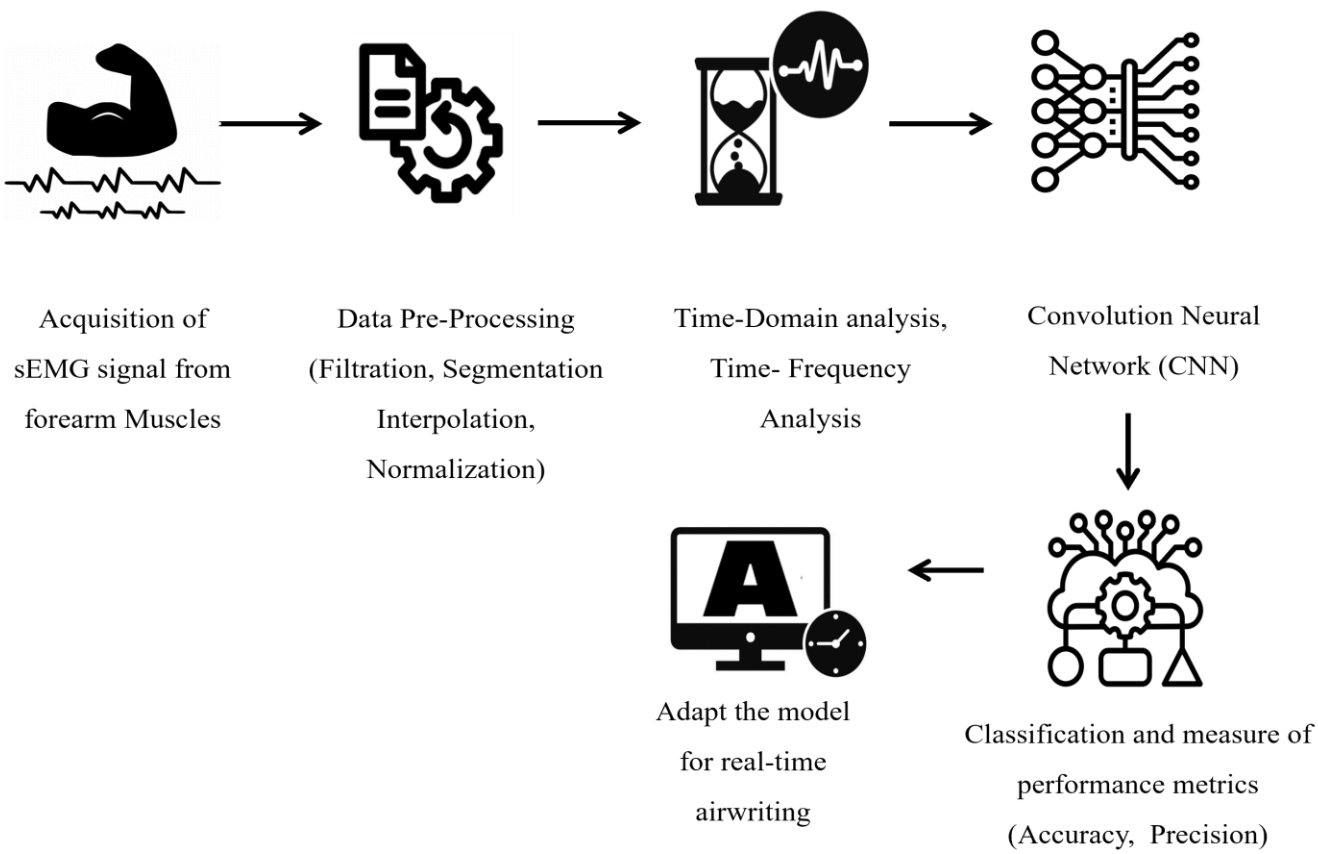


Figure 3.2: Block Diagram of proposed system for airwriting identification

3.1 Data Collection

Data collection is an integral part of research because it provides the basic sEMG signals required for training and testing the deep neural network for airwriting recognition in real time. The following are the structures of how I use participants, devices, tasks, and storage protocols in my research.

3.1.1 Participant Recruitment

In this study, 16 healthy individuals; 6 male and 10 female aged between 20 and 35 years participated. The participants were screened for pre-existing neuromuscular or motor impairments. Therefore, all participants were guaranteed to be free of such impairments leading to a homogeneous sample that would allow the collection of accurate sEMG signals during airwriting tasks.

3.1.2 Device Setup

sEMG signals were recorded from the forearm muscles of participants with the Delsys Trigno wireless sEMG system; this system was chosen because it has high precision and is wireless, thus in no way limiting participants' freedom of airwriting while at the same time affixing cables to physical bodies.

Figure 3.3: Trigno Base Station(left), EMG electrodes (Right)

3.1.3 Electrode Placement

sEMG electrodes were placed over the flexor and extensor muscles of the forearm. Since these muscles mostly control the movements of the hand and fingers as needed for writing, their activities have been adequately captured during the task of airwriting on the part of the participant.

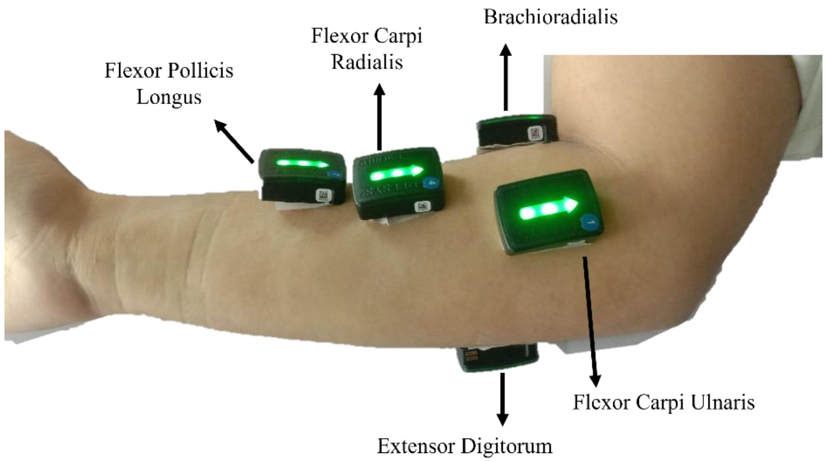


Figure 3.4: Configuration of electrode placement on the muscles of the participant

3.1.4 Airwriting Experiment

Participants had to write out uppercase English alphabets (A–Z) in a controlled setting. Each of the participants performed the task twice for every letter, so in total, there would be 52 samples per participant (2 trials for each letter). This scheme allows changes between trials to occur, in effect increasing the amount of information gotten from the dataset. Each experiment was conducted for 2 to 4 seconds based on how complex the word written might be. The experimenter allowed for as much time as necessary to ensure the writing movements to be smooth and not impeded. The sEMG signals of the airwriting experiments were in `. npy` format, since such multidimensional time series data are very efficiently stored in this format, which is beneficial for further processing.

3.1.5 Data Storage

The obtained data was systematically structured with a folder structure. Each participant's folder got assigned a folder, labelled by participant number (e.g., `Participant_1`, `Participant_2`, etc.). Inside each of these participant folders, the trials were saved under a standardized naming convention based on the letter and trial number (e.g., `A_TRIAL_1.npy`, `A_TRIAL_2.npy`, etc.). The `. npy` files contained raw sEMG signals sampled at a high sampling rate of 2000 Hz, which would help in having an accurate fine temporal resolution of muscle activity during the execution of each airwriting task. This offered very rich data capture that adequately helped in creating a robust dataset for the subsequent stages of preprocessing, feature extraction, and development of deep learning models.

3.2 Data Collection Protocol

Through the GUI developed from the use of the Tkinter library in Python, the process of data collection is made easier and automated for the sEMG signals. The GUI was, therefore, designed to aid subjects in providing trials related to the collection of muscle signals corresponding to hand gestures of writing alphabet characters. The system is integrated with the Delsys Trigno EMG system for the recording of surface electromyography from the muscles of the forearm. This stage primarily involved recording sEMG signals while subjects were writing capital English alphabets from A to Z. The application transmitted on the GUI screen instructions that flashed one character at a time with 2 s of muscle contraction and 1 s of relaxation. This type of pattern ensured that the same data is recorded uniformly among the trials.

3.2.1 GUI Design and Features

The features that emerged on the GUI were as follows:

- *Character Display:* The GUI displayed every character in a large readable format where the subject managed to write during the trial. Once the muscle contraction phase ended, it automatically transitioned into the rest phase.
- *Timer:* A timer came out as a countdown which provided the visual cues on the remainder of time left for each trial.
- *Progress Bar:* A progress bar was added to update the user on the completion status of the trial sequence.
- *Auto Saving of Data:* For each trial after its completion, the data is saved automatically in `.npy` format to store multidimensional EMG signals in an optimal fashion.

The Delsys Trigno EMG system was connected to the application, continuously acquiring sEMG signals with a sampling rate of 2000 samples per second. Signal acquisition employed 5 channels, processing data in real-time. For each subject, 52 samples were collected, as two trials were conducted for every one of the 26 characters. For every trial, the system recorded EMG signals from the active forearm muscles during both contraction and rest phases.

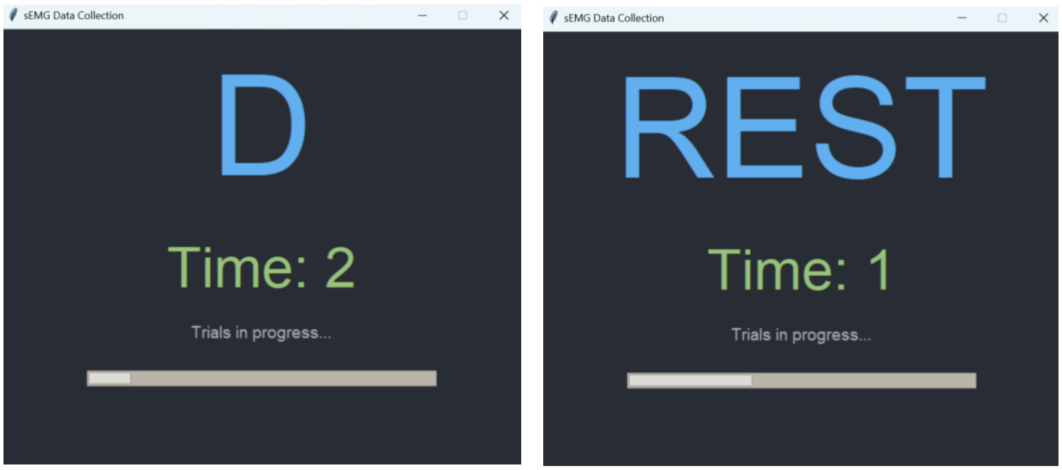


Figure 3.5: GUI Protocol for data collection showing displaying character for 2sec and then rest of 1sec

3.3 Data Preprocessing

3.3.1 Noise Removal

There are two main methods used in the removal of unwanted noise or artifacts from the raw sEMG signals: band-pass filtering and wavelet denoising.

