

TONGUE DRIVE SYSTEM ON FPGA PROCESSOR USING MACHINE LEARNING MODULE

Thesis Submitted for the Award of the Degree of

DOCTOR OF PHILOSOPHY

in

Electronics and Communication Engineering

by

Dasari Lakshmi Prasanna

Registration Number: 12014083

Supervised By

Suman Lata Tripathi (21067)

Electrical and Electronics Engineering

(Professor)

Lovely Professional University

Phagwara, Punjab, India

Co-Supervised by

Inung Wijayanto

Electrical Engineering

(Associate Professor)

Telkom University

Bandung, Indonesia



LOVELY PROFESSIONAL UNIVERSITY, PUNJAB

2026

DECLARATION

I, hereby declared that the presented work in the thesis entitled “**Tongue Drive System on FPGA Processor using Machine Learning Module**” in fulfilment of degree of **Doctor of Philosophy (Ph. D.)** is outcome of research work carried out by me under the supervision of Dr. Suman Lata Tripathi, working as Professor, in the School of Electrical and Electronics of Lovely Professional University, Punjab, India and Co-supervisor Dr. Inung Wijayanto, working as a professor, in the School of Electrical Engineering, Telkom University, Bandung, Indonesia. In keeping with general practice of reporting scientific observations, due acknowledgements have been made whenever work described here has been based on findings of other investigator. This work has not been submitted in part or full to any other University or Institute for the award of any degree.



(Signature of Scholar)

Dasari Lakshmi Prasanna

12014083

School of Electronics and Communication Engineering

Lovely Professional University

Punjab, India

CERTIFICATE

This is to certify that the work reported in the Ph. D. thesis entitled “**Tongue Drive System on FPGA Processor using Machine Learning Module**” submitted in fulfillment of the requirement for the award of degree of **Doctor of Philosophy (Ph.D.)** in the School of Electronics and Communication Engineering, is a research work carried out by Dasari Lakshmi Prasanna, 12014083, is bonafide record of her original work carried out under my supervision and that no part of thesis has been submitted for any other degree, diploma or equivalent course.



(Signature of Supervisor)

Suman Lata Tripathi

Professor

School of Electrical and Electronics Engineering

Lovely Professional University

Punjab, India



(Signature of Co-Supervisor)

Inung Wijayanto

Associate Professor

School of Electrical Engineering

Telkom University

Bandung, Indonesia

ABSTRACT

The World Health Organization (WHO) reports that over 2.5 billion people globally need assistive devices (AD). This figure is anticipated to increase to 3.5 billion by 2025. Increase in the traffic accidents, older people and people with certain medical conditions like paralysis, cerebral palsy, stroke, dementia or neuromuscular diseases often results in mobility difficulties. This has resulted in a significant increase in the demand for dependable and user-friendly assistive technology (AT). Conventional wheelchair such as joysticks, sip-and-puff systems, or eye-tracking gadgets is difficult to use for individuals with limited motor control, fatigue, or high cost of availability. Tongue Drive Systems (TDS) using magnetic sensors are introduced to avoid drawbacks of conventional AT. Traditional TDS also faces challenges like oral placement presents hygiene issues, prolonged usage makes user uncomfortable, and the precision of magnetic sensing may be compromised by sensor alignment and tongue movement, and hardware is also costly and complex. Speech is one of the most natural and easy for people to communicate, and it typically least affected even a serious accident. This makes it highly suitable and effective way for speech-based TDS to control assistive wheelchairs. This is easy to use and no need to wear heavy gadgets, user friendly, and cost effective.

The primary objective of this research is to design, develop, and implement a speech-based TDS on Field Programmable Gate Arrays (FPGA) processor with low power consumption and high accuracy speech recognition even in noisy conditions. Speech based assistive technology (AT) enables hands-free mobility by using a FPGA processor for real-time speech recognition. This research focuses on the implementation of an audio classification system using low power consumption, even in loud environments. Real-time, low-latency operation is the goal of this research, which combines FPGA-based hardware acceleration, sophisticated signal processing, and optimized machine learning models.

The first phase of the research focuses on improving machine learning models for classifying audio commands. Different machine learning classifiers studied and implemented for speech recognition. Naïve Bayes (NB), Stochastic Gradient Descent (SGD), Support Vector Machines (SVM), and Random Forest (RF) classifiers are

designed and compared the performance metrics for speech recognition. EMI filters are designed to suppress and eliminate the noise from audio signals. Active and passive filters are designed and calculated the gain, return loss and insertion loss. Collected 1400 audio samples of eight command words from 175 members for activating wheelchair. Command words – “back”, “call”, “down”, “help”, “left”, “right”, “straight” and “up” divided as eight classes. The dataset is collected in all kind of environment. Some dataset is collected in noise environment and other is in closed environment. The audio samples are balanced, labeled and organized according to command. The dataset was pre-processed using AUDACITY software by applying tempo, normalization, trimming and resizing to improve the quality of dataset for classification. Extracted the feature extraction 13 mel-frequency cepstral coefficients (MFCC) and energy for the speech classification. Multi-class classifier is designed calculated the performance metrics in MATLAB. One-vs-All SVM classifier are trained and tested with 80% and 20% dataset. The tested 20% data is of different people to check user independence.

The real-time speech-based TDS is accessed using Artix7 FPGA board. It has an inbuild ADMP421 MEMS microphone. Which collects speech signals in Pulse Density Modulation (PDM) format to guarantee high-fidelity speech acquisition. The PDM signal is converted to Pulse Code Modulation (PCM) using Cascaded Integrator-Comb (CIC). The data size is reduced while the signal quality is maintained, so that it will be used in noise environment. A passive low-pass EMI filter is employed post-CIC stage to eliminate undesired spectral components, reduce electromagnetic interference, and retain only the desired audio band.

The machine learning classifier is optimized for FPGA deployment and is made for real-time performance with low computational requirements. MATLAB HDL Coder is used for implementing SVM classifier on FPGA. The SVM model is mapped into a hardware-efficient implementation using fixed-point arithmetic, to reduce DSP block utilization over floating-point approaches. Feature selection reduces memory consumption and satisfying real-time requirements without sacrificing classification precision.

The final stage of this research involves implementing and testing the entire TDS system on the Xilinx Nexys 4 Artix-7 FPGA. Verilog HDL is used for entire

pipelines, which is then synthesized in Vivado and implemented on the FPGA. This pipeline includes audio acquisition, PDM-to-PCM conversion, EMI filtering, feature extraction, classification, and Pulse width Modulation (PWM) signal generation for motor control. For designing small prototype, we used H-bridge L298N motor driver which is connected to FPGA board and two DC motors of front wheels of wheelchair is connected to it. To supply the power for FPGA, motor driver and DC motor, a 12V power supply is connected.

Classifier- SGD, SVM, RF and NB are designed for speech detection. Trained on same dataset. The performance metrics of classifier SGD, SVM, RF and NB show the accuracy of 93%, 91%, 87% and 84% respectively. Precision, sensitivity, specificity and F1 score of SGD is 0.71, 0.69, 0.72, and 0.70, and SVM metrics values are 0.69, 0.78, 0.65 and 0.74, respectively. NB and RF values for precision 0.69 and 0.63, sensitivity 0.78 and 0.64, specificity 0.63 for both and F1 score is 0.73 and 0.63, respectively. These results show that SVM and SGD have better performance metrics. SVM has best recall means catches more positives and has the better F1 indicates better at balancing precision & recall for the positive class. SVM is better option to design TDS for more responsiveness means to not miss the user commands.

EMI filters are designed and compared the insertion loss, return loss and gain for both active and passive filters. Passive filters have better values of insertion loss of 1.9382 and gain is 0.912 than active filter. Multi-class one-vs-all SVM is designed for classifying command of eight words. Classifier is trained dataset which is collected in noisy and closed environments. SVM is trained on 1120 audio clips and tested on 280 samples. Calculated performance metrics and implemented trained SVM on FPGA.

The flow of FPGA implementation is audio acquisition using on board ADMP421 MEMS microphone, which give 1-bit PDM output. We designed CIC filter to convert the signal to PCM and reduce data size and maintain data quality. To suppress the unwanted spectral component, add low-pass passive EMI filter. Feature extraction block is added to extract 13 MFCC and energy features for classification and then PWM block is added for wheelchair control.

An efficient, scalable, and real-time Tongue Drive System for assistive mobility is developed by the integration of machine learning, noise suppression approaches, and FPGA-based hardware acceleration. A highly effective embedded voice processing

system is designed by integrating automatic HDL conversion of MATLAB HDL converter, FPGA implementation, and MATLAB-driven algorithm design. These findings advance low-power, FPGA-accelerated speech recognition, opening new avenues for assistive devices and autonomous mobility solutions.

ACKNOWLEDGEMENT

I am deeply grateful to my advisor, Dr. Suman Lata Tripathi, for her unwavering support, encouragement, and invaluable guidance during my doctoral journey. Her mentorship has been invaluable in the development of this research, and I am profoundly appreciative of her patience, insightful feedback, and expertise. Additionally, I would like to extend my appreciation to the School of Electronics and Communication Engineering at Lovely Professional University for its provision of a supportive research environment, continuous academic support, and access to essential resources. I am equally grateful to the administrative staff and faculty members for their support and motivation throughout my doctoral studies.

I wholeheartedly thank my co-supervisor, Dr. Inung Wijayanto, for his invaluable insights, technical expertise, and unwavering encouragement throughout this research. His constructive feedback and thoughtful suggestions have significantly enriched my work, and I am sincerely grateful for his support. I express my gratitude to Telkom University, Bandung, Indonesia, for granting me the chance to engage in the Research Exchange Program. This experience enabled collaboration with distinguished faculties, access to advanced resources, and international exposure, which greatly enhanced the quality and impact of the research.

I would want to express my sincere gratitude to my lab colleagues and research collaborators for their insightful technical talks, recommendations, and conversations that have substantially enhanced our work.

I am eternally grateful to my family and friends for serving as my steadfast support system. Your support, affection, and forbearance have inspired me to surmount obstacles. I am extremely appreciative of my parents' unwavering faith in me, as well as my siblings and friends' unwavering support and presence.

Lastly, I dedicate this thesis to all individuals with motor impairments, with the hope that advancements in assistive technologies will enhance their quality of life and independence.

TABLE OF CONTENTS

Declaration	i
Certificate	ii
Abstract	iii
Acknowledgement	vii
Table of Contents	viii
List of Figures	xiii
List of Tables	xviii
Abbreviation	xx
Chapter 1: Introduction	1-10
1.1 Background of Assistive Technology	1
1.2 Evaluation of Tongue Drive Systems	3
1.3 Development of AI and ML for Healthcare Applications	4
1.4 Advancement in Machine Learning	4
1.5 Machine Learning in VLSI	6
1.6 Research Gap	7
1.7 Problem Statement	8
1.8 Objectives	8
1.9 Role of International Collaboration in supporting Objective	8
1.10 Thesis Organization	9
Chapter 2: Literature Review	11-21

Chapter 3: Methodology	22-51
3.1 Sensor-based Tongue Drive Systems	23
3.1.1 Magnetic TDS	24
3.1.2 Capacitive TDS	26
3.1.3 Electromagnetic TDS	27
3.1.4 Optical TDS	28
3.1.5 Electromyographic TDS	29
3.1.6 Ultrasound based TDS	31
3.1.7 Speech based TDS	33
3.2 EMI filters	37
3.2.1 Passive EMI filters	37
3.2.2 Active EMI filters	38
3.2.3 Hybrid EMI filters	40
3.3 Feature Extraction	42
3.3.1 Time-domain Feature Extraction	42
3.3.2 Frequency-Domain Feature Extraction	44
3.3.3 Time-Frequency Feature Extraction	46
3.4 Artificial Intelligence and Machine Learning	48
3.4.1 Artificial Intelligence	48
3.4.2 Machine Learning	50
Chapter 4: Design and Implementation of Tongue Drive System on FPGA Processor	52-67
4.1 Implementation of speech-based TDS	53
4.2 Audio Acquisition and preprocessing	54

4.2.1	ADMP421 microphone	55
4.2.2	PDM to PCM Conversion using CIC	55
4.2.3	Signal Framing and Windowing	56
4.2.4	DC offset removal and amplitude normalization	56
4.3	EMI Filters	56
4.3.1	Passive EMI Filter	57
4.3.2	Active EMI Filter	58
4.4	Machine Learning Classifier	59
4.4.1	Data Collection	59
4.4.2	Dataset Organization and Labeling	60
4.4.3	Data Preprocessing using AUDACITY Software	60
4.4.4	Feature Extraction for Machine Learning Classifier	61
4.4.5	Implementation of Machine Learning Classifier	62
4.4.6	Multi-Class SVM Classifier	63
4.5	PWM generator	64
Chapter 5: International Research Collaborative Case Studies in Biomedical Signal and Image Processing		68-91
5.1.	Implementation of an Epileptic seizure detection	68
5.1.1	Dataset Description	70
5.1.2	Compressing EEG signal	70
5.1.3	Feature Extraction	70
5.1.4	SVM classifier for Epileptic seizure detection	71

5.1.5	MATLAB HDL Coder	73
5.1.6	Results and Discussion of Implementation of an Epileptic seizure detection	74
5.1.7	Conclusion of Epileptic Seizure Detection Implementation	78
5.2.	Classifying the Tongue Contours for Assistive Devices	79
5.2.1	Tongue Contours Classification	80
5.2.2	Pre-processing	80
5.2.3	Geometric Analysis	83
5.2.4	Convolution Neural Network (CNN)	83
5.2.5	Performance Metrics	84
5.2.6	Results of Tongue Contours Classification for Assistive Devices	87
5.2.7	Conclusion of Classification of Tongue Contours for Assistive Device	90
Chapter 6: Results and Discussion		92-122
6.1	Machine Learning classifiers	92
6.2	Active and Passive EMI filters	96
6.3	SVM 8x8 multi-class classifier	110
6.4	VIVADO results	115
6.4.1	Audio Acquisition	115
6.4.2	Passive Low-Pass EMI filter	117
6.4.3	Feature Extraction	118
6.4.4	Multiclass one vs all SVM Classifier	120
6.4.5	PWM generation	121