3.3.1.1 Band-pass filtering

A band-pass filtering operation picks out a specific range of frequencies in the signal. For sEMG data, this is between 20–450 Hz. The lower frequency range below 20 Hz contains mostly motion artifacts and low-frequency noise while the higher frequency range above 450 Hz may contain high-frequency noise such as electrical interference. The signal is filtered so that only the retained data corresponds to actual muscle activity, which occurs in that frequency range.

3.3.1.2 Wavelet Denoising:

This method uses DWT in the sEMG signal into multiple components of frequency referred to as wavelets. Depending upon the application of thresholding to wavelet coefficients, desirable or essential noise reduction could be achieved without losing the fundamental properties of the signal significantly. Importantly, wavelet denoising captures the nonstationary signals such as sEMG quite well because noise can be separated from muscle activity with various frequency scales that conventional filtering techniques may not do very easily. Wavelet denoising is one of the best signal quality improvement techniques[49], [50]. All these steps, combined in the process, can be summarized mathematically as follows: Select the wavelet function (for instance, Daubechies), and the level of decomposition N, Compute the wavelet decomposition of the noisy EMG signal s,

$$s = \sum_{j=0}^{N} A_j + \sum_{j=1}^{N} D_j \tag{3.1}$$

where A_j are the approximation coefficients and D_j are the detail coefficients.

Apply a threshold T to the detail coefficients for noise suppression. The most used techniques are: Soft Thresholding, hard thresholding

$$D_j^{denoised} = sgn(D_j).\max(|D_j| - T, 0) \tag{3.2}$$

$$D_j^{denoised} = \begin{cases} D_j & if\ |D_j| > T \\ 0 & otherwise \end{cases} \tag{3.3}$$

$$s_{denoised} = \textstyle\sum_{j=0}^{N} A_j + \sum_{j=1}^{N} D_j^{denoised} \tag{3.4}$$

3.3.2 Normalization

To make the data comparable from one participant to another participant and from one trial to another trial, Z-score normalization is applied on sEMG signals. In brief, normalization is a way to rescale the data having zero mean and unit standard deviation.

3.3.2.1 Z-score normalization

It is the technique that standardizes the amplitude of the signals. Following the subtraction of the mean, it divides the signals by the standard deviation of the data. In that regard, differences in muscle force, electrode placement, and the physical attributes of the human body were considered such that all signal values fall within a similar scale across all subjects and trials. This is particularly crucial in gesture recognition by sEMG as the variance of muscle activities between different subjects tends to have an impact on the performance of the model. Normalization of data allows the algorithm to focus more on what patterns to recognize that are associated with an airwriting activity rather than the differences in individuals.

3.3.3 Segmentation

The method after filtering and normalization of the signals of sEMG is to divide it into smaller time domains for further analysis. It is known as windowing.

3.3.3.1 Overlapping windowing

The continuous sEMG signal is divided into continuous, overlapping windows of fixed length, a duration of 200 ms with 50% overlap. That is, the signal is divided into 200 ms intervals and each interval is considered separately for further processing. A window size of 200 ms is widely accepted in gesture recognition tasks as it is large enough to capture all the relevant muscle activity for writing motions but not that large to lose the temporal resolution between the movement and differentiate between different letters. Data within those windows, the model should be able to discern patterns compatible with gestures or movements to recognize what letter one is airwriting.

The formula used for the sliding window is as follows:

$$window\ size = \frac{sampling\_rate \times windowsize(ms)}{1000} \tag{3.5}$$

This resulted in the production of a great number of windows per signal; it would depend on the overall length of the signal.

3.4 Feature Extraction Process

The research applies feature extraction on raw EMG signals to convert them into meaningful representations which will be used as inputs for machine learning models in recognizing alphabets. This will comprise the time-domain, frequency-domain, and time-frequency domains in the extraction process through applying the sliding window technique in segmenting the data.

3.4.1 Time-Domain Features

Time domain features were computed directly from the EMG signals in each window. They represent the basic statistical and morphological characteristics of the signal, such as amplitude, variance, and energy. Some of the most important extracted features are listed hereunder:

3.4.1.1 Mean

It is the average value that the signal has taken over the window.

$$Mean = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{3.6}$$

Where, N= number of samples in the window , X_i = Amplitude of the i-th sample

3.4.1.2 Standard Deviation

This represents the estimate of the variability of the signal values around their mean.

$$Standard\ Deviation = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{3.7}$$

Where, μ = Mean of Window

3.4.1.3 Root Mean Square (RMS)

RMS is the magnitude of the signal.

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i)^2} \tag{3.8}$$

3.4.1.4 Mean Absolute Deviation (MAD):

Mean Absolute Deviation (MAD) is the average absolute deviation of the signal from its mean.

$$MAD = \frac{1}{N}\sum_{i=1}^{N} |x_i - \mu| \tag{3.9}$$

3.4.1.5 Skewness

Skewness is a way of measuring how asymmetric the distribution of the signal is.

$$Skewness = \frac{1}{N}\sum_{i=1}^{N}(\frac{x_i - \mu}{\sigma})^3 \tag{3.10}$$

Where, σ =standard deviation.

3.4.1.6 Kurtosis

Kurtosis is a way of measuring how pointed the distribution of the signal is.

$$Kurtosis = \frac{1}{N}\sum_{i=1}^{N}(\frac{x_i - \mu}{\sigma})^4 \tag{3.11}$$

3.4.1.7 Autocorrelation

Autocorrelation denotes similarity of the signal with a lag of one sample to capture time dependencies.

$$Autocorr_{lag1} = \frac{1}{N-l}\sum_{i=1}^{N}(x_i - \mu)(x_{i+1} - \mu) \tag{3.12}$$

Where, l=Lag=1

3.4.1.8 Zero and Mean Crossings

Zero and Mean Crossings count the number of zero or mean crossings of the signal to have an idea about how oscillatory the signal is.

$$ZC = \sum_{I=1}^{N-1} 1(x_i . x_{i+1} < 0) \quad (3.13)$$

1 =indicator function =1, if the condition is true.

These time-domain features can capture the shape and behaviour of the signal as well, which is very crucial in the differentiation between the letters written in terms of sEMG signals.

3.4.2 Frequency-Domain Features

To analyze the spectral properties of the signal, frequency-domain features were extracted using the FFT. FFT transforms a signal from its time domain to its frequency components; hence, it reveals information on the periodicity and oscillations of the signal. The formula for FFT is:

$$X(k) = \sum_{n=0}^{N-1} x(n). e^{-j\frac{2\pi}{N}kn} \text{ , k=0,1, ………,N-1} \quad (3.14)$$

Some of the key frequency-domain features feature are:

3.4.2.1 FFT Energy

Total power in the signal across all its frequency components.

$$FFT\ Energy = \sum_{k=1}^{N} |X(k)|^2 \quad (3.15)$$

3.4.2.2 Maximum Power Frequency

It is the highest frequency at which the highest power value of the signal will come.

$$Max\ power\ Freq = \arg\max\ |X(k)| \quad (3.16)$$

3.4.2.3 Mean and Median Power Frequencies

This gives insight into the distribution of power across different values of frequency.

$$Mean\ Freq = \frac{\sum_{k=1}^{N} k.\,|X(k)|^2}{\sum_{k=1}^{N} |X(k)|^2} \tag{3.17}$$

3.4.2.4 Spectral Band Powers

It is the power of the signal within those different bands of frequencies, delta, theta, alpha, beta, and gamma bands, respectively, which are very important while analyzing the specific physiological or cognitive sEMG process.

$$Band\ power = \sum_{k=k_1}^{k_2} |X(k)|^2 \tag{3.18}$$

Where, k_1 and k2 = Frequency range for each band.

As can be seen, these frequency domain features are indeed useful, regarding capturing the frequency content, of muscular movements, just like the time domain features.

3.4.3 Time-Frequency Domain Features

Besides the characteristics in the time and frequency domains, time-frequency representations were used to grasp properties that account for both temporal and spectral information simultaneously. STFT was applied to each window. The time-frequency map described how the temporal behaviour of the signal varies over frequencies. It basically defines the average and spread of the magnitude of the signal in the time-frequency domain.

$$STFT\{x(t)\}(n,f) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)e^{-j2\pi fm} \tag{3.19}$$

w(n) =window function (Hann).

Some features that could be derived from the STFT are as follows:

3.4.3.1 STFT Entropy

This is the measure of randomness of the signal in the time frequency domain.

$$STFT\ Entropy = -\sum_{n=1}^{N} |Z(n,f)| \log (|Z(n,f)| + \epsilon) \quad (3.20)$$

3.4.3.2 STFT Mean Frequency and Phase

These capture the mean frequency and phase information of the signal as time evolves.

$$STFT\ Mean\ Magnitude = \frac{1}{N}\sum_{n=1}^{N} |Z(n,f)| \quad (3.21)$$

These signals sEMG can then be accounted to their non-stationary nature using time-frequency domain features, whereby the signal frequency components change dynamically with the variation in different hand movements.

3.5 Recursive Feature Elimination

Recursive feature elimination is a feature selection method, by recursively employing any learner to select a smaller subset of features. It works in an iterative form of training a model-with an example given on the Random Forest-and removing the worst feature at each iteration until a few desired features are reached. It was done through the method of Random Forest Feature Importance since this technique had shown potential in calculating the relevance of features in ensemble learning models as it measures how much of a reduction in impurity occurs for a feature on an ensemble of decision trees.

The Random Forest classifier was trained on the 26 letters of the alphabet based on features obtained from time-domain, frequency-domain, and time-frequency domain analyses of the voice signals. At training time, each feature was calculated as the reduction in Gini impurity if that feature were used to split at some point in a decision tree. The importance score for each feature during the testing phase, time information about the system was recorded. For the entire dataset, the mean importance score was found, and the ranking of the top 10 features was done in descending order.

$$I(f_j) = \sum_{t=1}^{T} \sum_{n=1}^{N_t} \Delta I_n(f_j) \quad (3.22)$$

where: T = Number of decision trees in the forest, N_t = Number of nodes for each tree t,

$\Delta I_n(f_j)$ = Reduction of Gini impurity on node (n) when splitting based on the feature (f_j)

The normalized scores of importance ranked features by relevance. Therefore, the higher the score of the features, the more influential they were for the increase in accuracy of classification, and the lower the score, the lesser the contribution, and because of that, they were considered for possible exclusion from the model to streamline it. I had to employ Random Forest Feature Importance, whereby I could extract the most significant features for consideration and even eliminate less useful or redundant ones while further optimizing performance in the model. That was improving efficiency and accuracy regarding a classification model while minimizing the dimension of data representation, focusing on important features that would indeed cause effects on an airwriting recognition task.

3.6 Deep Learning Architecture

At the heart of the deep learning structure was a multi-layered architecture that captures the complexity of the patterns in the EMG signals. They included various layers:

3.6.1 Convolutional Neural Network (CNN) Layers

Convolutional layers were used for the automatic extraction of local patterns and spatial hierarchies from the input data. CNN layers can be thought of as feature detectors that are able to capture both low-level and high-level information in the data. The architecture is built from the layers of CNN, learning the spatial patterns of EMG signals. The preprocessing layer of the EMG data automatically detects both low-level and high-level features in them.

3.6.1.1 Input Shape

The input shape for the pre-processed EMG signal fed into the 3D CNN as arrays where each input represents a time-frequency representation after STFT for a specific channel of EMG signals.

3.6.1.2 Convolution Operation

The convolution application places spatial filters over the input data, producing feature maps. In mathematical terms, this operation is described as:

$$S(i,j) = \sum_{m=1}^{M} \sum_{n=1}^{N} X(i+m, j+n).W(m,n) + b \tag{3.23}$$

Where, X = Input EMG signal, W= learned filter,

S(i, j) = Resulting feature map, b = bias term.

Max-pooling was applied after the convolution operation for dimensionality reduction with all the features needed.

3.6.2 Training Process

The framework utilized a labeled EMG signal dataset corresponding to alphabet letters, splitting the dataset into three sets: training set, validation set, and test set. Since the cross-entropy loss function is going to be used in model training. Adam optimizer with learning rate adjustment based on moment estimates was used to optimize the deep learning model. Dropout regularization was further added by randomly deactivating a proportion of neurons in the training to account for overfitting.

3.6.3 Evaluation Metrics

Accuracy, precision, recall, and F1-score values were used to validate the model on the test dataset. Moreover, confusion matrices were calculated in order to evaluate classification errors for all cases of different alphabet letters.

3.7 Real Time Alphabet Prediction

Real-time classification is the capacity of a trained deep learning model to classify incoming EMG signals immediately when they are acquired. Usually, such ability requires multiple critical stages.

3.7.1 Signal Acquisition and preprocessing

The consecutive collection of EMG data from sensors or devices in real time. Such data is typically streamed into a processing unit (such as a computer or microcontroller).Incoming signals apply to the same preprocessing as in the training phase, such as:

- Normalization
- Filtering techniques are used as a denoising process
- Resample to a specified shape as expected from the input by the model.

3.7.2 Feature Extraction

It involves extracting the relevant features from the preprocessed signals, often using techniques such as STFT or wavelet transforms, thus transforming time-domain signals into frequency domain representations.

3.7.3 Model Classification and Output Display

The pre-processed feature-extracted data is passed to the trained deep learning model. As a result of its training, the model now outputs each class's probability. In this experiment, the class that obtained the highest probability is the class used as the output by the model. One would always want to see the output in the real-time sense. This could be done with visual displays: a text changing within a certain range, lighting effects once it has been recognized, etc.

3.7.4 System Architecture

To classify, a CNN-based deep learning model was developed to develop the real-time prediction model. For signal acquisition, Delsys Trigno EMG was used in combination with 8 channels at 2000 samples per second. As preprocessing techniques before applying the trained model for prediction, bandpass filtering was conducted at 20-450 Hz, and notch filtering removed the powerline interference of 50 Hz, while wavelet denoising ensured the improvement of the signal-to-noise ratio and removal of artifacts from the signals. These preprocessed signals were then formatted for the model's prediction.

3.7.5 Real-Time Prediction Procedure

The screen displayed the character of an alphabet that the participants tried to write. This process was accompanied by the collection of sEMG signals. The real-time prediction was based on them. There is a 2-second delay between each character when the deep learning model predicts based on the most recent batch of EMG data. The target and predicted characters are then presented to the user immediately afterward on the system's screen.

3.8 GUI Design for Real-Time Prediction

A good GUI in real-time prediction should mainly be intuitive, responsive, and presentable. The following are some basic GUI elements for real-time EMG signal prediction:

3.8.1 Layout

The interface should present a clean and organized layout of real-time information, predicted letters, and status indicators provide plots or graphs of the real-time EMG signal waveforms, to be viewed by the user in displaying their own signals' waveforms.

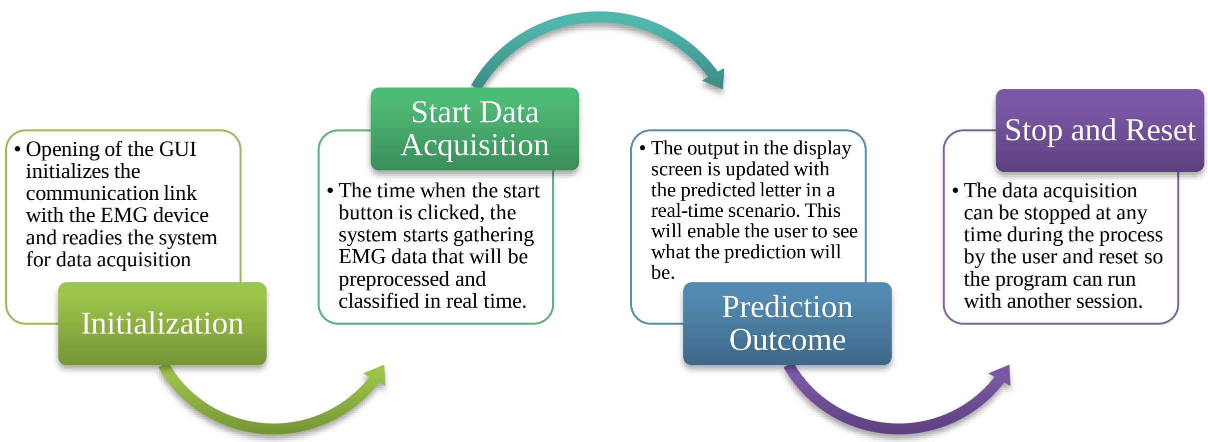


Figure 3.6: GUI Workflow

3.8.2 Input Controls and Output Displays

A button by which the user may start and stop the data acquisition. Offers the user an opportunity to reset prediction state, if desired. A prominent text area that shows the most recently predicted letter. optionally, print the confidence score of the prediction, like the probability of the predicted class.

3.9 Performance Matrix

In deep learning, basically, the performance matrix often refers to the collection or table of multiple performance metrics for evaluating the effectiveness of a model. The values include accuracy, precision, recall, F1 score, specificity, among others, all derived from the confusion matrix. The matrix also serves as a structured way in comparison with the performance of various models on multiple metrics.

- Interpretation of the results. It tries to point out inadequacies in the models. For example, too few positives are missed by the model if recall is too low.
- Optimizing Models. This can be used to tune in terms of specific priorities-for example improved precision at the cost of decreased recall if some types of errors are more expensive than others.

3.9.1 *Accuracy*

The most intuitive of performance measures is that of accuracy, and it is defined as the number of correct predictions divided by all predictions. However, accuracy can sometimes give a false sense of great performance in an imbalanced dataset where one class occurs much more frequently than others.

$$Accuracy = \frac{\boldsymbol{True\ Positive(TP)+True\ Negatives(TN)}}{\boldsymbol{Total\ Predictions}} \tag{3.24}$$

$$= \frac{\boldsymbol{TP+TN}}{\boldsymbol{TP+TN+FP+FN}} \tag{3.25}$$

3.9.2 Precision

Precision is also referred to as Positive Predictive Value. It measures the number of correct positive predictions the model is making. This metric is useful when there is a significant cost for false positives. High precision is where the model correctly classifies positive. Precision is especially useful when there exist examples like a medical test where a false positive entails unnecessary additional tests.

TP (True Positives): The model correctly predicted the positive class.

TN (True Negatives): The model correctly predicted the negative class.

FP (False Positives): The model wrongly predicted the positive class.

FN (False Negatives): The model wrongly predicted the class to be negative.

$$Precision = \frac{TP}{TP + FP} \quad (3.26)$$

3.9.3 *Recall*

Recall, also known as Sensitivity or True Positive Rate, illustrates how good the model is in classifying positive cases. It can be useful when the cost of missing positive cases, which is false negatives, is high for example, in fraud detection or in medical diagnostics. High recall means the model accurately predicts most of the actual positive cases. Recall is useful when a false negative has serious implications.

$$Recall = \frac{TP}{TP + FN} \quad (3.27)$$

3.9.4 *F1 Score*

The F1 Score is the harmonic mean of precision and recall; hence it means making a balance between both values. Where there are imbalanced datasets, it proves useful to create a trade-off between precision and recall.

$$F1 = 2.\frac{Precision \times Recall}{Precision + Recall} \quad (3.28)$$

3.9.5 *Confusion Matrix*

A Confusion Matrix is a tabular representation which provides a great idea of how well a model is performing. The same reports give the number of true positives, true negatives, false positives, and false negatives, which are all crucial in the calculation of the other metrics.

CHAPTER 4: RESULTS

The aim of this research work was to build and test an air written English alphabet recognition system in which that has been achieved through the representation of the sEMG signals from the muscles of the forearm. This is done by using deep learning for classification without any error. It has results associated with the effectiveness of preprocessing techniques, feature extraction methods, and performances associated with different deep models for recognizing air written characters with accuracy, besides the effect of signal variations across participants and different writing styles.