Chapter 7: Conclusion and Future Scope	123-125
References	126
APPENDIX-A	152
APPENDIX-B	154
Certification	156

LIST OF FIGURES

Figure No.	Figure Title	Page No.
2.1	Magnetic placement	14
2.2	Block diagram of TDS Transmitter	15
2.3	Block diagram of TDS Receiver	16
2.4	Working model of generating electrical signals to control the device	17
2.5	EM sensor placement for data collection	20
3.1	Flowchart of speech- based TDS implementation on FPGA	22
3.2	Magnetic sensor-based TDS	25
3.3	Optical sensors for TDS	29
3.4 (a)	Photo sensor	30
3.4 (b)	EMG placements for electromyographic TDS	31
3.5	Ultrasound transducer placement for collecting ultrasound images	33
3.6	Basic Passive low-pass filters	38
3.7	Basic passive high-pass filter	39
3.8	Basic passive band-pass filter	39

3.9	Basic active low-pass filter	40
3.10	Basic active band-pass filter	40
3.11	Hierarchy of Artificial Intelligence	49
3.12	Machine Learning Types and algorithms	51
4.1	Block diagram of speech-based TDS design	53
4.2	MEMS omnidirectional ADMP421 Microphone	54
4.3	Single stage CIC filter	55
4.4	Passive EMI filter	57
4.5	Active EMI filter	58
4.6	Wire connections of FPGA and L298N motor driver to two DC motors of wheelchair	66
5.1	Block diagram of epileptic seizure detection	70
5.2	Steps involved in HDL Workflow Advisor	73
5.3	Workflow advisor of MATLAB HDL Coder	74
5.4	Classification of tongue contours using hybrid approaches	80
5.5	Pre-processing blocks for ultrasound video of tongue contours	80
5.6	Block diagram of classifying the tongue contours	83

5.7	Classification of tongue contours using Convolution Neural Network (CNN)	84
5.8	Frame extraction of tongue contours from ultrasound video	86
5.9	Images of input and output ultrasound of tongue contours	87
5.10	Command window of geometric feature extraction	88
5.11	Graph of tongue area and perimeter over time	89
6.1	Raw audio signal	92
6.2	Bar chart of comparison of Performance Metrics of MLCs	94
6.3	ROC diagram of SGD	94
6.4	ROC diagram of NB	95
6.5	ROC diagram of SVM	96
6.6	ROC diagram of RF	96
6.7	Stage-I Passive EMI low-pass filter	97
6.8	Stage-II Passive EMI low-pass filter	98
6.9	Stage I passive EMI High-pass filter response	99
6.10	Stage-II Passive EMI filter response	99
6.11	Response of Active EMI high-pass filter	100

6.12	Response of Active EMI low-pass filter	101
6.13	The output response of active EMI filters time and Frequency domain	102
6.14	Output waveform of filter design	105
6.15	Power summary of filter design on FPGA	106
6.16	Device package	106
6.17	The schematic of filter design	107
6.18	The output image of FPGA – delayed_xout	107
6.19	The output image of yout	108
6.20	Confusion matrix for 20 testing sets for 800 audio clips	110
6.21	Runtime of each objective	111
6.22	Overall Accuracy and each class performance metrics	111
6.23	SVM confusion matrix of one vs all SVM multiclass classifier	112
6.24	Evaluation time of one vs all 8x8 multiclass classifier	113
6.25	Minimum objective vs Number of function Evaluation graph	114
6.26	ROC curve for 8x8 metaclassifier	115

6.27	Audio acquisition block with CIC filter and PDM-to-PCM Converter	116
6.28	Schematic PDM to PCM conversion	116
6.29	Audio acquisition power summary	117
6.30	EMI waveforms	117
6.31	Schematic of EMI filter	118
6.32	Power utilization summary for EMI filter	118
6.33	Waveforms of Feature extraction	119
6.34	Power utilization summary of feature extractions	120
6.35	Schematic of feature extraction	120
6.36	Waveforms generated for SVM classifier	121
6.37	Power summary for SVM classifier	121
6.38	Control signals are generated for L298N motor driver.	122

LIST OF TABLES

Table No.	Table Title	Page No.
3.1	Comparison of different sensor-based TDS	35
4.1	Command and their action	64
4.2	Control signal generation according to the commands	65
5.1	Explanation and Expressions of Performance metrics for SVM Classifier	72
5.2	Performance Metrics of dataset of Neurology and sleep center, New Delhi	75
5.3	Performance Metrics of dataset of Bonn University	75
5.4	Utilization summary of seizure detection modules for New Delhi dataset	76
5.5	Utilization summary of seizure detection modules for Bonn University	76
5.6	Comparison of utilization of FPGA implementation of seizure detection	77
5.7	Metrics of SRAD and NLM noise reduction	88
5.8	Performance metrics of classifier using CNN	90
6.1	Comparison of Performance metrics for different classifiers	93

6.2	Parameters comparison of Active and Passive EMI filters	103
6.3	Utilization summary of FPGA implementation	109
6.4	Slice and block counts	109

ABBREVIATIONS

AB	Able bodied
AD	Assistive Devices
AFE	Analog front-end
AI	Artificial Intelligence
ALS	Amyotrophic Lateral Sclerosis
AT	Assistive Technology
BCI	Brain-Computer Interface
CIC	Cascaded Integrator-Comb
CNN	Convolutional Neural Networks
CP	Cerebral Palsy
CMOS	Complementary Metal-Oxide-Semiconductor
CPU	Central processing unit
DSP	Digital Signal Processing
DT	Decision Tree
ECS	Environmental Control Systems
EEG	Electroencephalogram
EHR	Electronic Health Records

EMG	Electromyography
EMI	Electromagnetic Interface
FPGA	Field Programmable Gate Arrays
HDL	Hardware Description Language
HLS	High-Level Synthesis
HMI	Human-Machine Interface
IoT	Internet of Things
LUT	Look Up Tables
MFCC	Mel-Frequency Cepstral Coefficients
ML	Machine Learning
MLC	Machine Learning Classifier
NB	Naïve Bayes
NLP	Natural Language Processing
PCM	Pulse Code Modulation
PDM	Pulse Density Modulation
PWM	Pulse width Modulation
RBF	Radial Basis Function
RF	Random Forest
RNN	Recurrent Neural Networks
SCI	Spinal Cord Injuries

SGD	Stochastic Gradient Descent
SNR	Signal-to-noise-ratio
SVM	Support Vector Machine
TDS	Tongue Drive System
VHDL	VHSIC Hardware Description Language.
VLSI	Very Large-Scale Integration
WHO	World Health Organization

Chapter 1

Introduction

The development of Assistive Technology (AT) has become increasingly high in recent years, as individuals with disabilities are depending on these devices for their daily activities. The main intention of AT development is to help disabled people to live more independently [1]. According to World Health Organization (WHO) in 2023, mentioned that around 1.3 billion people are experiencing disabled in the world [2]. Numerous disabled individuals depend on caregivers for every small thing, such as people with cognitive disabilities, mobility impairments, hearing impairments and visual difficulties [3]. Individuals with disabilities limit their social engagement because of these impairments and lack of confidence.

People who are suffering from polio, spinal cord injuries (SCI), amyotrophic lateral sclerosis (ALS), cerebral palsy (CP), Parkinson's disease, stroke survivors, neuromuscular disorders, quadriplegia, arthritis and elderly people will have difficulties of mobility and muscular weakness [4]. They need special care and help for day-to-day life to do daily tasks; without caregivers or helpers it will become difficult to perform or live their life. Assistive Devices (ADs) help disabled individuals live independently, improves their quality of life, and enhance them to engage with others and participation in social events.

1.1 Background of Assistive Technology

AT drives a wide range of instruments, techniques, and systems made to empower individuals with disabilities, AT improves their social interaction, independence, and overall quality of life. Due to advancements in engineering, technology, and computer science in recent years, the AT to significantly broadened for developing the systems in healthcare sector especially rehabilitation and assisting the individuals [5]. Initially, assistive instruments were rudimentary and operated manually, such as crutches, wheelchairs, and hearing aids. Advancement in the AT systems have developed in such a way that it made easy for the individual with disabled to access or use AT systems without putting lot of effort. Advance technologies may be

customized to meet the diverse and intricate requirements of users due to advancements in wireless communication, sensors, artificial intelligence, and microelectronics [6].

The development and utilization of AT are designed to support individuals with disabilities in their everyday activities, education, employment, and social interactions [7]. The prevalence of individuals with disabilities increases alongside the aging global population and the rising incidence of chronic illnesses [8]. Due to this, AT is important in facilitating the independent and productive lives of these individuals [9]. Mobility aids [10] (such as prostheses and powered wheelchairs), sensory aids [11] (such as screen readers and hearing aids), communication devices [12] (such as speech-generating devices), and cognitive aids [13] (such as memory aids and learning software) are examples of common forms of AT.

Mobility equipment, including motorized wheelchairs, are operated using joysticks [14], voice instructions [15], or tongue inputs [16] for navigation. Speech-generating devices (SGDs) are communication aids that convert user input via touch screens, switches, eye tracking, or tongue sensors into spoken words [17]. Hearing aids and cochlear implants process acoustic signals by amplifying or converting them into electrical signals for auditory processing [18]. Visual aids translate digital text or images from computers or cameras into speech or tactile Braille. Examples of these include screen readers and Braille displays [19]. Cognitive AD incorporates location data, reminders, and scheduling features to aid users with everyday chores [20]. Home appliances and electronics can be controlled by environmental control systems that react to voice instructions, tongue movements, or switch inputs [21]. Computer access tools that allow for interaction using head, mechanical, or breath motions include adaptable keyboards and sip-and-puff switches [22]. Educational aids facilitate reading, writing, and understanding by processing text, speech, or user responses [23]. Adaptive gaming consoles and other recreational devices make games more available by mapping buttons, using pressure, or using gestures [24]. Ultimately, assistive robots can assist with tasks such as nutrition, mobility, or companionship by responding to voice, joystick, or touchscreen inputs.

Human-Machine Interface (HMI) systems are a key aspect of AT, allowing users to communicate with electronic equipment using alternate input methods when conventional alternatives such as keyboards or joysticks are not available. Eye tracking

[25], brain-computer interfaces (BCIs)[26], and tongue drive system (TDS) [27] are a few of these technologies. TDS takes advantage of the tongue's remaining functioning, particularly in those with severe spinal cord injuries (SCI), to provide a dependable, fatigue-resistant way of control. The tongue is especially well-suited for assistive control since it can move quickly and precisely and frequently maintains function even in the most severe cases of paralysis [28].

1.2 Evaluation of Tongue Drive Systems

In the past two decades, the TDS has made remarkable progress, establishing itself as a promising AT for individuals with severe motor impairments, particularly those with high-level spinal cord injuries [29]. TDS provides a consistent, hands-free, wireless replacement for conventional input methods using the tongue's natural strength, sip and puff, and eye tracking even in quadriplegia [30]. The first TDS used magnetic-based devices, in which a small permanent magnet penetrated to the tongue could be tracked with external sensors on headgear or dental retainers. For tasks like computer and wheelchair control, these systems shown amazing accuracy, low latency, and utility [31]. New possibilities that have evolved over time include capacitive-based systems, which sense changes in capacitance when the tongue approaches sensor surfaces, and inductive or resistive systems, which depend on tongue contact or proximity to conductive pads [32]. Recently, compact intraoral TDS devices incorporated within dental retainers have been designed with the intention of enhancing user comfort, mobility, and aesthetics without compromising signal quality [33]. The latest research is hybrid systems combining TDS with voice commands or brain-computer interfaces (BCIs) to boost control and adaptation in demanding environments. Sensor miniaturizing, low-power wireless connectivity, machine learning based signal processing, and FPGA-based real-time control platforms have helped TDS evolve. The viability of TDS as a common assistive interface for independent living is growing as research tackles issues including interference, calibration, and long-term wearability. Internal TDS and arch-shaped intraoral TDS are newly developing research, which uses small magnetic sensors pierced on tongue or dental retainer to calculate the tongue movement. Which requires sensors to be inside the mouth.

1.3 Development of AI and ML for Healthcare Applications

Machine Learning (ML) and Artificial Intelligence (AI) have revolutionized the fields of healthcare, including medical diagnostics, treatment planning, and patient care. Large-scale medical data is analyzed by AI technology to improve decision-making, increase the accuracy of diagnoses, and expedite the delivery of healthcare [34]. AI has recently surpassed human experts in fields such as image analysis, illness prediction, and customized treatment, thanks to developments in deep learning and massive data [35]. Earlier attempts had concentrated on rule-based systems.

Nowadays, ML algorithms are crucial to radiology and pathology because they enable automated image analysis for quicker and more precise illness identification [36]. Electronic health records (EHRs) are analyzed by artificial intelligence (AI) systems in natural language processing (NLP) to give doctors useful information for diagnosis and prognosis [37]. Virtual health attendants and robotic surgery enabled by AI enhance patient engagement and surgical accuracy, while predictive analytics facilitate early disease detection and epidemic forecasting [38].

AI-enabled wearable technology allows for real-time health monitoring and provides patients with chronic illnesses with ongoing assistance [39]. However, as the use of artificial intelligence grows, it is essential to address concerns such as data protection, model openness, and regulatory compliance. Achieving the full potential of AI in healthcare and opening the door to completely automated diagnostics and precision medicine requires ethical AI development and continuous research.

1.4 Advancement in Machine Learning

Rapid advances in ML have improved signal processing, precision, and flexibility, which has greatly improved AT, especially TDS. Traditional TDS designs frequently had limited adaptability and higher error rates as a result of individual user variations, as they relied on predefined rule-based algorithms to detect tongue movements or speech inputs [40]. However, ML has transformed TDS functionality by enabling real-time learning, adaptive calibration, and noise-tolerant signal categorization.

ML-based approaches, such as Support Vector Machines (SVMs), RF, DT and Gradient Boosting have enabled the precise detection and classification of tongue

movements, speech signals, and bioelectric patterns, making TDS more effective for people with severe motor disabilities [41]. The technology can generalize among users due to models built on extensive datasets of tongue motion patterns and voice instructions, while yet providing customization based on individual characteristics [42]. The use of feature extraction techniques, such as Mel-Frequency Cepstral Coefficients (MFCCs) for speech-based TDS and magnetic sensor signal analysis for tongue-tracking systems, has enhanced command recognition rates and reaction times [43].

A significant aspect of ML in TDS is real-time adaptive learning, where the system continuously enhances its ability to recognize accuracy based on user behavior [44]. Individuals with different speech patterns, tongue flexibility, or neurological disorders that impact consistent input would significantly benefit from this. Research has been conducted on optimizing control systems using reinforcement learning approaches to minimize unintentional activations and ensure that TDS-based ADs, such as computer interfaces, wheelchairs, or prosthetics, respond appropriately to user commands.

ML combining with TDS is transforming AT. ML driven prediction models lower false positives and boost robustness, which enhances the TDS. These algorithms use time-series data, movement patterns, and contextual information to separate between deliberate orders and involuntary motions or background noise [45]. TDS is appropriate for practical usage in intelligent ADs like FPGA-based ML accelerators with low latency and can rapidly process data. In addition to mobility, environmental control systems, brain-computer interfaces (BCIs), and speech recognition have broadened the applications of TDS. More independence is being promoted by the ability for users to control communication devices, robotic arms, and home automation using voice commands.

Deep learning-based adaptive models, cloud-based personalization, and AI-powered self-calibration will be the primary focus of future advancements in AI and ML [46]. Artificial intelligence-based error correction, Internet of Things-enabled sensors, and efficient ML algorithms on embedded hardware will enhance the usability and accessibility of TDS [47]. These advancements are intended to enhance the use of TDS and facilitate its seamless integration into daily life, so enabling individuals with disabilities to exert more influence on their environment.

1.5 Machine Learning in VLSI

The use of ML into Very Large-Scale Integration (VLSI) has transformed real-time data processing, circuit design, and energy efficiency. ML identifies patterns in extensive datasets, so improving conventional VLSI methodologies and informing design choices towards more intelligence [48]. Techniques like Reinforcement Learning and Neural Networks are employed to optimize circuit behavior, reduce power consumption, and enhance transistor placement, hence improving performance [49].

MLs most significant applications in VLSI involve Field-Programmable Gate Arrays (FPGAs). Ideal for real-time ML requirements, these reconfigurable systems excel in parallel processing, minimal latency, and energy efficiency [50]. Integrating Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) inside ML models on FPGAs, utilizing High-Level Synthesis (HLS) and Hardware Description Languages (HDL), facilitates expedited inference and real-time processing [51]. FPGA-based ML accelerators offer flexibility and energy efficiency compared to conventional CPUs, with substantial applications in image processing, speech recognition, and embedded artificial intelligence systems [52].

Optimizing FPGA-based ML presents several issues, including latency, memory bandwidth, and resource use. Quantification and pruning facilitate the simplification of ML models for FPGA deployment [53]. Hardware and software co-design methodologies enhance the efficiency of programmable logic and processing systems by delineating activities inside architectures such as Intel Stratix and Xilinx Zynq.

The Nexys 4 Artix-7 FPGA offers a versatile platform for executing ML models while optimizing performance, power economy, and hardware adaptability. The Artix-7 FPGA accelerates ML algorithms for embedded artificial intelligence, real-time data processing, and categorization with its programmable logic and DSP slices [54]. Employing Verilog, VHDL, or HLS, ML models such as SVMs, RFs, and DT, etc. can enhance FPGA performance. Efficient matrix multiplication facilitated by on-chip DSP blocks is essential for deep learning applications such as image processing, speech recognition, and biomedical signal analysis. Quantization and hardware-friendly approximations are employed to simplify models, thereby providing a more practicable

approach to ML inference on FPGA resources [55]. The Nexys 4 Artix-7 FPGA has been utilized in low-power AI accelerators, real-time voice processing, and embedded classification systems, demonstrating its capacity to implement ML models for edge AI applications with minimal latency and good energy economy.

1.6 Research Gap

TDS have been adopted as popular for individuals with severe disabilities; however, their effectiveness and user adoption are restricted by a few obstacles. Wearable sensors, intraoral magnetic piercings, and head-mounted devices are the centerpieces of many TDS implementations now in use. These methods can be uncomfortable for extended periods of time and may result in irritation or pain from pressure. Furthermore, hygienic issues are brought up by magnetic tongue piercings, which need to be cleaned often to avoid bacterial accumulation and lower the risk of mouth infections. Systems dependent on outside magnetic sensors are also prone to interference from nearby metallic objects, therefore compromising dependability and accuracy. Moreover, real-time assistive applications still present major difficulties, particularly for those with limited mobility because of battery consumption, false activations, and low processing speeds.

The FPGA-based speech-driven TDS that features a built-in microphone may offer a solution to these issues. The Nexys 4 Artix-7 FPGA features an inbuilt MEMS omnidirectional microphone that facilitates direct, non-invasive audio capture, eliminating the need for wearable sensors or magnetic components. This method lowers maintenance requirements while simultaneously improving user comfort, hygiene, and system longevity. Furthermore, FPGA-based processing offers several advantages, such as hardware-level augmentation of ML models, real-time signal processing, parallel computation, and reduced power consumption.

FPGAs are optimal choice for real-time speech classification and wheelchair control applications due to their deterministic performance, reduced latency, and high-speed processing, which distinguish them from traditional microcontrollers or software-based methods. This research aims to deliver a highly efficient, comfortable, and sanitary TDS with FPGA-based ML and speech recognition, addressing the limitations of current wearable and magnetic systems.

1.7 Problem Statement

The thesis aims to address problems in existing TDS designs by developing an FPGA-driven, ML speech-based control TDS. This system seeks to deliver real-time data processing, categorization precision, and enhanced flexibility for assistive mobility devices. The emphasis will be on improving important parts for effective FPGA implementation, such as the EMI filter, SVM classifier, and PWM control. Quantization, pruning, and parallel processing will be investigated in the study using the Nexys 4 Artix-7 FPGA to investigate hardware-efficient ML techniques. The goal is to develop a robust, real-time TDS speech recognition system to enhance wheelchair control.

1.8 Objectives

- ❖ Design and evaluate the different ML model for audio data classification.
- ❖ Design a low power tongue drive system for processing the data with less noise and high accuracy and reduce the bit transmission rate for transmitting the data.
- ❖ Increase the accuracy of ML models by keeping reduced the number of computations and memory in the ML classifier by using appropriate algorithm.
- ❖ Hardware (FPGA) implementation of low power tongue drive system and calculate the power consumption, area and time.

1.9 Role of International Collaboration in supporting Objectives

The main goals of my thesis are to develop and construct a speech-based Tongue Drive System on an FPGA. However, international research collaborations played a big role in helping me reach these goals. Working together on Electroencephalography (EEG) seizure detection with SVM classifiers on FPGA gave us hands-on experience with hardware-efficient ML, which helped us reach our goal of implementing FPGA. Research on tongue contour extraction utilizing ultrasound images and CNN provided significant insights into articulatory feature extraction, enhancing classifier design and optimization goals. These collaborations did not represent independent objectives,

rather, they functioned as enablers, augmenting both the methodology and technical depth of the proposed work.

1.10 Thesis Organization

Thesis is organized as below:

- ❖ Chapter 1 – Introduce an overview of AT, with a special emphasis on TDSs, and advancement in the ML, for people with profound impairments. It specifies the research problem, objectives, and scope of the research.
- ❖ Chapter 2 – Provides a thorough analysis of current FPGA- based implementations, signal processing techniques, and TDS designs. It illustrates the deficiencies of existing systems and pinpoints areas for further investigation.
- ❖ Chapter 3 – Explained in detail the objectives, methodology and the data acquisition, EMI filtering, feature extraction, and SVM classification techniques employed in the TDS. Also describes the procedure for implementing an FPGA based Verilog.
- ❖ Chapter 4 – Detailed implementation of the real-time execution of TDS on the Xilinx Nexys 4 Artix-7 FPGA. It encompasses hardware integration, Verilog programming, power analysis, and system evaluation.
- ❖ Chapter 5 – Collaborative case studies in Biomedical signal and Image processing presents collaborative work on implementation of EEG seizure detection using SVM on FPGA board and other is Tongue contours extraction from Ultrasound Images. Although these projects were independent, they provided expertise in implementation of SVM on FPGA and tongue movement analysis using ultrasound, which can helpful in future development and getting insight idea of using other approaches for ADs. These collaborations have been crucial in enhancing methodologies and advancing the technical aspects of the speech-based Tongue Drive System.
- ❖ Chapter 6 – Discuss the results and examines experimental findings, encompassing classification accuracy, reaction time, power consumption, and real-time performance.

- ❖ Chapter 7 – Concludes the key contributions of the research and discusses its limitations. It suggests future work directions, including AI-driven adaptive learning, cloud-based monitoring, and enhanced miniaturization for wearable applications.

Chapter 2

Review of Literature

ATs and ADs are essential for improving the quality of life of older citizens, individuals with disabilities, and those with mobility challenges. Brain-computer interfaces (BCIs), speech-controlled wheelchairs, prosthetic limbs, exoskeletons, and smart home automation systems are examples of better and more efficient solutions made possible by developments in biomedical engineering, artificial intelligence, and embedded systems [56]. The World Health Organization (WHO) estimates that over 2.5 billion individuals worldwide require AT, and that figure is only expected to increase by 2020 about 3.5 billion as the population ages and the frequency of physical and neurological disorders rises [57]. Sensor-based, Artificial Intelligence AI-driven, and IoT-enabled systems that provide more independence and adaptability are gradually replacing or upgrading more conventional devices like manual wheelchairs and hearing aids [58]. Speech-controlled and tongue-controlled systems offer intuitive solutions for those with significant motor limitations, facilitating hands-free operation and seamless contact with their environment. ADs have also been revolutionized by FPGA-based real-time processing, which enables low-power, high-speed computation in adaptive control, motion analysis, and voice recognition systems [59]. Nonetheless, obstacles including accessibility, price, and customization remain, highlighting the necessity for continuous research and innovation to create cost-effective, energy-efficient, and user-friendly ATs that can serve a broader global population.

The purpose of ADs is to empower people with disabilities by improving their ability to communicate, navigate, and interact with their surroundings. These devices are categorized into many types that are customized to fit specific needs, such as environmental control systems, mobility aids, communication aids, and prosthetic/rehabilitation devices. Mobility aids, such as powered wheelchairs and exoskeletons, provide independence through user-friendly controls including joysticks, touch interfaces, voice commands, Electroencephalogram (EEG), Electromyography (EMG) or ultrasound images [60]. The use of communication aids, such as text-to-speech devices and eye-tracking systems, enables individuals with speech impairments

to establish communication with others by speech recognition, gestures, or gaze-based inputs. Environmental control systems (ECS) enable users to regulate smart home gadgets, doors, and lighting via voice commands, facial recognition, or EEG-based controls, promoting a sense of autonomy. Prosthetic and rehabilitation systems that use bio-signals, motion sensors, or neural implants to smoothly adapt to real-life settings include bionic limbs and brain-controlled robotic arms [61]. Advancements in ML, sensor fusion, and real-time processing are enhancing the speed, accuracy, and user-friendliness of these devices. By combining AI, IoT, and embedded computing, AT is evolving to provide people with mobility disabilities with increased autonomy, self-confidence, and quality of life.

Individuals with severe motor disabilities, such as quadriplegia, amyotrophic lateral sclerosis (ALS), and spinal cord injuries, often face challenges in mobility and independent living. Assistive technologies such as Brain-Computer Interfaces (BCI) and Eye-Gaze Systems have been explored for mobility control; however, these solutions are often expensive, complex, and require extensive calibration [62].

A TDS is an innovative hands-free control mechanism that leverages tongue movement to operate external devices, such as wheelchairs, computers, and robotic systems [63]. The tongue is an ideal control organ because it retains mobility even in individuals with severe motor impairments. Most traditional TDS implementations rely on magnetic sensors, capacitive touch sensors, or electromyography (EMG) electrodes to detect tongue movements and translate them into control commands [64].

The Tongue Drive system is designed to empower individuals with severe impairments who rely on assistive equipment, leveraging the remarkable versatility of the tongue. This innovative technology involves attaching an array of Hall-effect magnetic sensors to the outer edge of the teeth using a dental retainer. These sensors detect the magnetic field produced by a small permanent magnet affixed to the tongue. The system translates these signals into precise control for tasks such as moving a computer cursor, operating a motorized wheelchair, or interacting with mobile devices. Unlike conventional switch-based systems with limited on/off functionality, the Tongue Drive offers users smooth and proportional control by analyzing multiple sensor outputs. Using FEMLAB simulations and experimental data, authors evaluated the effects of the magnet's orientation and position on sensor performance. A prototype

system, built from off-the-shelf components and integrated with a graphical user interface in LabVIEW, demonstrated the system's potential. Authors focused on developing a compact, battery-powered wireless mouthpiece, streamlining the design for enhanced usability [65].

In this paper [66], authors evaluated the performance of the TDS as both a computer input device and a powered wheelchair controller. Tasks were designed based on the ISO 9241-9 standard for pointing and selection, utilizing Fitts' Law, and an obstacle course navigation task. Nine participants with pre-existing tongue piercings volunteered for the trial, which spanned five sessions over five weeks, allowing for an analysis of the TDS learning curve and its current limitations. Participants used titanium tongue rings resembling barbells, each equipped with a securely sealed rare earth magnetic tracer in the top ball. Results shows that the performance measured improved significantly from the first session to the last session.

In paper [67], author presented the dual-mode TDS of wireless and wearable AD. Which targets the enhancing speed, usability, adaptability and autonomy through the tongue motion and speech for utilizing the computers for significant impairments. The dTDS was employed by using magnetic tracer, ergonomic and compact magnetic sensors, and wireless headset for monitoring tongue movements to control computer. The headgear/headset has a tiny microphone for the wireless recording of users' voices. Preliminary assessments involving 14 able-bodied participants indicated that the dTDS headset, in conjunction with commercially available speech recognition software, exhibited markedly superior performance compared to techniques dependent solely on tongue movement or vocalization, especially for tasks necessitating both pointing and text input.