4.1 Data Collection and Pre-processing

Sixteen subjects were recorded to collect the sEMG signals, each of whom wrote the 26 alphabets in uppercase English for 10 repetitions in thin air. The muscle activities of the forearm were recorded by surface electromyography sensors. Figure 4.1 illustrates the signal for Subject 1, 'W_TRIAL_1.npy'.

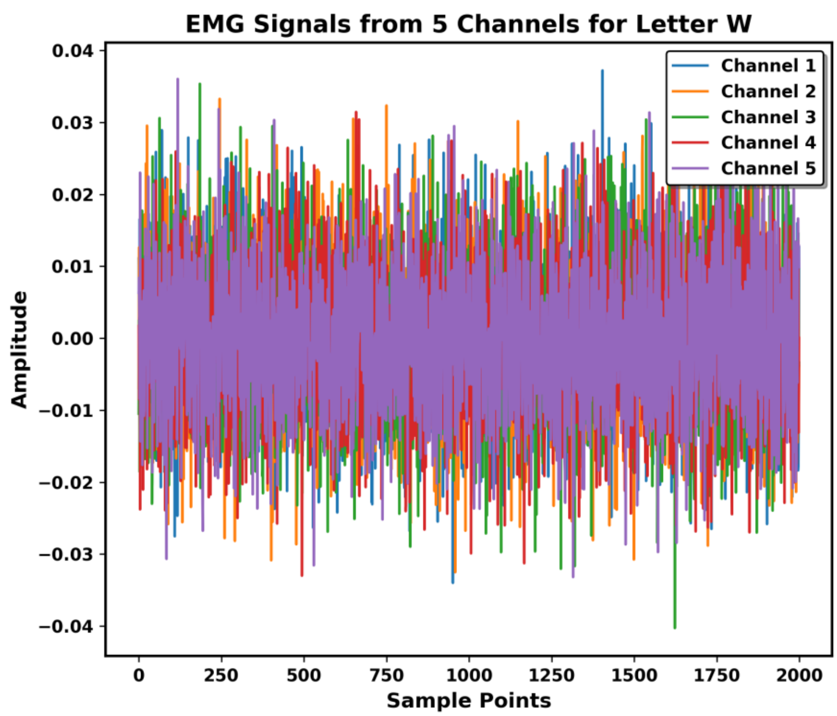


Figure 4.1: Superimposed EMG signal of 5 channels for alphabet W

It was a preprocessing stage wherein lots were done to reduce the noise and make sure the data is like each other. The process of filtering these signals begins with a band-pass filter at cutoff frequencies ranging from 20 Hz to 450 Hz to eliminate strong low-frequency noise and high-frequency artifacts. Moreover, it incorporates notch filtering to filter out 50 Hz of powerline interference. Finally, wavelet denoising techniques are applied further to clean the signals by preserving significant patterns and discarding random noise. Finally, it normalized the data to a range of 0 to 1 so that the data used in the tests, and the number of subjects, become uniform and were subjected to more certain feature extraction as well as the training of the model. Figure 4.2 illustrates the filtered signal for 'B_TRIAL_1.npy'.

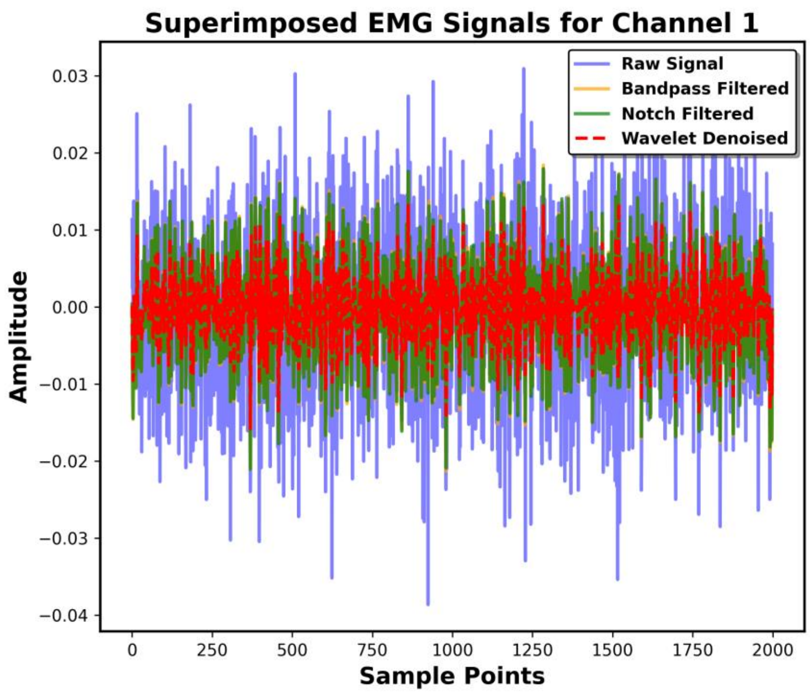


Figure 4.2: Superimposed Signal for Raw EMG signal, Bandpass filtration, notch filtration and denoised signal for alphabet B

4.2 Feature Extraction

This processing results in a wide variety of time-domain, frequency-domain, and time-frequency features through the sliding window approach feature extraction process. The sliding window size was set to 100 ms with an overlap of 50% for each signal, ensuring that all the signals analyzed formed a comprehensive representation of the muscle activity captured during airwriting gestures. Figure 4.3 presents all the quantities obtained from the time domain, frequency domain and time-frequency domain of the electromyography signals.

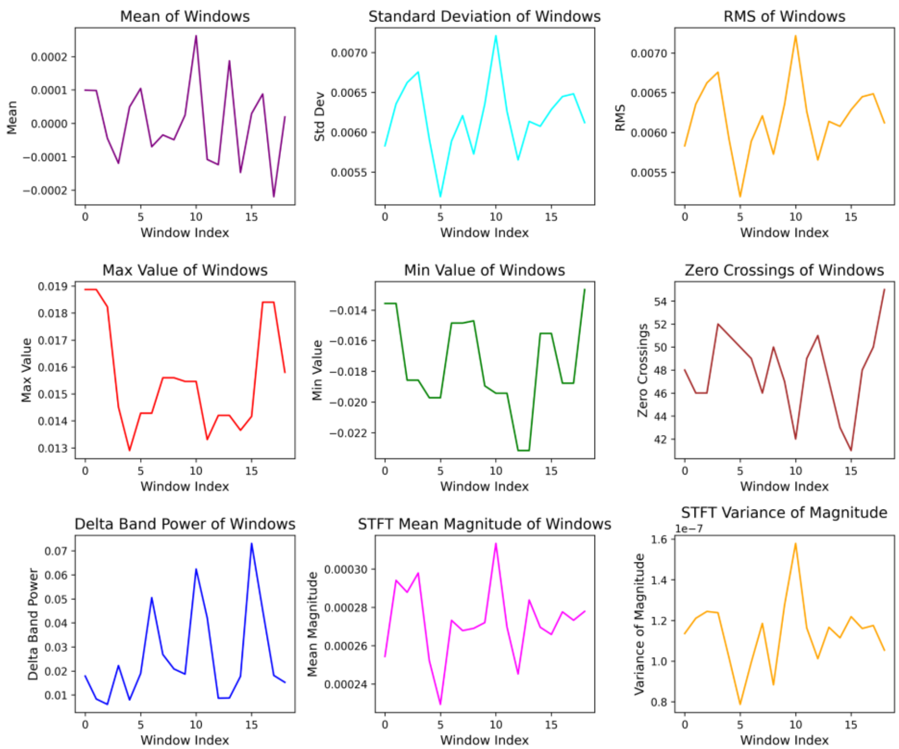


Figure 4.3: A windowed segment of signal from a single sEMG electrode

The picture presents all the key characteristics such as mean magnitude, variance of magnitude, maximum magnitude indicating overall strength and variability of muscle contractions, the mean frequency, and the entropy providing information on preferred frequencies as well as on complexity of signal respectively. More insights can be gained about the timing and variability of muscle activations from these measures, mean phase, and phase variance. Hence, these features become important for analyzing hand movements from EMG signals and further deepen knowledge about the dynamics in muscles involved with different tasks.

4.2.1 Correlation Matrix

This heatmap correlation matrix provides the abundant context of linear relationships between various features extracted from signal data as well as to what extent one feature may be related to the target variable.