Author [68] introduced a novel technology which was employed by a tiny magnet pierced to the tongue and monitored by magnetic sensors placed on headset or mouth brace. Commands like forward, turning left and right or stop interpreted from collected signals and transferred to computer wirelessly. The Tongue Operated Robotic Wheelchair, utilizing TDS concept, enables patients with significant impairments to maneuver their wheelchair with basic tongue movements. The hardware comprises a transmitter equipped with Hall Effect sensors, an encoder, and a microprocessor, in addition to a receiver with an RF receiver and decoder. The algorithm utilizes a round-

robin polling technique to continuously read sensor information. Figure 2.1 shows the placement of sensors and their threshold values. Each sensor possesses a threshold: S1 activates forward movement over 750, S2 initiates a left turn above 600, S3 prompts a right turn above 700, and S4 halts the vehicle above 700. Flag variables monitor sensor activity and reset upon completion of actions. Research indicates that TDS facilitates seamless, accurate navigation, presenting a user-centric, non-intrusive approach applicable across many bodily postures. It is economical, accommodates various commands, and promotes autonomy. Challenges including possible pain from oral components, weariness with prolonged usage, and technological difficulties such as signal delays and system complexity.

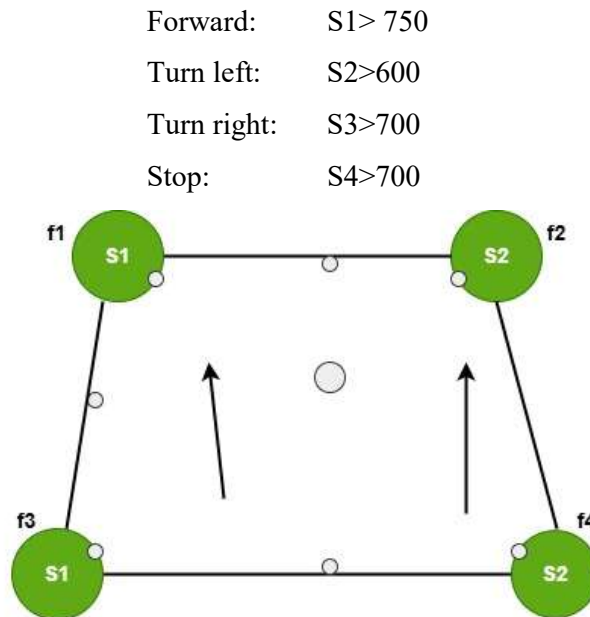


Figure 2.1 Magnetic placement [68]

The working model of TDS were divided into two, one for transmitting the data and the other is receiving the data. Figure 2.2 shows the block diagram of Transmitting the data at transmitter side of TDS. Which consists of four sensors to collect the data when the pierced tongue with magnet moves. Which sends the collected data to 32-bit microcontroller LPC2148. Power supply is connected to the microcontroller. RF encoder is connected to RF transmitter. This data which is saved in microcontroller is transmitted to the receiver by RF transmitter.

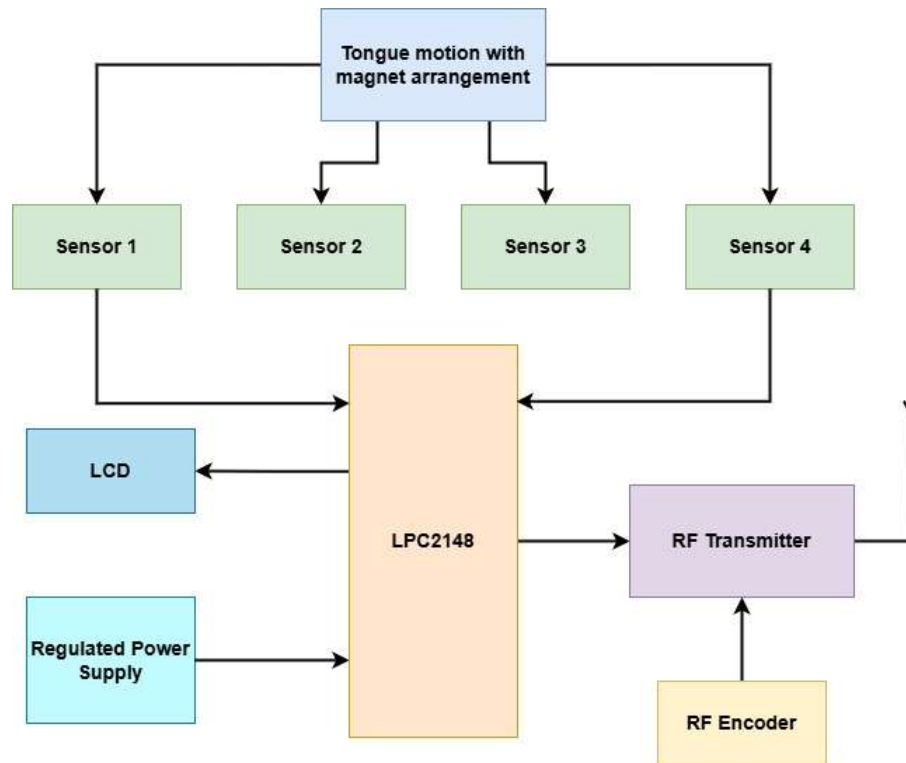


Figure 2.2 Block diagram of TDS Transmitter [68]

The above Figure 2.3 explains the Receiver block diagram for TDS, the data which is transmitted through the RF transmitted by RF receiver block and decodes the data using RF decoder. Which is sent to 32-bit microcontroller. Microcontrollers are connected to power supply. This microcontroller gives the command signals to motor driver L293D. Motor driver drives the two DC motors of wheels. Motor 1 and Motor 2 are two DC motors connected to L293D motor driver.

The article highlights the Hall effect and the design of a TDS, providing an innovative control system for powered wheelchairs (PWCs). The system translates tongue motions into control signals with a magnetic sensor configuration that monitors a tiny permanent magnet attached to the user's tongue. Hall effect sensors detect these motions, converting them into commands that control the wheelchair's speed and direction. A unique interface including an LCD display offers immediate feedback and displays current commands. The control devices, using H-bridge circuits and relays, control the wheelchair's motors. Signal processing analyzes the motions of the tongue, converting them into signals for accurate directional control. The technology provides

immediate input via the LCD, enabling users to control several items, including televisions, fans, and lights, independently [69]. The block diagram of figure 2.4 explains the generation of electrical signals using magnetic sensors for controlling the assistive motor driver using hall effect.

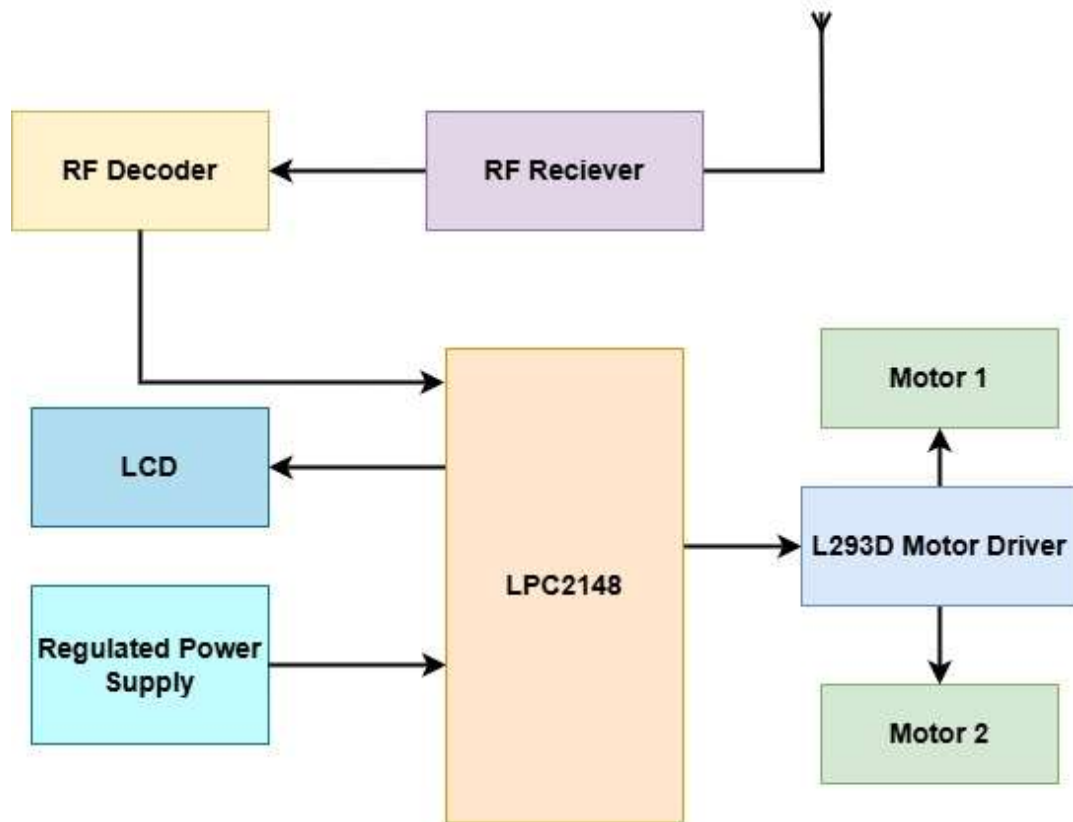


Figure 2.3 Block diagram of TDS Receiver [68]

The article highlights the Hall effect and the design of a TDS, providing an innovative control system for powered wheelchairs (PWCs). The system translates tongue motions into control signals with a magnetic sensor configuration that monitors a tiny permanent magnet attached to the user's tongue. Hall effect sensors detect these motions, converting them into commands that control the wheelchair's speed and direction. A unique interface including an LCD display offers immediate feedback and displays current commands.

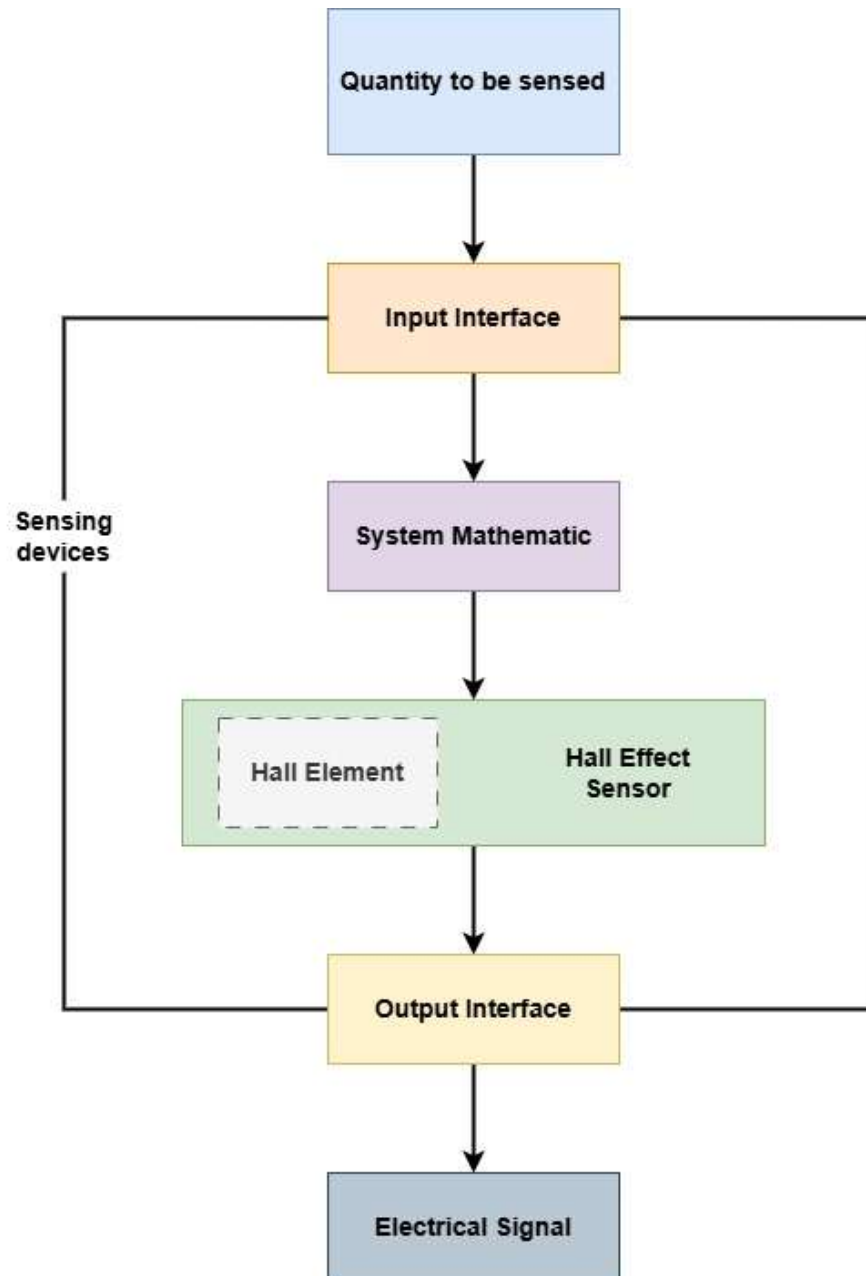


Figure 2.4 Working model of generating electrical signals to control the device [69]

The control devices, using H-bridge circuits and relays, control the wheelchair's motors. Signal processing analyzes the motions of the tongue, converting them into signals for accurate directional control. The technology provides immediate input via the LCD, enabling users to control several items, including televisions, fans, and lights, independently [69].

The author presented the iTDS-2, an arch-shaped intraoral TDS intended for the buccal shelf. It incorporates a SoC for the amplification and digitization of magnetic sensor data, wirelessly transferring it to a TDS-UI using an inductive coil or planar inverted-F antenna. The device employs a dual-band radio (27 MHz or 432 MHz) for wireless communication, utilizing a super-regenerative receiver to assess link quality and switch bands as necessary. An energy-efficient FPGA produces data packets and oversees digital control. The TDS-UI analyzes sensor data on a smartphone via an SSP algorithm, issuing orders to designated devices. Human subject trials demonstrated a 22% improvement in performance compared to the previous model [69].

In this paper [70], effectiveness of the TDS was assessed through studies with nine able-bodied (AB) individuals and 11 individuals with tetraplegia (TP). AB participants utilized a mouse, keypad, and TDS, whereas TP participants just employed the TDS. Tasks including Fitts' law tapping, wheelchair operation, phone calling, and weight shifting, executed across five to six sessions. Although AB participants were acquainted with the keypad, their tapping task throughputs were just 1.4 times greater with the keypad than with the TDS. Completion durations for wheelchair driving activities varied between 157 seconds and 180 seconds across three control techniques. TP participants executed phone-dialing and weight-shifting tasks in 81.9 seconds and 71.5 seconds, respectively, utilizing tongue movements. Performance enhancements were statistically significant or exhibited favorable trends over the sessions, with most learning transpiring during the first and second sessions. Ongoing practice indicated the possibility of more performance improvement with the utilization of the TDS.

The author designed a TDS for a real-time voice-controlled wheelchair device for those with mobility issues. The system incorporates speech recognition and obstacle detection with a microcontroller-based configuration, providing a cost-effective option for those with physical disabilities. The system utilizes an Arduino UNO and employed a motor driver for wheelchair movement, ultrasonic sensors for obstacle detection, and a VR3 voice recognition module. The device interprets voice instructions including "forward," "backward," "left," and "right," and ceases operation upon detecting an obstruction. The findings indicate that the wheelchair proficiently reacts to spoken commands and circumvent obstacles, rendering it a pragmatic and advantageous option for practical use [71].

The study presents a non-invasive, multifaceted AD that utilizes the tongue's inherent dexterity and the brain's sensitivity to aid or augment upper limb functions. The study introduces innovative tongue-activated architecture tailored for wheelchairs, computer interfaces, robotic rehabilitation, and smart gadgets, surpassing traditional systems reliant on eye tracking, hand motions, or surgical interventions. It enhances magnetic-based TDS by ensuring accurate and reliable functionality without intrusive procedures, while integrating tongue input with vocal and head movements to offer optimal versatility. This methodology demonstrates enhanced accuracy, speed, and user-friendliness across several contexts, including virtual reality and space, compared to previous methods that concentrated on restricted tasks [72].

The Tongue Drive System (TDS), a wireless, minimally invasive assistive tool that converts voluntary tongue motions into user-defined commands, was created by the authors to increase accessibility for people with physical limitations. Designed as a mild and pleasing dental retainer, a new intraoral variation of the TDS called iTDS was created. Four 3-axis magneto resistive sensors incorporated onto the printed circuit board (PCB) form a system-on-a-chip (SoC) with a programmed analog front-end (AFE) intended to detect magnetic field variations. Within the ISM band, the SoC has a dual-band transmitter operating at 27 and 432 MHz that sends data to an external TDS universal interface composed of generally accessible components. With technologies like telephones, computers, and motorized wheelchairs, this interface makes smooth connection possible. Comprising a power management integrated circuit (PMIC), the SoC charges a 50 m Ah Li-ion battery, regulates voltages, and maximizes power consumption. Appropriate for daily usage, the system runs at an average power of 3.7 mW, is built using a 0.5- μ m CMOS technology, and allows its setup utilizing a 13.56 MHz inductive connection [73].

Small sensor coils are placed on the tongue to track tongue movements using electromagnetic (EM) sensors, which belong to Electromagnetic Articulography (EMA), where it measures the position of parts of the mouth [74]. When speaking or swallowing, these coils precisely track the position and movement of the tongue in three dimensions by utilizing an electromagnetic field produced by external transmitters [75]. The ability to capture tiny, continuous, multidimensional tongue motions is the primary

advantage of this approach, and it is crucial for applications requiring extensive articulatory data [76].

There have been some limitations, using Audas for data collection/input to the system. The accuracy of EMA depends a lot on where and how the sensors are set up. Even little alterations might lead to huge inaccuracies in results. Also, items like metal objects or electromagnetic interference in the vicinity could damage the system's functionality by changing the magnetic field and making the data less reliable [77],[78]. The placement of sensors inside the mouth is shown in figure 2.5, EM sensors are placed on upper lip, lower lip, left corner of mouth, tongue tip and tongue dorsum to analyze the tongue. The author collected large kinematic dataset [79]. To acquire accurate readings, most users must calibrate their device, which can be a pain to set up and make the system less comfortable and useful.

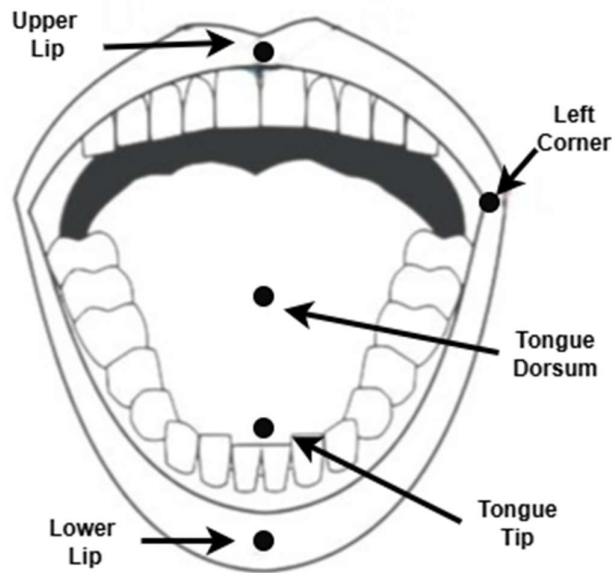


Figure 2.5 EM sensor placement for data collection [79]

Alternative methods of data modalities have been investigated; there are multiple ways to collect the data of speech and tongue articulatory motion. Ultrasound tongue imaging (UTI) employs sound waves to display the shape and motion of the tongue in real time, rendering it a non-invasive technique [80]. Magnetic Resonance Imaging (MRI) gives detailed three-dimensional images of the tongue and vocal tract [81]. Nonetheless, it is not very beneficial for continuous speech because of its cost

and complexity of use. Electropalatography (EPG) monitors tongue contact with the palate, rendering it valuable for analyzing consonant articulation [82]. Optical motion capture utilizes cameras and reflective markers to document visible motions of the tongue and lips [83]. Using microphones for acoustic research provides speech output without limitations. It is painless as it requires no sensors or attachments to the tongue or palate. There are several advantages of using an FPGA with an integrated microphone for speech-based operations. It reduces the requirement for exact sensor alignment, simplifies hardware setup, and is less susceptible to environmental influences. Additionally, advancements in signal processing and classification algorithms enable precise command recognition, providing a robust and user-friendly solution for TDS ADs.

This research presents an alternative speech-based TDS for wheelchair operation utilizing FPGA-based signal processing and an integrated microphone. In contrast to traditional TDS that need intrusive sensors, our approach enables users to provide spoken commands using tongue-articulated speech, thereby offering a non-invasive and user-friendly alternative.

Chapter 3

Methodology

Over the past ten years, speech-based control for ADs has experienced notable developments. For those with mobility challenges, a TDS employing speech recognition offers a viable method for wheelchair navigation using voice instructions. This chapter describes a methodical approach for creating a TDS using FPGA technology to run a voice-command wheelchair.

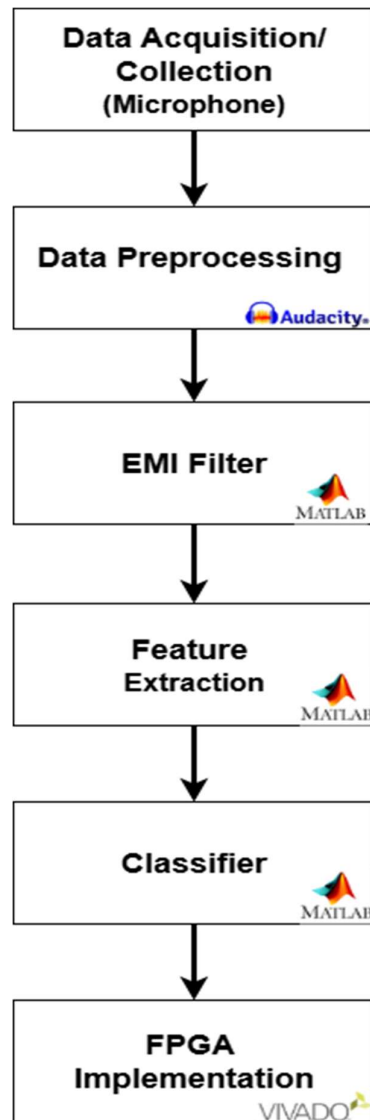


Figure 3.1 Flowchart of speech- based TDS implementation on FPGA

The FPGA implementation of speech-based TDS main blocks are data collecting, pre-processing, EMI, feature extraction, classification, and Implementation. Figure 3.1 shows the flow diagram of speech-based TDS implementation, the process spans hardware and software development. Data collection is done using microphone and preprocessing of dataset uses AUDACITY software for tempo change, normalization, formatting, trimming and sizing. Every movement is precisely contextualized by careful description. The primary goal is to provide physically challenged individuals with a low-latency, trustworthy, and energy-efficient solution using the real-time processing capability of FPGAs. Usually, high sensitivity and noise-reducing microphones with assistance technology collect speech signals for them. Because of its small size and dependability, MEMS microphones have been investigated as embedded systems' tool. Studies show that onboard microphones included into FPGA boards such as Nexys 4 offer real-time data gathering with lowest latency. The proposed speech-based TDS captures audio signals utilizing an onboard microphone built into the FPGA development board. The microphone is securely mounted on the hardware platform and captures the audio signals produced by the tongue during experimental sessions. This architecture enables immediate real-time signal processing within the FPGA. No need of wearing uncomfortable wearable devices to capture the commands.

3.1 Sensor-based Tongue Drive Systems

Sensor-based TDS detects tongue movements or physiological data and translates it into commands for assistive technologies like computers or wheelchairs. These systems utilize magnetic, capacitive, optical, electromagnetic, MEMS, auditory or speech-based, ultrasonic, and electromyographic sensors as their data collection methods. Magnetic TDS identifies a magnet positioned on the tongue by the utilization of Hall-effect or magneto resistive sensors [84]. Capacitive TDS use variations in electrical capacitance to detect touch or proximity [85]. Optical methods detect the tongue by light reflection or obstruction [86]. MEMS-based systems monitor inertial motion with accelerometers and gyroscopes; electromagnetic TDS identifies field fluctuations utilizing coils and sensors [87]. While ultrasonic systems utilize echolocation to depict the tongue's shape [88], MEMS microphones are employed in

speech-based acoustic TDS to capture spoken instructions. In electromyography TDS, surface electrodes monitor muscle activity [89]. These systems enhance the effectiveness and accessibility of AT because to their adaptation to users' motor skills, real-time functionality, portability, and non-invasive characteristics.

3.1.1 Magnetic TDS

Systems built on magnetic sensors are among the first and most well-known types of TDS. These devices track the movement of a tiny permanent magnet attached to the tongue using neighboring magnetic sensors like magnetometers or Hall-effect sensors. By turning their tongue gestures into separate instructions, people with quadriplegia or ALS can considerably benefit from their capacity to manage objects such computers, wheelchairs, and smart home appliances. Using the tongue's intact motor control, an organ innervated by cranial nerves and unaffected by spinal cord damage. Magnetic TDS allows patients to interact with AT devices as computers or motorized wheelchairs. The basis of this system is a tiny permanent magnet usually made of neodymium because of its compact form and strong magnetic field. This magnet is affixed to the tongue often via tongue piercings or other biocompatible procedures. Usually found all over the user's face, magnetic field sensors like three-axis magnetometers or Hall-effect sensors are positioned around the cheeks. These sensors monitor changes in the tongue's magnetic field and are installed on wearable technology such as glasses or headphones [90].

Whenever the user moves their tongue, a magnet moves along with it. These fluctuations in the magnetic field are detected by sensors, which then translate the position of the tongue into a command. The magnetic field $\vec{B}$ generated by a magnetic dipole (the tongue magnet) is given by

$$\vec{B} = \frac{\mu_0}{4\pi} \cdot \left[\frac{3\vec{r}(\vec{m} \cdot \vec{r})}{|\vec{r}|^5} - \frac{\vec{m}}{|\vec{r}|^3} \right] \quad (3.1)$$

Where μ_0 is permeability of free space= $4\pi \times 10^{-7}$ T-m/A, $\vec{m}$ is magnetic moment vector of the magnet, $\vec{r}$ is vector from the magnet to the sensor, $|\vec{r}|$ distance between magnet and sensor, $\vec{B}$ is magnetic flux density at sensor location [91].

$$B = \frac{\mu_0 \cdot m}{2\pi \cdot r^3} \quad (3.2)$$

A simplified formula for field strength from a dipole is given below eq (3.2). Where m is magnetic moment and r is distance from the magnet and sensor.

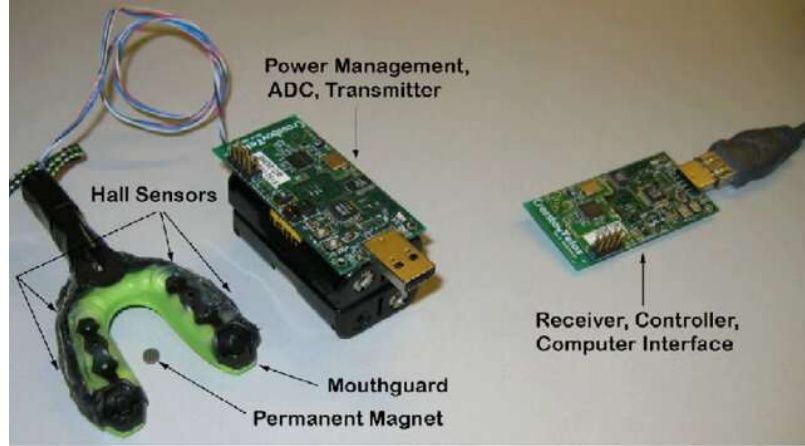


Figure 3.2 Magnetic sensor-based TDS

The magnetic field detected by sensors is altered because of the magnet's position being shifted by the movement of the tongue. Each tongue move has its own spatial imprint due to these variances. Sensors record the variations as analog voltage signals, which are then analyzed by a microcontroller, FPGA, or embedded system after being converted to digital from analog voltage signal using ADC [92]. Figure 3.2, shows the sensors and hardware used for magnetic sensor-based TDS. Algorithms like vector mapping and ML can read these signs and turn them into commands to the device like "stop," "turn left," and "move forward."