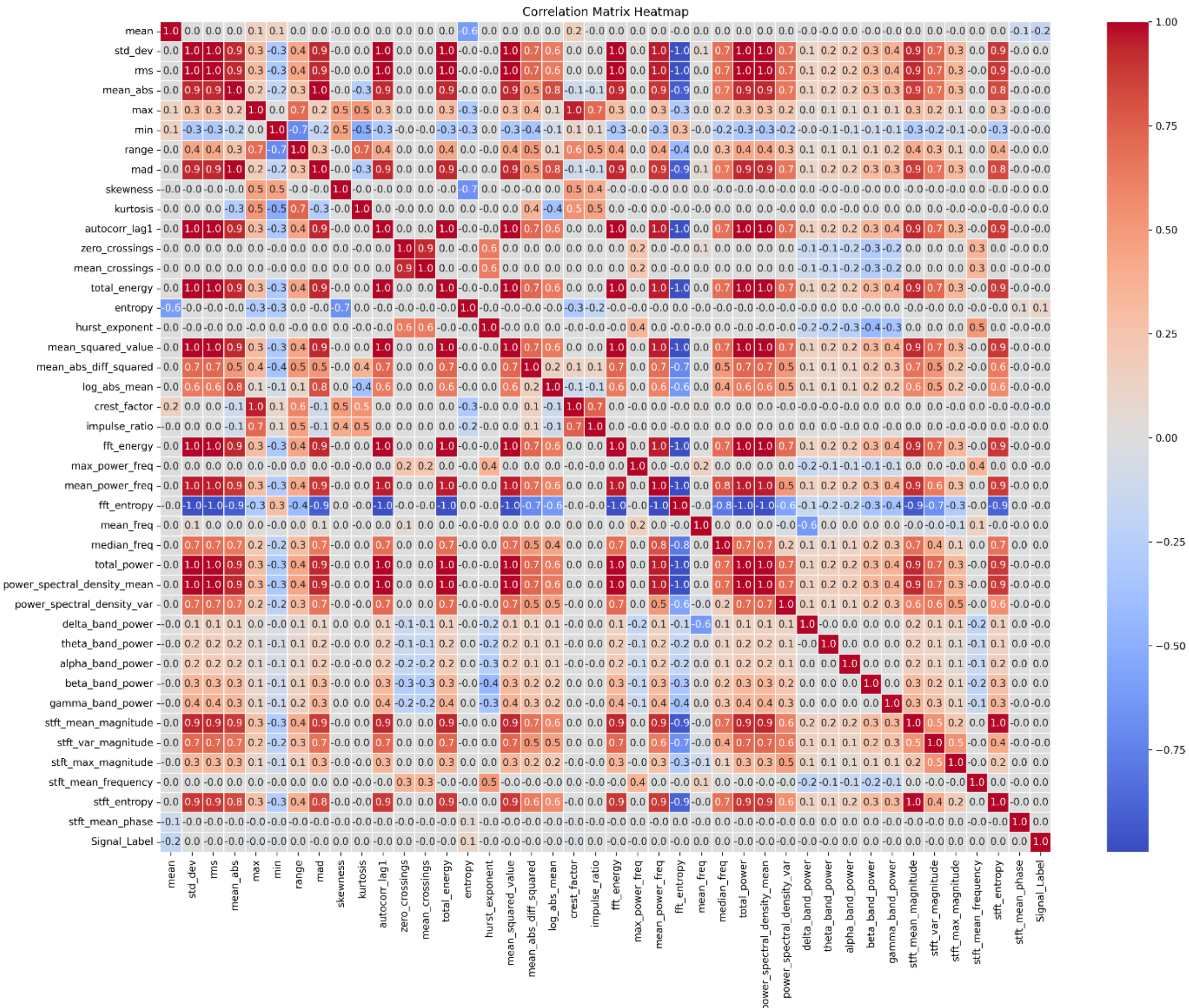


Figure 4.4: Heatmap correlation matrix provides the abundant context of linear relationships between various features extracted from signal data

The key features of correlation matrix show in figure 4 are as follows:

- Most of Amplitude Features show very good correlations amongst themselves (mean, max, min) thereby showing redundancy.
- Frequency-related metrics like `fft_entropy`, `fft_energy` are strongly and highly positively and negatively correlated among them.
- Total power and energy are very correlated, but their relation to `Signal_Label` is different, making them very important feature for your model.
- Redundant Feature: I drop some of the features that are highly correlated with each other, like `mean_abs_0` and `mean`.

- Target Relation: `fft_entropy`, `gamma_band_power`, `mean_power_freq` features are meaningful in predicting the signal labels since they relate with `Signal_Label`.

4.3 Recursive Feature Elimination

The importance of features extracted from the EMG signal dataset was ranked using Recursive Feature Elimination by applying RFE with a Random Forest Classifier. This eliminated features from being considered less significant in identifying those that highly contribute toward the model's predictive accuracy. The top 10 most relevant features that the RFE process derived are: stft_phase_var, mean_abs_diff, stft_mean_phase, skewness, stft_max, mean_abs_diff_squared, log_abs_mean, median_freq, stft_mean_freq, and min. These were the features that bore the best relevance to the model's performance. For instance, stft_phase_var was found at 0.0352 while mean_abs_diff and stft_mean_phase came out at 0.0352 and 0.0351, respectively. The dominated features that have a crucial role in model performance are stft_phase_var and other features. When iteratively removing the least important features, the accuracy remained stable until around 60% removal of the least important features. Thereafter, a minimal increase in accuracy was observed, hence suggesting a retention of a core set of the top stft domain features as the optimal for this classification problem. The RFECV process determines the top 14 features ensuring consistent performance for accuracy, as shown in figure 4.5.

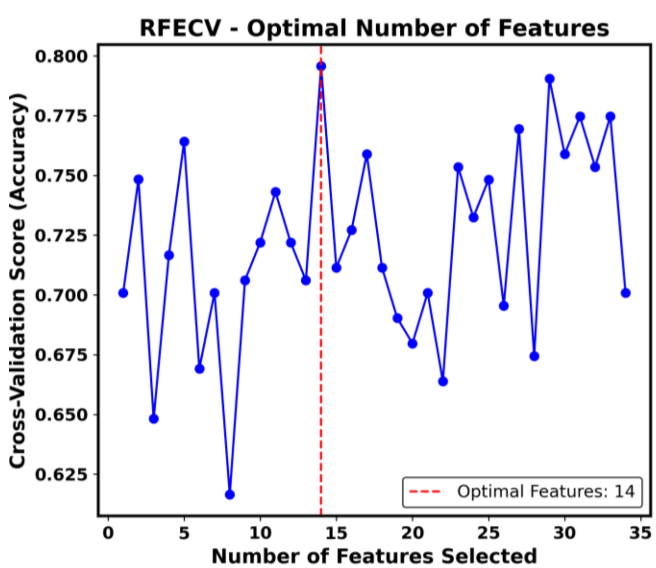


Figure 4.5: Optimal Feature selection using Recursive Feature Elimination with cross validation

The features selected by RFECV are according to their importance for the classification accuracy shown in figure 4.6.

Figure 4.6: Importance of all features based on random forest mode

4.4 Deep Learning Framework

To confirm the feasibility of the proposed alphabet recognition framework using EMG signals, 5-fold cross-validation experiments have been carried out. A CNN classifier has applied a model for classification of upper-case letter classes from preprocessed EMG data for 26 letters: A-Z. Results of accuracy for each fold and average accuracy are calculated for this experiment. The training and testing of the CNN model for alphabet recognition from surface electromyography signals were done with 5-fold cross-validation to avoid any biasing arising from a single train-test split and ensure a more comprehensive assessment of the generalization capacity of the model. For each fold, the dataset was divided into 80% of training data and 20% of testing data, with five iterations in total, so that every sample appears exactly once in the testing set.

It composed two convolutional layers and max-pooling layers. In brief, CNN successfully extracted the spatial and temporal features of various letters' sEMG signals. The model input consisted of a time-frequency representation of sEMG signals obtained from STFT. Each input sample had the shape (79, 101, 5), representing 79 windows, 101 frequency components, and 5 channels of EMG data.

4.4.1 *Accuracy of Each Fold*

The accuracy values of the model for all the 5 folds are as depicted in Figure 4.7. The model operated at accuracies of 88% to 90% across all folds with no fluctuation, therefore the performance is consistent with very negligible variance. The small inconsistencies observed between the folds can be result of the natural variability of the sEMG signal such as muscle activation, electrode placement, or the level of noise in the signal. In this experiment, the overall average accuracy for the entire five folds was calculated as an average, the alphabet letters from the sEMG signals were classified correctly by the model with an 89.65% accuracy, which implies good generalization capability towards unseen data. The consistent performance across the folds points out to the fact that the model effectively learned the patterns representing movements of the different hands associated with the letters of the alphabet.

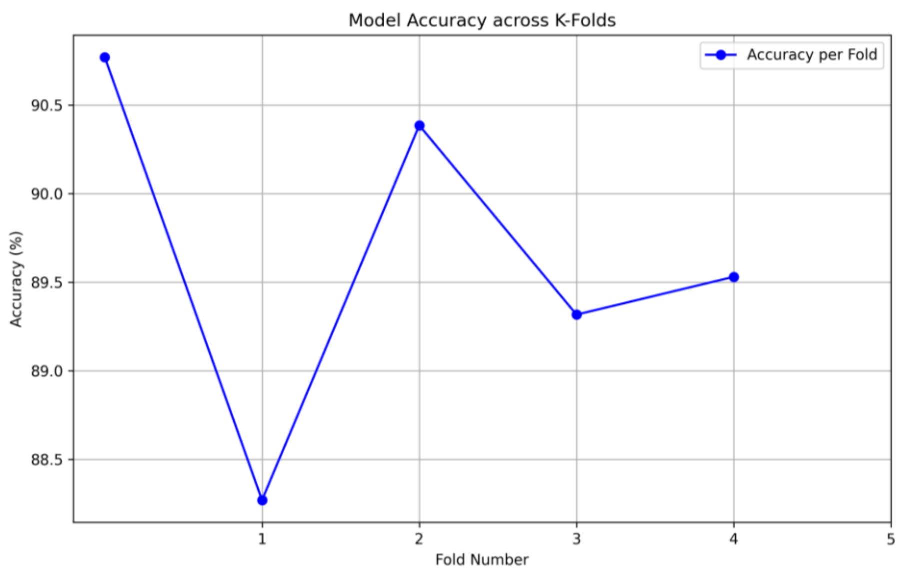


Figure 4.7: Accuracy of 1DCNN model for all the 5 folds

4.4.2 *Model Training and Early Stopping*

The model was trained using a learning rate of 0.001 for an Adam optimizer and the loss function was sparse categorical cross entropy due to the multi-class classification problem with 26 classes: letters A-Z. The early stopping of the training with validation accuracy stopped overfitting by halting training after the validation performance leveled off. This helped prevent overfitting the training data; however, the model achieved a high accuracy over the test data. The high accuracy of the model suggests that the CNN architecture is capturing very well the finest features in the EMG data, those associated with the fine motor movements assigned to each letter of the alphabet. Thus, the time-frequency features derived through the STFT and the spatial information through the different EMG channels have yielded a rich representation for the model to learn about.

4.5 Real Time Alphabet Prediction

A real-time prediction system of the English alphabet using sEMG signals collected from forearm muscles was designed on tkinter. The architecture of the prediction system is based on a deep learning model that specifically classifies sEMG signals into 26 categories as used in the alphabets from A-Z. The proposed system was tested with participants writing selected alphabets using specific forearm muscle movements while the signals were continuously recorded and processed in real time. The latency between the signal collection and display of the predicted character was a major performance metric. The average observed latency during testing was about 60 milliseconds, which is well within an acceptable range for real-time applications.