The system recognizes commands and produces digital output signals to operate AD such as motorized wheelchairs, computers, and smart home appliances. In contemporary FPGA-based systems, this process operates in real time with high precision and little latency. The utilization of FPGA modules for real-time signal processing, filtering, feature extraction, and classification enhances the method's responsiveness, scalability, and energy efficiency compared to conventional microcontroller-based systems [93]. This hardware-based method is essential for safety-critical applications where precision, reliability, and promptness are paramount.

3.1.2 Capacitive TDS

Capacitive TDS operates beneath the mouth and is an AT reliant on fluctuations in capacitance resulting from the tongue's interaction with electrodes. Devices such as dental retainers, intraoral instruments, or mouthguards incorporate these embedded electrodes [94]. The AD allows the disabled to control the devices, like wheelchair, computers, mobiles, smart homes and assistive robots by converting certain tongue motions to electrical signals.

Capacitive sensing detects variations in the electric field between two conductive components (plates) as the tongue approaches or contacts them. Because of its electrical conductivity and dielectric properties, the human tongue is ideal for AD [95]. The capacitance formula is presented below eq (3.3), Where ϵ is permittivity of the material (tongue/ saliva/ air), A area of the electrode plate, d is distance between electrodes and C is the capacitance.

$$C = \frac{\epsilon \cdot A}{d} \quad (3.3)$$

The capacitance (C) rises when the tongue gets closer to a sensor pad because the distance (d) gets smaller. Real-time measurements are employed to quantify this variation. After calculating the capacitance (C) between the tongue and the implanted electrodes using the parallel-plate formula, the capacitive TDS begins signal analysis. A custom-fitted intraoral dental retainer with a grid of capacitive sensors is worn on the upper palate by the user. The capacitance at specific electrodes is measurably increased as the tongue moves toward or contacts them, resulting in a reduction in the distance (d). Capacitive sensing circuits, such as RC oscillators or ADC-based capacitance-to-voltage converters, detect these changes by translating raw capacitance variation into electrical signals. These signals are received by a microprocessor inside the device, which then filters and normalizes mouth noise, moisture artifacts, and electrical interference. After processing, the data is transformed into a 2D spatial "capacitance map," which displays the closeness of each electrode to the tongue. From this map, feature vectors that describe the characteristics of gestures (such pressure, duration, and position) are extracted and fed into a ML classifier (e.g., SVM, decision tree, or neural

network). Several tongue motions, including swipes, taps, and holds, are recognized by the classifier, and it connects these gestures to specific commands. Bluetooth or RF modules wirelessly transmit these recognized orders to an external device like a wheelchair, computer, or smartphone. The outcome is a seamless and intuitive control system that enables individuals with severe physical limitations to interact with their environment solely through tongue movements. Some systems even employ feedback loops like haptic pulses or auditory tones to ensure command execution and increase user trust [96]. Capacitive TDS serves as a non-invasive, silent, and unobtrusive alternative to magnetic systems, particularly in AT applications.

3.1.3 Electromagnetic TDS

A wireless intraoral AT called electromagnetic TDS translates tongue motions into commands for device control by means of interactions between electromagnetic fields. Its working basis is electromagnetic induction and mutual coupling of receiver and transmitter coils [97]. These sensors detect changes in the electromagnetic field produced by tongue movements inside the oral cavity. Usually situated on a dental retainer or helmet close to the cheekbones or palate, a sensor array of receiver coils is found; generally, the arrangement is with a tiny electromagnetic transmitter implanted in a tongue stud or magnet. Tongue movements modulate the relative locations of the transmitter and receiver coils, therefore changing the magnetic flux passing through the receiving coils [98].

These changes lead to variations in the induced voltage, which are recorded and analyzed for device control. The induced voltage (V) in the receiver coils is derived by Faraday's Law, which presented in eq. (3.4), Where V is induced voltage, N is number of turns in the coil, Φ is magnetic flux and $\frac{d\Phi}{dt}$ is rate of change of magnetic flux due to tongue motion. The magnetic flux calculated by eq. (3.5).

$$V = -N \frac{d\Phi}{dt} \quad (3.4)$$

$$\Phi = B \cdot A \cdot \cos(\theta) \quad (3.5)$$

Where B is magnetic field strength, A is area of the coil and θ is angle between magnetic field and coil plane.

The receiver coils detect changes and send the signals to a conditioning circuit, which removes ambient noise and improves the essential data. Important characteristics including amplitude, duration, and spatial coil activation patterns are subsequently extracted from these signals through analysis [99]. The features are fed into a classifier, often based on ML techniques like Support Vector Machines (SVM) or neural networks, to determine the user's intended command (e.g., move forward, turn left, select). The detected command is wirelessly sent to external devices, such computers, Internet of Things gadgets, or motorized wheelchairs. Enabling precise, low-latency, hands-free operation, this system is appropriate for individuals with severe motor impairments. Continuous location tracking is supported by electromagnetic TDS, which may provide greater gesture resolution than magnetic or capacitive TDS.

3.1.4 Optical TDS

The Optical Tongue Controlled Assistive Device (OTCAD) is a non-invasive, wireless technology that helps people who are paralyzed. Optical TDS are a new form of HMI that employs light technology to monitor tongue motions [100]. Unlike magnetic or capacitive systems, which depend on electrical capacitance changes or magnetic fields, optical TDS monitors the location and mobility of the tongue using visible or infrared light in conjunction with photodiodes or image sensors. This method provides a non-invasive, contactless solution, which may make it more pleasant and sanitary for prolonged usage.

$$I = P / 4\pi r^2 \quad (3.6)$$

Where I is the light intensity measured by photodetector, P is power of the emitted light, r is the distance from light source to the tongue.

$$\Delta x, \Delta y = \text{optical flow}(I_t, I_{t+1}) \quad (3.7)$$

Where I_t and I_{t+1} are two sequential frames captured by camera.

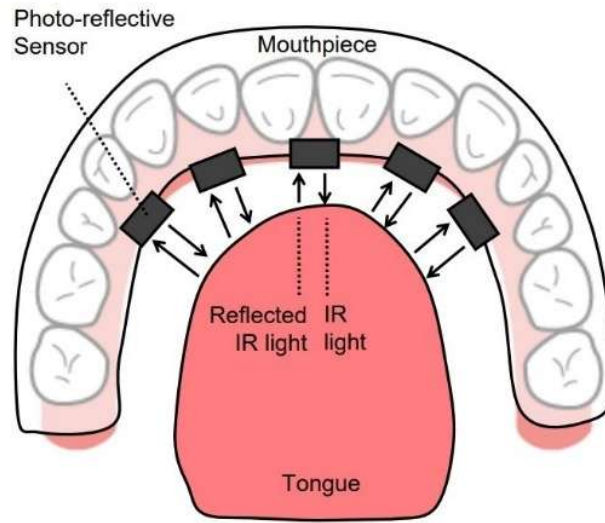


Figure 3.3 optical sensors for TDS

Optical TDS monitor tongue movements by using light reflection or occlusion. The oral cavity or a reflective marking on the tongue are illuminated by a light source, such as an LED or infrared (IR) emitter. The tongue's movement modifies the intensity, angle, or pattern of light, which is caught by photo detectors or camera sensors. These changes are used to calculate the position and orientation of the tongue using methods like the inverse square law of light or optical flow techniques, Figure 3.3 shows the optical sensor which is used to collect the tongue movement for TDS [101]. The system associated with tongue movement commands like left, right, select, or stop, etc. Optical TDS provides non-invasive, high-resolution tracking, but it can be effect by saliva, food obstruction, and ambient illumination. It needs a clean line of sight for optical sensors to get accurate performance [102].

3.1.5 Electromyographic TDS

An electromyographic tongue drive system (EMG-TDS) is an innovative AT that enables individuals with significant motor impairment to operate devices using tongue movements. In contrast to conventional TDS systems that employ magnetic sensors, EMG-TDS utilizes surface electromyography (sEMG) to detect electrical activity generated by tongue muscle contractions [103]. This advanced technology offers a non-invasive, wireless, wearable solution that enables individuals to control

wheelchairs, computers, and other gadgets with more autonomy and enhanced quality of life. Figure 3.3 (b) shows the working of electromyographic sensor for collecting tongue movements.

$$RMS = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} EMG_{l,n}(p)^2} \quad (3.8)$$

Where N represents the number of samples in one analysis window, l represents the number of electrodes and $EMG_{l,n}$ represents the EMG signal of the l^{th} electrode in the n^{th} sample of p^{th} analysis window.

$$CC_l^n(p) = \frac{1}{N} \sum_{k=0}^{N-1} \log |X_l^k(p)| e^{j2\pi kn/} \quad (3.9)$$

$$X_l^k(p) = \sum_{n=0}^{N-1} EMG_{l,n}(p) e^{-j2\pi kn} \quad (3.10)$$

The EMG-TDS functions by collecting muscle signals from the suprahoid area (located under the jaw) using a flexible sensor array equipped with numerous electrodes. Specific tongue motions by the user elicit muscular activation that produces EMG signals [104]. The signals are further amplified and filtered, with characteristics including Root Mean Square (RMS) and cepstral coefficients (CC) recovered. A Support Vector Machine (SVM) classifier analyzes these attributes to associate them with specific instructions, facilitating accurate control of equipment such as wheelchairs.

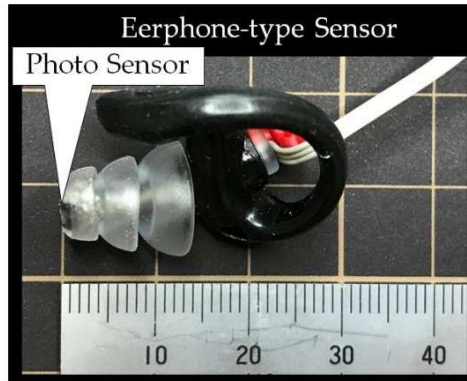


Figure 3.4 (a) Photo sensor [103]

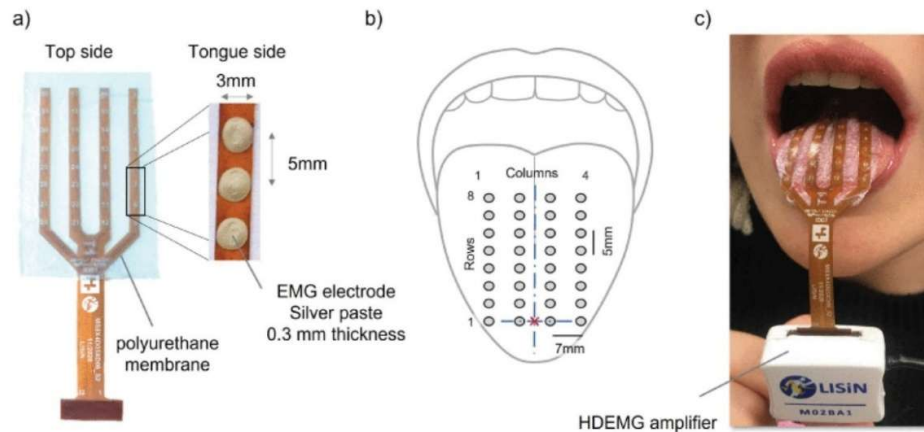


Figure 3.4 (b) EMG placements for electromyographic TDS [104]

EMG-TDS offers advantages such as being non-invasive, speech-friendly, compact, lightweight, and capable of real-time control. However, it is sensitive to electrode placement and motion artifacts, requires regular calibration, and may exhibit variability across users due to differences in muscle anatomy, as well as signal noise caused by skin impedance or sweat. Figure 3.4 (a) & (b), shows the hardware parts of photo sensor and EMG sensors which used to collect the data for TDS.

3.1.6 Ultrasound-based TDS

The Ultrasound-based TDS is an innovative AT that enables patients with significant physical limitations to control computers, wheelchairs, and other equipment with basic tongue motions [105]. This TDS system utilizes ultrasound imaging to track the tongue's position and movement in real-time, in contrast to previous systems that rely on magnetic or electromyographic signals.

The technology is intended to be intuitive and unobtrusive. An ultrasound transducer is positioned behind the chin, aimed at the floor of the mouth to get dynamic sagittal-plane pictures of the tongue during movement. The photos are further processed utilizing modern methods. Figure 3.5 shows the placement of ultrasound transducer for collecting the tongue motions [106].

The distance (d) between the tongue and the transducer is calculated from the below formula ToF, which is time taken for ultrasound signal to travel to the tongue and back. The calculation of d helps to map the exact position of the tongue in space in real time [107].

$$ToF = \frac{2d}{v} \quad (3.11)$$

$$d = \frac{v \cdot ToF}{2} \quad (3.12)$$

When multiple ultrasound transducers are used, triangulation helps locate the tongue in a 2D or 3D space:

$$(x, y, z) = f(d_1, d_2, d_3 \dots) \quad (3.13)$$

Where d_1 , d_2 , and d_3 are distance measured from 3 different sensor positions and f is the geometric function depending on the transducer layout. To accurately find ToF, the system uses cross-correlation between the transmitted and received signals given below:

$$R_{xy}(\tau) = \int x(t)y(t + \tau)dt \quad (3.14)$$

Where $x(t)$ is the transmitted signal, $y(t)$ is the received signal and τ time delay is, ToF corresponds to the maximum of $R_{xy}(\tau)$.

Initially, edge detection methods, such Sobel and Canny, detect the shape of the tongue. Subsequently, categorization techniques, such as template matching or ML algorithms, align the identified form with a predetermined gesture repository [108]. Normalized cross-correlation is frequently employed in template matching, whilst ML methods like Support Vector Machines (SVM) and Convolutional Neural Networks (CNNs) are developed to identify tongue patterns. In further steps, the system converts these tongue motions into distinct user instructions, providing a smooth and intuitive interface for those with restricted mobility [109].

Ultrasound-based TDS provides a non-invasive method as it eliminates the need for foreign items in the mouth cavity. Because of this, it is appropriate for both speaking and oral consumption. It is appropriate for a variety of tongue motions as it gathers high-resolution spatial information on the tongue's direction and shape [110]. Additionally, the device is becoming increasingly wearable and small thanks to developments in portable ultrasound technology.

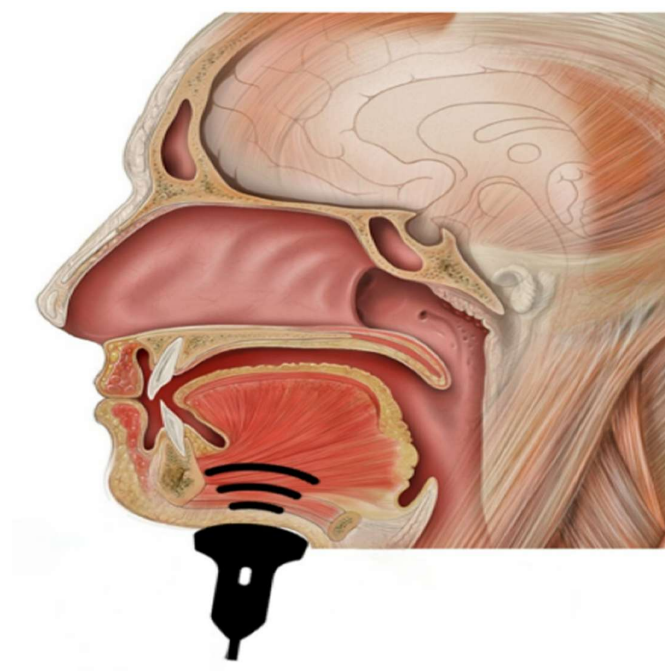


Figure 3.5 Ultrasound transducer placement for collecting ultrasound images [110]

However, issues persist. Although compact, the ultrasonic device necessitates meticulous alignment and continuous chin contact, potentially leading to pain with prolonged usage. Effective integration of hardware and software is crucial for real-time image processing [111]. Moreover, human anatomy, probe movement, and illumination conditions in image-based feedback may all influence the system's variability.

Ultrasonic-based TDS systems have a lot of promises for developing incredibly accurate, adaptable, and user-friendly control interfaces for those with severe impairments, despite these drawbacks. It is imperative to conduct continuous research to resolve computational and ergonomic challenges.

3.1.7 Speech-Based Tongue Drive System

Speech-Based or Micro-Electro-Mechanical Systems (MEMS) TDS are assistive technologies that utilize speech or acoustic signals captured by external microphones to control various ADs, including wheelchairs, prosthetics, smart home systems, and communication aids [112]. In contrast to systems dependent on intraoral hardware, speech-based TDS systems are non-invasive and need just microphone sensors. These systems predominantly utilize MEMS microphones, like the Knowles SPH0645LM4H-

B, ADMP421, and INMP441, owing to their compact dimensions, little power consumption, and elevated sensitivity [113]. The microphones generally have digital outputs, including I²S or Pulse Density Modulation (PDM), rendering them suitable for incorporation into embedded systems such as FPGAs or microcontrollers [114]. Their omnidirectional capability guarantees superior audio recording irrespective of head movement. The MEMS microphone produces PDM output, which is converted to Pulse Code Modulation (PCM) via a CIC filter for subsequent processing. The conversion formula given below:

$$x[n] = \sum_{k=0}^{N-1} p[k] \cdot h[n - k] \quad (3.15)$$

Where $p[k]$ is PDM data, $h[n]$ is filter coefficient, and $x[n]$ is the converted PCM signal.

The process begins when the user states a certain command, like "left," "right," or "help." Often sampled for accuracy between 16 kHz and 48 kHz, the microphone records the unprocessed voice sound. Pre-processing phases for the raw signal include noise reduction, normalizing, and sometimes silence removal. After cleaning, the audio moves to the feature extraction phase when important properties including pitch, energy, and cepstral coefficients are determined [115]. Training a classifier to link speech patterns with certain instructions requires these features. The classifier's output then operates a device, including moving a wheelchair or orienting a wheelchair to the left, so fitting for hands-free, voice-activated ADs [116]. Table 3.1 shows the comparison of different sensors, their working principle, power usages, advantages and disadvantages. A speech-based TDS offers a major non-invasive advantage as well. It's more hygienic and neater as there's no need to fit any sensors or devices into the user's mouth. Microphones are cheap when compared to other sensors as they are cheap and somewhat common [117]. For those with significant motor limitations, this approach allows the gadgets to operate without using your hands, which is crucial. Natural language communication is also encouraged, which helps customers become more at ease with technology and calls for less training.

Table 3.1 Comparison of different sensor-based TDS

TDS	Sensor	Working Principle	Invasiveness	Cost	Power usage	Advantages	Dis-advantages
Magnetic	Hall-effect and magnetometers	The magnet pierced on the tongue detects a magnetic field whenever the tongue moves.	Semi-invasive	Medium	Moderate	Robust to facial motion, high spatial resolution.	Magnets are required to be pierced into the tongue, and interference with magnets can cause the device to malfunction.
Capacitive	Capacitive grid or electrodes	Detects the capacitance charge from tongue	Invasive	Low	Low	Compact, low-power, and responsive	System is affected by oral moisture and positioning.
Electromagnetic	EM coils and field detectors	Whenever there is a movement, induced field is detected from active coils.	Invasive	High	High	No need of tongue mounted batteries	Need large components, and high interference.

TDS	Sensor	Working Principle	Invasiveness	Cost	Power usage	Advantages	Dis-advantages
Optical	IR LEDs and Photodiodes	Detectors sense tongue position-based light modulation.	Semi-invasive	Medium	Moderate	Fast detection and high accuracy	A line-of-sight is required, and saliva can affect the system.
Electromyographic	EMG surface electrodes	Detects the bioelectric signals (tongue muscle activity)	Non-invasive	High	Moderate	Natural control interface, high signal fidelity	Crosstalk from nearby muscles, electrode displacement.
Ultrasound	Ultrasonic transducers	Real-time echo imaging of tongue position or tongue contours.	Non-invasive	Very high	High	High accuracy, medical imaging	It is large, energy-intensive and sensitive to the movement.
Speech based (Acoustic)	MEMS Microphones	Processes speech commands via machine learning classifiers	Non-invasive	Very low	Very low	Implementation is compact and very low hardware cost	Not usable by non-verbal patients, performance drops in noisy environment.

3.2 EMI filters

Electromagnetic Interference (EMI) is a major problem in modern electronics, especially with high-speed systems, for example digital signal processors, MEMS microphones and communication systems [118]. EMI adds the noise to the raw signal which reduced the quality of the signal causing interference in both analog and digital systems. EMI filters are designed to eliminate or reduce the unwanted noise from the signal and transmit the desired signal for further processing [119]. EMI filters are designed of passive, active, or hybrid components depending on the intended use and frequency requirements.

3.2.1 Passive EMI filters

Passive EMI filters are commonly utilized to reduce electromagnetic interference (EMI) by allowing just the essential signal to pass, hence minimizing unwanted noise or interference. Capacitors, inductors, and resistors are passive components that enable these filters to operate independently of an external power source [120]. Passive EMI filters primarily reduce undesirable frequencies without sacrificing the quality of the intended transmission. Various passive EMI filters are available, contingent upon the frequency characteristics and component combinations employed. All-Pass Filters (APF), Band-Pass Filters (BPF), Band-Stop Filters (BSF), High-Pass Filters (HPF), and Low-Pass Filters (LPF) are the most often seen filters [121]. This section describes filters consisting of various combinations of resistors, capacitors, and inductors, along with their specific uses and designs. Passive Low-Pass Filter: Which uses the passive element for allowing only low frequency signals.

Passive High-Pass Filter: The filters which allow high frequency signals by suppressing low frequency signals. Designed using passive elements like resistance, capacitor and inductor.

Passive Band-Pass Filter: The filter is designed only using passive elements to pass only certain frequency signals.

Passive Band-Stop Filter: The filter is designed to stop only particular signals by utilizing passive elements.

Passive All-Pass Filter: Filter is designed in a way allowing all frequency signals by eliminating noise and using passive elements.

3.2.2 Active EMI filters

Usually operational amplifiers (Op-Amps), active EMI filters mix passive parts such resistors, capacitors, and inductors with active components. The key benefits of active EMI filters over passive ones are amplification, improved selectivity, and dynamic frequency response of which passive filters cannot offer [122]. Using these filters is somewhat common to lower electromagnetic noise or interference in high-performance systems. Active filters must be driven to operate from outside power sources. Active EMI filters come in numerous forms, arranged based on frequency response and interference they are meant to lower [123].

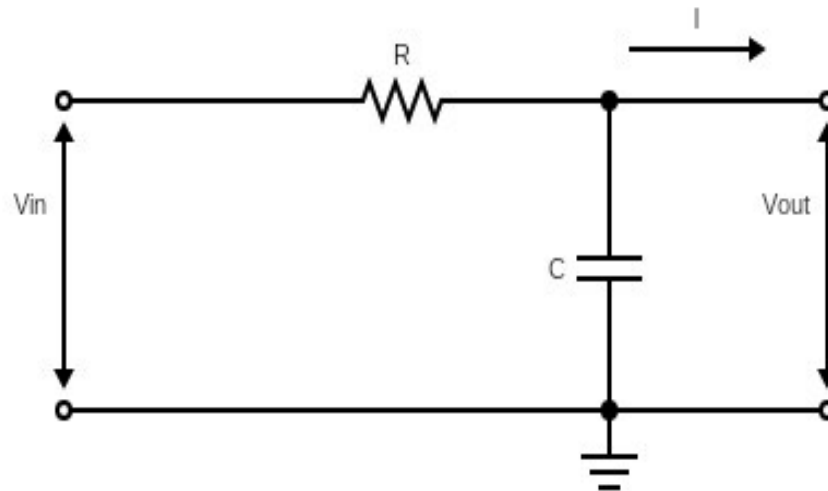


Figure 3.6 Basic Passive low-pass filters

Active filters are designed similar to passive by additionally adding active elements like power amplifiers. The types of active EMI filters are active low-pass (ALP), active high-pass (AHP), active band-pass (ABP), active band-stop (ABS), and active all-pass (AAP) designed similarly as passive filters with passive and active elements [124].

The representation of basic passive low-pass filter presented at figure 3.6, which is capacitor followed by resistance. For high-pass filter resistance followed by capacitor, which is showed in figure 3.7.

Basic passive band-pass filter is represented in figure 3.8 using passive elements to pass only certain range of frequency signals. Which has a series connection of resistance (R_1) and capacitor (C_1) and parallel connection of resistor (R_2) and Capacitor (C_2). Output is collected at the V_{out} nodes.

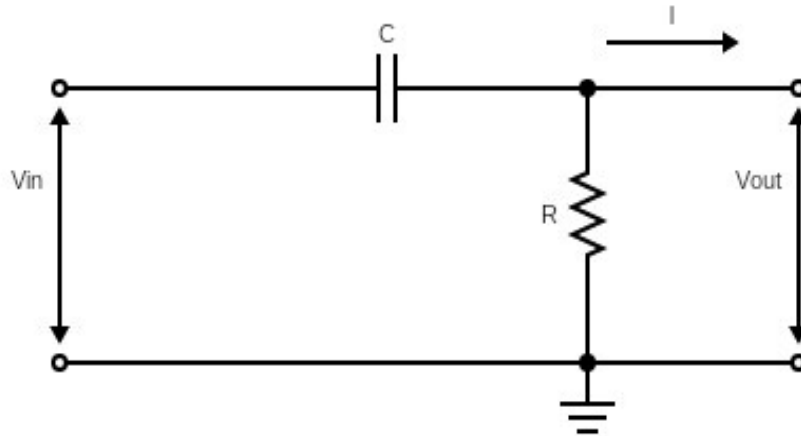


Figure 3.7 Basic passive high-pass filter

Active filters are designed using passive and active elements. Figure 3.9 shows the active low-pass filter, which uses feedback op-amp and RC low-pass connection. The output of the filter is connected at the output nodes. Figure 3.10 shows an active band-pass filter. Which is similar to the passive band-pass filter with extra op-amp.