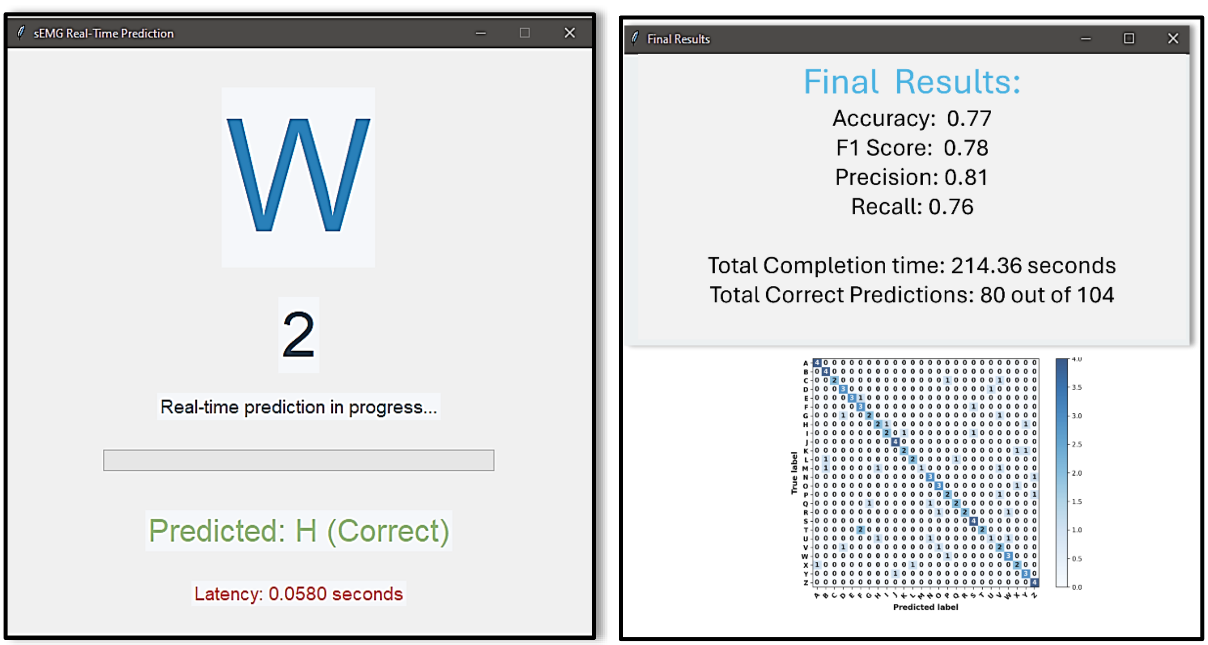
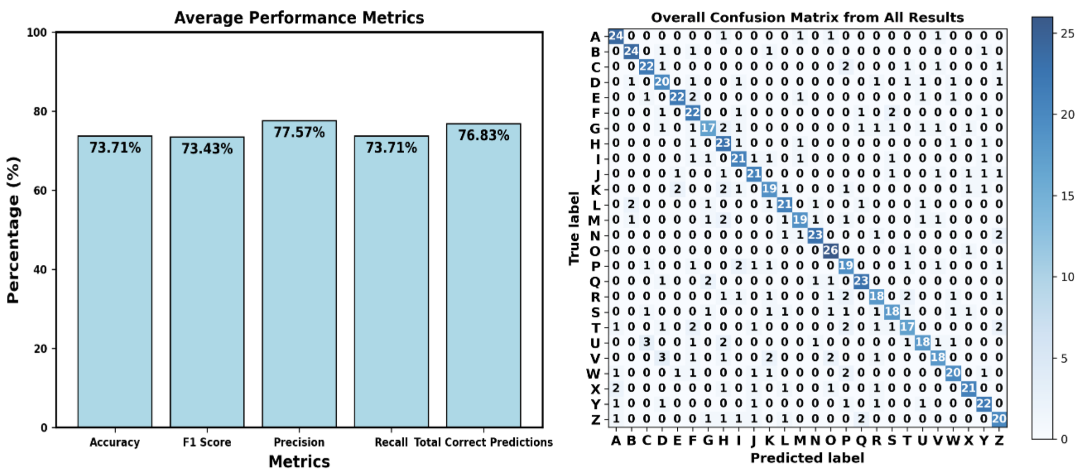


Figure 4.8: Tkinter protocol for real time alphabet prediction

The performance of the model was evaluated on grouped data, this analysis examines how effectively the model generalizes in real-time when tested on subjects whose data was part of the training set. Performance metrics, such as accuracy, precision, F1 Score, and recall, were also calculated. This model showed an average accuracy of 73.71% across all tested subjects with an excellent generalization toward patterns that it has seen earlier, even under real-time conditions as shown in figure 4.9 a. A detailed summary of model performance in recognizing 26 capital English letters using sEMG data from airwriting movements of 7 subjects can be examined in the confusion matrix in Figure 4.9 b. Many letters have good classification accuracy, as indicated by the high values along the diagonal. Letters like "A," "B," and "O" have significantly high correct predictions.



a b

Figure 4.9: (a) Average performance metrics of real time alphabet prediction on grouped data (b) Confusion matrix representing performance of model for recognizing each of the 26 uppercase English alphabets based on surface electromyography data. Each row corresponds to the true labels (actual classes) and each column corresponds to the predicted labels (the output from the classifier). Diagonal cells are the correct classifications while the off-diagonal cells represent misclassifications.

The average latency for each alphabet is shown in figure 4.10; it ranges from roughly 0.058 to 0.066 seconds. Some letters, such as E, M and T, have noticeably higher latency, because their shapes are less distinct or complex when written in the air. On the other hand, letters like I and L have lower latency, which means the algorithm can identify them more quickly. Figure 4.11 represents the average error rate for every letter, describing accuracy challenges in air-writing recognition. Error rates range from 0.15 to 0.37, and letters like G, T, U and V have higher error rates due to similarities in their patterns of air writing with other alphabets. In fact, lower error rates for letters like A, B, I, and O also show that these are easier and more likely to be recognized as true instances. Letters that are notable with a high error rating need to be identified and corrected for better accuracy in the airwriting recognition system.

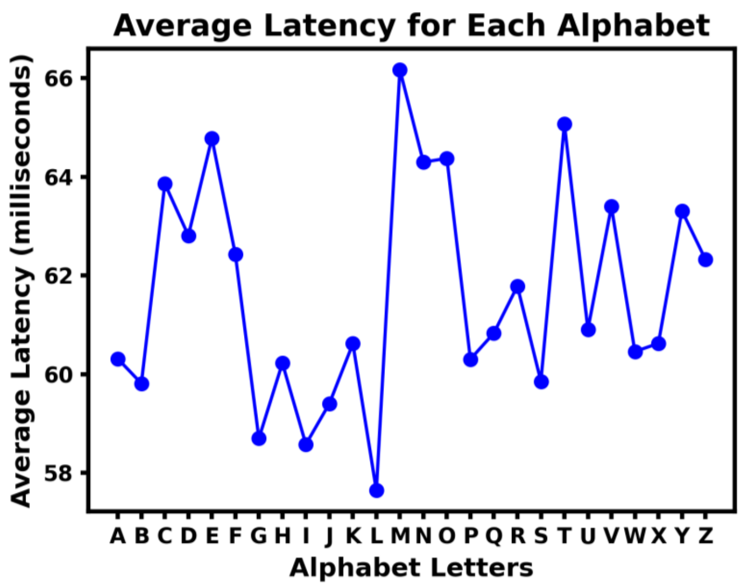


Figure 4.10: Average latency in seconds for the classification

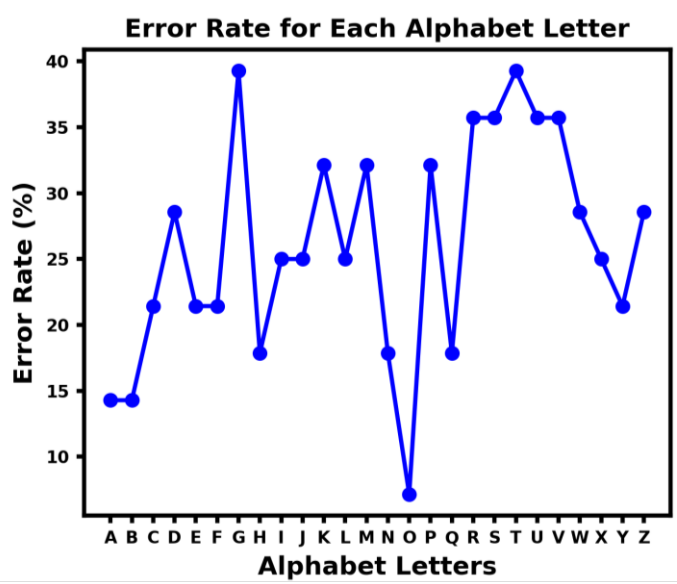


Figure 4.11: Error rate for each alphabet, obtained as the number of misclassified instances of a letter divided by the total number of instances of that letter.

Real-Time Prediction was also performed on Individual Data in which model performance was evaluated on real-time data specific to a single subject, testing adaptability to individual characteristics. During the individual-specific testing phase, the accuracy of model using EMG data of a single subject was 77%. This degree of accuracy suggests that, even in real-time situations, the model was able to adjust to distinct, subject-specific muscle patterns. Classification accuracy of model for real time alphabet prediction on individual data for each alphabet letter is shown in Figure 6. Significant differences between letters are highlighted by the plot, suggesting that some letters like L, I, T, and R, are predicted more accurately than others.

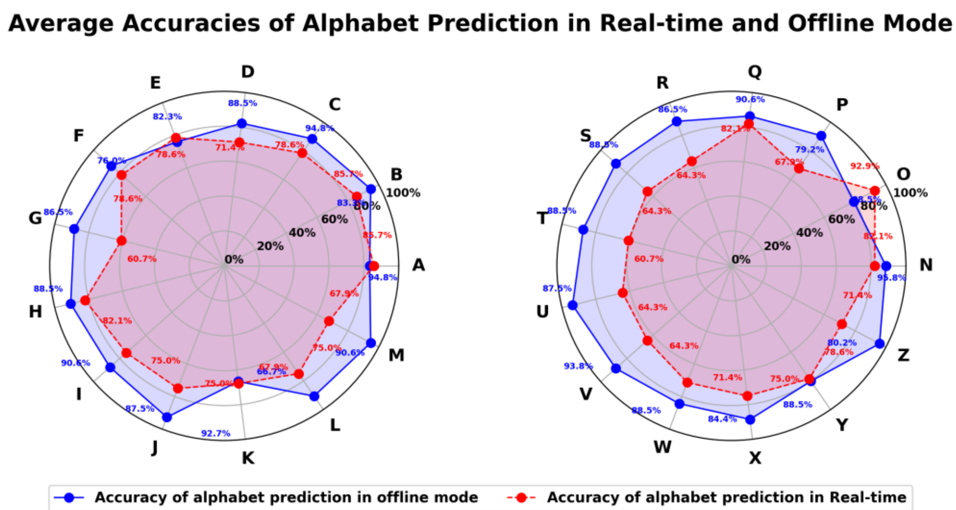


Figure 4.12: Radar plot of the Accuracy per Alphabet Letter for three different letter sets representing accuracy at which each letter was classified correctly, with values closer to the edge being more accurate.

Thus, preprocessing, feature selection, and model architecture play a crucial role in strong performance of model in classifying air-written alphabet letters. The best performing model was the 1DCNN model, which demonstrated strong generalization with an average real-time accuracy of 73.71% across subjects and an accuracy of 89% on training data. Individualized models for specific subjects performed even better, with an accuracy of 77% and an accuracy of 92% on training data. Letters like A, B, I, and O were classified with high accuracy, while letters such as G, T, U, and V were misclassified due to their muscle pattern similarities.

CHAPTER 5: SUMMARY OF RESEARCH WORK

The analysis of the processed data showed that the preprocessing techniques significantly improved the quality and consistency of sEMG signals across trials and subjects. SNR comparison for various wavelet denoising techniques show improved SNR using db9 wavelet. Various feature reduction techniques such as PCA, LDA and RFECV were employed to determine optimal features. Among them, RFECV provides the best result by maintaining high classification accuracy. This technique indicates the most relevant features while preventing overfitting by using cross-validation tests at every step of feature elimination. Unlike PCA, which focuses on variance, or LDA, which focuses on linear separability, RFECV iteratively removes the least impactful features that improve model performance, hence resulting in a very optimized set of 14 features. This efficiency improvement is harder to achieve by PCA and LDA because they tend to retain more generalized features, which could bring noise or irrelevant information into the model. Among various deep learning algorithms, 1DCNN achieves high classification accuracy in both offline and real-time alphabet prediction as shown in table 1. While LSTM-based DNN and the Fully Connected RNN had considerably lower accuracies at 71% and 62% respectively. The best performing 1DCNN model was trained and validated to deploy in real time alphabet predictions.

Table 5.1: Performance metrics for the deep learning model across different testing conditions

Performance Metrics	Offline Grouped Data	Offline Individual data	Real-time Grouped Data	Offline Grouped Data
Accuracy	89.81%	92%	73.71%	77%
Precision	91.39%	90%	77.57%	81%
Recall	89.81%	94%	73.71%	76%
F1 Score	90.06%	92%	73.43%	78%

The results highlight that both offline grouped and offline individual settings yield good results for the model. Performance measurements for real-time implementations had maintained consistency in precision, recall, F1 score, latency and error rate with means of 79.28 ± 1.71 %, 74.85 ± 1.14 %, and 75.72% ± 2.28%, 0.06 ± 0.01 seconds, 0.22 ± 0.07% respectively, across a range of testing situations. Although the accuracy is higher on offline individual data because the model had adjusted its recognition to unique signal patterns of each user, lessening the effect of inter-subject variability. Grouped data, on the other hand, adds greater unpredictability because individual signal patterns vary, which significantly affects performance. It is noted that the system shows promising results as most of the alphabets like A, B, and O have predicted correctly as shown in figure 6b but there are also significant misclassifications between letters such as U was confused with C, E with F, and Q with G because of their similar shapes, different writing styles of subjects and overlapping sEMG signal patterns. All the alphabets take different time from receiving the EMG signal until the result of the prediction is displayed that is calculated as latency. Some letters, such as E, M and T, have noticeably higher latency, because their shapes are less distinct or complex when written in the air. On the other hand, letters like I and L have lower latency, which means the algorithm can identify them more quickly. The figure demonstrates error rate per alphabet that describes the number of misclassified instances of a letter divided by the total number of instances of that letter, range from 15 to 27%, and letters like G, T, U and V have higher error rates due to similarities in their patterns of air writing with other alphabets. In fact, lower error rates for letters like A, B, I, and O also show that these are easier and more likely to be recognized as true instances. Letters that are notable with a high error rate need to be identified and corrected for better accuracy in the airwriting recognition system. However, the improved accuracy of model in individual-specific testing points to be a promising direction for personalized air-writing systems that adapt to the unique characteristics of each user, enhancing their potential for real-world applications. Thus, the proposed deep learning-based solution outperforms other existing research and airwriting recognition systems in terms of accuracy and real-time applicability. Due to its potential for real-time predictions and its simplicity of incorporation into assistive technology for people with speech impairments, this study promises greater real-world application. It has the potential to be used in assistive technology, rehabilitation, and human-computer interaction with further developments in feature extraction and model optimization.

CHAPTER 6: CONCLUSION AND FUTURE RECOMMENDATIONS

In this paper, an effective framework is proposed for the recognition of 26 uppercase English alphabet letters using sEMG-based airwriting by analyzing unique muscle activation pattern. The quality of signal was enhanced by using wavelet decomposition for denoising, thus improving the strength of the model in identification of muscle activation patterns related to each letter. The feature set was optimized using the RFECV technique, which reduced computational complexity while maintaining high classification accuracy.

Among the tested model, the best performing model was the 1DCNN, with an accuracy of 89% on training data and an average real-time accuracy of 73.71% across subjects. Individualized models for specific subjects performed even better, with an accuracy of 77% in real-time and an accuracy of 92% on training data. This system represents a significant improvement over earlier sEMG-based airwriting models, which were less accurate and lacked real-time capabilities. By achieving high accuracy in both training and real-time settings, this model addresses precision and adaptability by presenting a more realistic solution for real-world airwriting applications. This study highlights the potential of deep learning models for real-time, user-specific airwriting detection and provides foundations for sEMG-based airwriting applications. This method has the potential to be used in assistive technology, rehabilitation, and human-computer interaction with further developments in feature extraction and model optimization.

Future research could focus on employing hybrid deep learning models or advanced feature extraction techniques to improve model accuracy, especially for letters with comparable muscle patterns. For real-world applications, going beyond single-letter recognition to word-level recognition is also important requiring a system that can identify continuous letter sequences and integrate language models to enhance accuracy.

REFERENCES

[1] A. Phinyomark, E. Campbell, and E. Scheme, "Surface Electromyography (EMG) Signal Processing, Classification, and Practical Considerations," in *Biomedical Signal Processing: Advances in Theory, Algorithms and Applications*, G. Naik, Ed., Singapore: Springer, 2020, pp. 3–29. doi: 10.1007/978-981-13-9097-5_1.

[2] Y. Sun *et al.*, "Intelligent human computer interaction based on non redundant EMG signal," *Alexandria Engineering Journal*, vol. 59, no. 3, pp. 1149–1157, Jun. 2020, doi: 10.1016/j.aej.2020.01.015.

[3] J. Park *et al.*, "Assessing workload in using electromyography (EMG)-based prostheses," *Ergonomics*, vol. 67, no. 2, pp. 257–273, Feb. 2024, doi: 10.1080/00140139.2023.2221413.

[4] E. R S *et al.*, "A SURVEY ON AIR WRITING CHARACTER RECOGNITION AND TRANSLATION," *IJEAST*, vol. 7, no. 12, pp. 36–46, Apr. 2023, doi: 10.33564/IJEAST.2023.v07i12.006.

[5] A. Choudhury and K. K. Sarma, "Visual Gesture-Based Character Recognition Systems for Design of Assistive Technologies for People With Special Necessities," in *Research Anthology on Physical and Intellectual Disabilities in an Inclusive Society*, IGI Global, 2022, pp. 264–285. doi: 10.4018/978-1-6684-3542-7.ch014.

[6] C. Amma, M. Georgi, and T. Schultz, "Airwriting: A wearable handwriting recognition system," *Personal and Ubiquitous Computing*, vol. 18, Jan. 2014, doi: 10.1007/s00779-013-0637-3.

[7] F. Duan, X. Ren, and Y. Yang, "A Gesture Recognition System Based on Time Domain Features and Linear Discriminant Analysis," *IEEE Transactions on Cognitive and Developmental Systems*, vol. PP, pp. 1–1, Dec. 2018, doi: 10.1109/TCDS.2018.2884942.

[8] S. Mohammadi and R. Maleki, "Air-writing recognition system for Persian numbers with a novel classifier," *Vis Comput*, vol. 36, no. 5, pp. 1001–1015, May 2020, doi: 10.1007/s00371-019-01717-3.

[9] "A Real-Time Classification Model for Bengali Character Recognition in Air-Writing," in *Computer Vision and Image Analysis for Industry 4.0*, Chapman and Hall/CRC, 2023, pp. 109–119. doi: 10.1201/9781003256106-10.

[10] J. G. Beltrán Hernández, J. Ruiz Pinales, P. López Rodríguez, J. L. López Ramírez, and J. G. Aviña Cervantes, "Multi-Stroke handwriting character recognition based on sEMG using convolutional-recurrent neural networks," Aug. 2020, doi: 10.3934/mbe.2020293.

[11] A. Tripathi, A. K. Mondal, L. Kumar, and P. AP, "SCLAiR: Supervised Contrastive Learning for User and Device Independent Airwriting Recognition," *IEEE Sensors Letters*, vol. 6, no. 2, pp. 1–4, Feb. 2022, doi: 10.1109/LSENS.2021.3139473.

[12] A. K. Singh and D. Koundal, "A Temporal Convolutional Network for modeling raw 3D sequences and air-writing recognition," *Decision Analytics Journal*, vol. 10, p. 100373, Mar. 2024, doi: 10.1016/j.dajour.2023.100373.

[13] Z. Yang *et al.*, "Dynamic Gesture Recognition Using Surface EMG Signals Based on Multi-Stream Residual Network," *Frontiers in Bioengineering and Biotechnology*, vol. 9, 2021, doi: 10.3389/fbioe.2021.779353.

[14] A. Tripathi, A. P. Prathosh, S. P. Muthukrishnan, and L. Kumar, "SurfMyoAiR: A Surface Electromyography-Based Framework for Airwriting Recognition," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023, doi: 10.1109/TIM.2023.3248084.

[15] M. A. Ozdemir, D. H. Kisa, O. Guren, and A. Akan, "Hand gesture classification using time–frequency images and transfer learning based on CNN," *Biomedical Signal Processing and Control*, vol. 77, p. 103787, Aug. 2022, doi: 10.1016/j.bspc.2022.103787.

[16] A. Tripathi, A. K. Mondal, L. Kumar, and P. AP, "SCLAiR: Supervised Contrastive Learning for User and Device Independent Airwriting Recognition," *IEEE Sensors Letters*, vol. 6, no. 2, pp. 1–4, Feb. 2022, doi: 10.1109/LSENS.2021.3139473.

[17] A. Tripathi, A. K. Mondal, L. Kumar, and A. P. Prathosh, "ImAiR: Airwriting Recognition Framework Using Image Representation of IMU Signals," *IEEE Sensors Letters*, vol. 6, no. 10, pp. 1–4, Oct. 2022, doi: 10.1109/LSENS.2022.3206307.

[18] F. Al Abir, M. Siam, A. Sayeed, Md. A. Hasan, and J. Shin, "Deep Learning Based Air-Writing Recognition with the Choice of Proper Interpolation Technique," *Sensors*, vol. 21, p. 8407, Dec. 2021, doi: 10.3390/s21248407.

[19] J. G. Beltrán Hernández, J. Ruiz Pinales, P. López Rodríguez, J. L. López Ramírez, and J. G. Aviña Cervantes, "Multi-Stroke handwriting character recognition based on sEMG using convolutional-recurrent neural networks," Aug. 2020, doi: 10.3934/mbe.2020293.

[20] M. I. Rusydi, Oktrison, W. Azhar, S. W. Oluwarotimi, and F. Rusydi, "Towards hand gesture-based control of virtual keyboards for effective communication," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 602, no. 1, p. 012030, Aug. 2019, doi: 10.1088/1757-899X/602/1/012030.

[21] S. Sharma and R. Gupta, "On the Use of Temporal and Spectral Central Moments of Forearm Surface EMG for Finger Gesture Classification," in *2018 2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)*, Sep. 2018, pp. 234–239. doi: 10.1109/ICMETE.2018.00059.

[22] A. Tripathi, P. AP, S. P. Muthukrishnan, and L. Kumar, "TripCEAiR: A Multi-Loss minimization approach for surface EMG based Airwriting Recognition," Mar. 19, 2023, *arXiv*: arXiv:2212.02870. doi: 10.48550/arXiv.2212.02870.

[23] S. Rampichini, T. M. Vieira, P. Castiglioni, and G. Merati, "Complexity Analysis of Surface Electromyography for Assessing the Myoelectric Manifestation of Muscle Fatigue: A Review," *Entropy*, vol. 22, no. 5, May 2020, doi: 10.3390/e22050529.

[24] M. Al-Ayyad, H. A. Owida, R. De Fazio, B. Al-Naami, and P. Visconti, "Electromyography Monitoring Systems in Rehabilitation: A Review of Clinical Applications, Wearable Devices and Signal Acquisition Methodologies," *Electronics*, vol. 12, no. 7, Art. no. 7, Jan. 2023, doi: 10.3390/electronics12071520.

[25] F. A. Abir, Md. A. Siam, A. Sayeed, Md. A. M. Hasan, and J. Shin, "Deep Learning Based Air-Writing Recognition with the Choice of Proper Interpolation Technique," *Sensors (Basel)*, vol. 21, no. 24, p. 8407, Dec. 2021, doi: 10.3390/s21248407.

[26] A. Qedear, A. AlMatrafy, A. Al-Sowat, A. Saigh, and A. Alayed, "Real-Time Air-Writing Recognition for Arabic Letters Using Deep Learning," *Sensors*, vol. 24, no. 18, Art. no. 18, Jan. 2024, doi: 10.3390/s24186098.

[27] Q. Li and R. Langari, "EMG-based HCI Using CNN-LSTM Neural Network for Dynamic Hand Gestures Recognition," *IFAC-PapersOnLine*, vol. 55, no. 37, pp. 426–431, Jan. 2022, doi: 10.1016/j.ifacol.2022.11.220.

[28] J. M. Vaz and S. Balaji, "Convolutional neural networks (CNNs): concepts and applications in pharmacogenomics," *Mol Divers*, vol. 25, no. 3, pp. 1569–1584, 2021, doi: 10.1007/s11030-021-10225-3.

[29] Z. Ding, C. Yang, Z. Tian, C. Yi, Y. Fu, and F. Jiang, "sEMG-Based Gesture Recognition with Convolution Neural Networks," *Sustainability*, vol. 10, no. 6, Art. no. 6, Jun. 2018, doi: 10.3390/su10061865.

[30] D. Boateng *et al.*, "Air-Writing Recognition Enabled by a Flexible Dual-Network Hydrogel-Based Sensor and Machine Learning," *ACS applied materials & interfaces*, Sep. 2024, doi: 10.1021/acsami.4c10168.

[31] Y. Xue, D. Zhang, L. Li, S. Li, and Y. Wang, "Lightweight multi-scale convolutional neural network for real time stereo matching," *Image and Vision Computing*, vol. 124, p. 104510, Aug. 2022, doi: 10.1016/j.imavis.2022.104510.

[32] C. Amma and T. Schultz, "Airwriting: bringing text entry to wearable computers," *XRDS*, vol. 20, no. 2, pp. 50–55, Dec. 2013, doi: 10.1145/2540048.

[33] M. Chen, G. AlRegib, and B.-H. Juang, "Air-Writing Recognition—Part I: Modeling and Recognition of Characters, Words, and Connecting Motions," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 3, pp. 403–413, Jun. 2016, doi: 10.1109/THMS.2015.2492598.

[34] T. Yanay and E. Shmueli, “Air-writing recognition using smart-bands,” *Pervasive and Mobile Computing*, vol. 66, p. 101183, Jul. 2020, doi: 10.1016/j.pmcj.2020.101183.

[35] U.-H. Kim, Y. Hwang, S.-K. Lee, and J.-H. Kim, “Writing in the Air: Unconstrained Text Recognition From Finger Movement Using Spatio-Temporal Convolution,” *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 6, pp. 1386–1398, Dec. 2023, doi: 10.1109/TAI.2022.3212981.

[36] S. Mukherjee, Sk. A. Ahmed, D. P. Dogra, S. Kar, and P. P. Roy, “Fingertip detection and tracking for recognition of air-writing in videos,” *Expert Systems with Applications*, vol. 136, pp. 217–229, Dec. 2019, doi: 10.1016/j.eswa.2019.06.034.

[37] D. Bachmann, F. Weichert, and G. Rinkenauer, “Review of Three-Dimensional Human-Computer Interaction with Focus on the Leap Motion Controller,” *Sensors*, vol. 18, no. 7, Art. no. 7, Jul. 2018, doi: 10.3390/s18072194.

[38] A. K. Jain, R. P. W. Duin, and J. Mao, “Statistical pattern recognition: a review,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, Jan. 2000, doi: 10.1109/34.824819.

[39] C. Xu *et al.*, “Multi-loss Regularized Deep Neural Network,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 12, pp. 2273–2283, Dec. 2016, doi: 10.1109/TCSVT.2015.2477937.

[40] X. Tao, J. Kong, M. Jiang, and T. Liu, “Unsupervised Domain Adaptation by Multi-Loss Gap Minimization Learning for Person Re-Identification,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 7, pp. 4404–4416, Jul. 2022, doi: 10.1109/TCSVT.2021.3135274.

[41] W. Zhong, L. Jiang, T. Zhang, J. Ji, and H. Xiong, “Combining multilevel feature extraction and multi-loss learning for person re-identification,” *Neurocomputing*, vol. 334, pp. 68–78, Mar. 2019, doi: 10.1016/j.neucom.2019.01.005.

[42] L. Guo, Z. Lu, and L. Yao, “Human-Machine Interaction Sensing Technology Based on Hand Gesture Recognition: A Review,” *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 4, pp. 300–309, Aug. 2021, doi: 10.1109/THMS.2021.3086003.

[43] S. Jiang, P. Kang, X. Song, B. P. L. Lo, and P. B. Shull, “Emerging Wearable Interfaces and Algorithms for Hand Gesture Recognition: A Survey,” *IEEE Reviews in Biomedical Engineering*, vol. 15, pp. 85–102, 2022, doi: 10.1109/RBME.2021.3078190.

[44] K. Sabry and S. AlShawi, “Information systems for higher education: interactive design perspective,” *Transforming Government: People, Process and Policy*, vol. 3, no. 2, pp. 163–180, Jan. 2009, doi: 10.1108/17506160910960559.

[45] U.-H. Kim, Y. Hwang, S.-K. Lee, and J.-H. Kim, “Writing in the Air: Unconstrained Text Recognition From Finger Movement Using Spatio-Temporal Convolution,” *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 6, pp. 1386–1398, Dec. 2023, doi: 10.1109/TAI.2022.3212981.

[46] W. Wei and L. Ren, “From Unimodal to Multimodal: improving sEMG-Based Pattern Recognition via deep generative models,” arXiv.org. Accessed: Sep. 28, 2024. [Online]. Available: https://arxiv.org/abs/2308.04091v2

[47] X. Gao, R. K. Saha, M. R. Prasad, and A. Roychoudhury, “Fuzz testing based data augmentation to improve robustness of deep neural networks,” in *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering*, in ICSE ’20. New York, NY, USA: Association for Computing Machinery, Oct. 2020, pp. 1147–1158. doi: 10.1145/3377811.3380415.

[48] S. K. Singh and A. Chaturvedi, “An efficient multi-modal sensors feature fusion approach for handwritten characters recognition using Shapley values and deep autoencoder,” *Engineering Applications of Artificial Intelligence*, vol. 138, p. 109225, Dec. 2024, doi: 10.1016/j.engappai.2024.109225.

[49] C. Ouyang, L. Cai, B. Liu, and T. Zhang, “An improved wavelet threshold denoising approach for surface electromyography signal,” *EURASIP Journal on Advances in Signal Processing*, vol. 2023, no. 1, p. 108, Oct. 2023, doi: 10.1186/s13634-023-01066-3.

[50] M. Srivastava, C. L. Anderson, and J. H. Freed, “A New Wavelet Denoising Method for Selecting Decomposition Levels and Noise Thresholds,” *IEEE Access*, vol. 4, pp. 3862–3877, 2016, doi: 10.1109/ACCESS.2016.2587581.