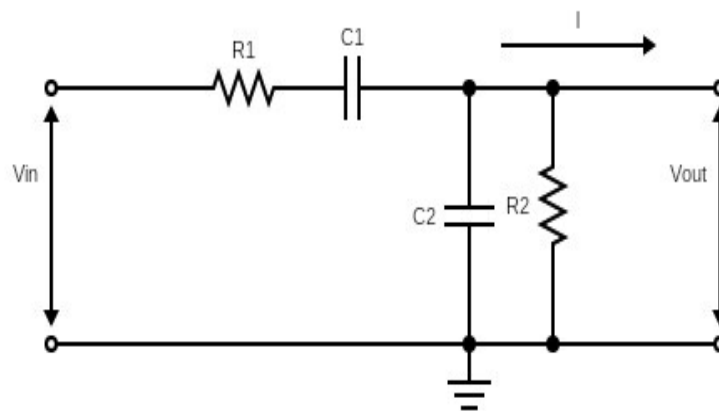


Figure 3.8 Basic passive band-pass filter

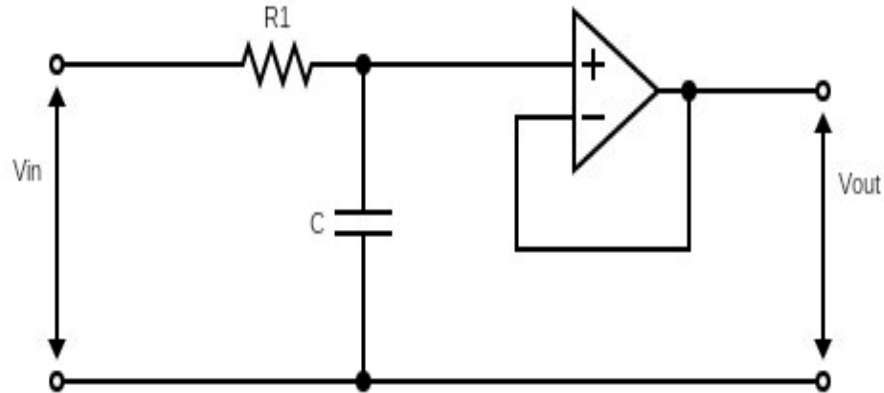


Figure 3.9 Basic active low-pass filter

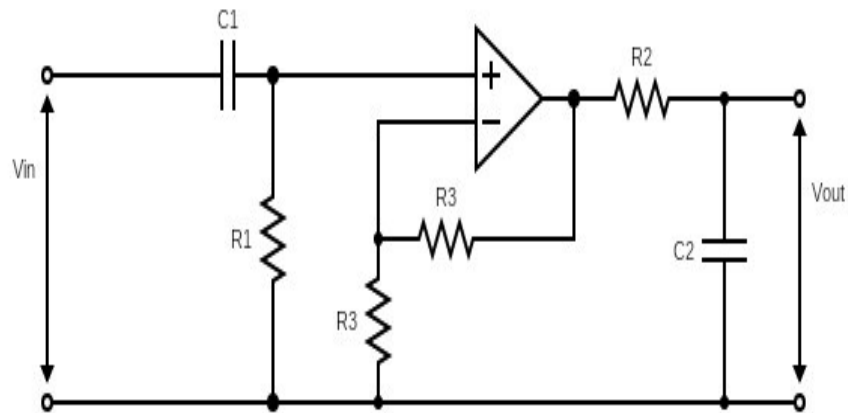


Figure 3.10 Basic active band-pass filter

3.2.3 Hybrid EMI filters

EMI is a significant obstacle in the evolution of sensitive electronic systems such as audio acquisition circuits. Integration of active and passive components allows exact control over filter properties like gain, roll-off rate, and cutoff frequency [125]. Hybrid EMI filters are applied to effectively lower EMI over a wide frequency spectrum. Among these filters are the benefits of both passive and active filtering techniques [126].

Passive devices including inductors (L), capacitors (C), and resistors (R) reduce high-frequency noise; active elements such operational amplifiers (Op-Amps) or transistors give signals that decrease lower-frequency noise. This dual approach generates a compact and effective filter able to reduce EMI over a wide range. In audio systems particularly, hybrid EMI filters are crucial as active filters struggle with high-frequency noise while passive filters are useless and heavy at lower frequencies [127]. By hybrid EMI filters, which combine many technologies, exceptional noise attenuation and signal reflection capabilities are attained across the operational frequency range. Series hybrid filters, shunt hybrid filters, feedforward hybrid filters, and feedback hybrid filters define the design and the guiding ideas for operation of hybrid EMI filters [128].

In the series hybrid filters, passive and active filtering components are coupled to offset voltage harmonics in the series. Often this arrangement calls for an active filter in series with a passive filter. They are suitable for many different power systems uses as this approach guarantees efficient harmonic filtering and voltage control [129]. Combining the advantages of both shunt active power filters (APFs) and passive filters helps a shunt hybrid EMI filter effectively reduce harmonic distortion and raise power factors. Shunt hybrid technique [130] helps to more successfully and efficiently handle certain power quality problems. The system obtains corrective current from active power circuits. Reactive power and harmonic currents are thereby canceled out. Furthermore, very efficient in handling nonlinear loads generating harmonics are shunt active power filters (SAPFs) [131]. Active series and active/passive shunt filters make the most suitable hybrid filter architecture for uses requiring a mix of voltage and current-based compensation. Two other possible hybrid filter uses are AC-motor drives and power supply [132].

Combining the benefits of active and passive filtering techniques allows feedback and feedforward hybrid EMI filter to provide a larger frequency range and maybe better noise suppression capability. Specifically at lower frequencies, it utilizes an active feedback/feedforward loop for targeted noise cancellation and a passive filter for general noise reduction [133].

3.3 Feature Extractions

Feature extraction is a fundamental operation in signal processing that converts unprocessed data into a clear and meaningful representation therefore enabling effective classification and recognition tasks [134]. This means identifying important parts or patterns in the signal that highlight significant features including energy, frequency characteristics, and temporal fluctuations. Time-domain, frequency-domain, and hybrid-domain approaches abound in the feature extractions. Essential phase in processing speech signals for a TDS, feature extraction turns unprocessed speech data into meaningful commands for the AD [135]. The TDS's MEMS microphone provides audio input that may be examined across numerous blocks to derive data necessary for system efficacy.

3.3.1 Time-Domain Feature Extraction

The signal is analyzed without conversion in time-domain. Real-time systems depend on this computationally fast approach [136]. By means of analysis of variations in the amplitude and characteristics of the signal over time, these features provide significant insights into signal patterns intimately connected to user goals [137]. Time-domain features of TDS help to identify individual voice instructions, therefore enabling AT to be managed.

Mean: The average value of the variable in the dataset and equation of mean is mentioned eq. (3.16). Where μ represents the mean of the data, N represents the number of samples and x_i is the variable value [138].

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (3.16)$$

Median: Middle value of the variable in the dataset and eq. (3.17) shows the equation of median [139].

$$\text{median}(x) = \begin{cases} X\left[\frac{n+1}{2}\right], & \text{if } n \text{ is odd} \\ \frac{X\left[\frac{n}{2}\right] + X\left[\frac{n}{2}+1\right]}{2}, & \text{if } n \text{ is even} \end{cases} \quad (3.17)$$

Standard deviation: which measures the variability or dispersion inside a dataset [140].

Zero-Crossing Rate (ZCR): In the time domain, the ZCR provides a low computational cost approximation of a signal's frequency content. Accordingly, ZCR is extensively employed in a variety of signal processing applications, such as spectral estimation, voice activity detection (VAD), voiced/unvoiced speech discrimination, speech recognition, audio & sound classification, fluid mechanics, Doppler frequency estimation, and time-delay estimation, including systems such as the TDS [141]. ZCR calculates the rate at which a signal passes the zero-amplitude threshold, or the number of times the signal changes sign (from positive to negative or vice versa) during a particular time-period.

$$ZCR = \frac{1}{T-1} \sum_{t=1}^{T-1} |sgn(t) - sgn(t-1)| \quad (3.18)$$

$$sgn(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (3.19)$$

Speech activity detection (SAD), classifying speech signals, and differentiating between voiced and unvoiced speech are all made easier with this straightforward yet powerful measure [142]. Above eq. 3.18 represents the formula of ZCR, where 'T' is length of the signal, and 'sgn(x)' is the sign function, its values are presented in eq. 3.19.

Root Mean Square: Which represents the magnitude or strength of the signals [143]. The equation of the RMS is given below- eq. 3.20.

$$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^N |x_n|^2} \quad (3.20)$$

Variance: statistical feature extraction sample variance formula is given below eq 3.21. where x_i represents individual data points, $\bar{x}$ represents mean of the data, and N is the number of data points [144].

$$S^2 = \sum (x_i - \bar{x})^2 / N - 1 \quad (3.21)$$

Covariance: Which measures the degree of two variables, which changes together. The formula is given below eq. 3.22. If the value of covariance is positive, which means the variables tend to increase together or decrease together. If the value of covariance is negative the variables move in opposite directions. The equation eq. (3.22) represents a covariance of two variables, where x_i and y_i represents individual data points, $\bar{x}$ and $\bar{y}$ represents the mean of the data, and N is number of data points [145].

$$Cov(X, Y) = \frac{\sum (x_i - \bar{x}) * (y_i - \bar{y})}{N} \quad (3.22)$$

Min and Max: The min and max values represent the minimum and maximum value of features in the dataset [146].

3.3.2 Frequency-Domain Feature Extraction

Frequency-domain extraction is essential for speech recognition systems, particularly in speech-based assistive technologies such as TDS. Speech signals are non-stationary, so their statistical features vary with time. It is considered as a quasi-stationary. The frequency domain shows pitch, formants, harmonics, and spectral envelopes [147]. By considering frequency analysis techniques on speech signal provides notable patterns corresponding to specific commands or signals.

Spectral Centroid: The spectral centroid is one of the frequency domain feature extractions in digital signal processing. Which frequently employed to denote the brightness or "liveliness" of a sound, which represents the "center of mass" or average frequency of a spectrum. The weights represent the amplitudes (or power) of each frequency component, essentially constituting the weighted average of the frequencies within a sound spectrum [148]. Equation of centroid is given below eq. (3.23), where $f(k)$ represents frequency at k and $|X(k)|$ represents magnitude spectrum.

$$Centroid = \frac{\sum_{k=0}^{N-1} f(k) \cdot |X(k)|}{\sum_{k=0}^{N-1} |X(k)|} \quad (3.23)$$

Spectral Bandwidth: In the electromagnetic spectrum, spectral bandwidth refers to the range of frequencies or wavelengths included by a signal. Typically quantified as the full width at half maximum (FWHM) of a spectral line or peak, it represents the breadth of a spectral band [149]. This bandwidth is crucial for applications like spectroscopy and radar, influencing signal resolution, data transmission, and other factors. The equation is given below:

$$Bandwidth = \sqrt{\frac{\sum_{k=0}^{N-1} (f(k) - Centroid)^2 \cdot |X(k)|}{\sum_{k=0}^{N-1} |X(k)|}} \quad (3.24)$$

Spectral Roll-off: Spectral roll-off is a technique for spectral feature extraction that identifies the frequency below which a certain percentage of the total spectral energy is contained. It is commonly employed to characterize the energy distribution of audio signals or to differentiate between voiced and unvoiced speech [150].

Spectral Flux: Spectral flux is the rate of change in the power spectrum of a signal over time, therefore quantifying the speed with which the frequency content of a signal is changing. Usually applying the L2-norm of the difference between the normalized spectrum, it is computed by comparing the power spectrum of one frame of a signal with the preceding frame [151]. Equation is mentioned in eq. (3.25), where $X_t(k)$ represents magnitude of the Fourier transform of the current frame n at k frequency.

$$Flux = \sum_{k=0}^{N-1} (|X_t(k)| - |X_{t-1}(k)|)^2 \quad (3.25)$$

Spectral Flatness: In the frequency spectrum of a signal, spectral flatness is a metric measuring uniformity of energy distribution. Low spectral flatness shows concentrated energy, typical of tonal sounds; a high spectral flatness suggests a very uniform distribution, implying a noise-like signal [152].

$$Flatness = \frac{(\prod_{k=0}^{N-1} |X(k)|)^{1/N}}{\frac{1}{N} \sum_{k=0}^{N-1} |X(k)|} \quad (3.26)$$

Mel Frequency Cepstral Coefficients: MFCCs frequently utilized in speech recognition and audio analysis, are a set of features derived from the spectral envelope of an audio signal [153]. The below equation eq. (3.27) represents MFCC formula, where $n=1, 2, \dots, N$.

$$MFCC(n) = \sum_{k=1}^K \log(E_k) \cdot \cos \left[\frac{n(k-0.5)\pi}{K} \right] \quad (3.27)$$

3.3.3 Time-Frequency feature extrication

Time-frequency feature extraction involves analyzing signals to get critical insights on the temporal development of their frequency content. The short-time Fourier transform (STFT), wavelet transform, and Hilbert-Huang transform are methodologies employed to yield representations that elucidate both temporal and spectral dimensions [154]. Typically represented as spectrograms or scalograms, these illustrations are then employed to derive features for diverse signal processing applications, pattern recognition, or classification.

Short-Time Fourier Transform: The Short-Time Fourier Transform (STFT) analyzes signals by segmenting them in the time-frequency domain and subsequently applying the Fourier Transform to each segment. This enables the observation of how the frequency content of a signal fluctuates over time, in contrast to the traditional Fourier Transform. The equation of STFT is given below eq. (3.28), where $\omega(\tau - t)$ represents a window function and $X(t, f)$ represents the time-frequency [155].

$$X(t, f) = \int_{-\infty}^{\infty} x(\tau) \cdot \omega(\tau - t) \cdot e^{-j2\pi f\tau} d\tau \quad (3.28)$$

Wavelet Transform: A wavelet transform is a mathematical approach that provides time-frequency analysis by decomposing a signal or function into several frequency components at different scales. In contrast to the Fourier transform, which examines the entire signal simultaneously, it is particularly advantageous for signals having time-varying frequency content. Wavelet transformations enable the analysis of signals by decomposing them into smaller, wave-like oscillations known as wavelets, which capture local features and discontinuities. The formula of continuous wavelet

transform is given eq. (3.29), where φ is mother wavelet and a and b represents scale and time shift [156].

$$W(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t) \cdot \varphi^* \left(\frac{t-b}{a} \right) dt \quad (3.29)$$

Discrete Wavelet Transform: The Discrete Wavelet Transform (DWT) is an effective technique for extracting time-frequency information by partitioning signals into several frequency bands across time. It facilitates the examination of the temporal variation in the signal's frequency content by providing enough resolution in both time and frequency domains [157].

$$W_{\varphi}(j, k) = \frac{1}{\sqrt{M}} \sum_k f(x) \varphi_{j,k}(x) \quad (3.30)$$

Wavelet Packet Decomposition: Wavelet Packet Decomposition (WPD) facilitates comprehensive analysis and feature selection by partitioning signals into many frequency bands, hence enabling efficient time-frequency feature extraction. WPD disrupts both the detail and approximation coefficients, in contrast to the DWT, hence providing a more comprehensive representation of the signal's frequency content. This leads to more precise feature extraction and enhanced performance in many applications, such as defect detection and signal categorization [158].

Hilbert- Huang Transform: The Hilbert-Huang Transform (HHT) is a time-frequency analysis method that uses Empirical Mode Decomposition (EMD) followed by Hilbert spectrum analysis (HSA) on each intrinsic mode function (IMF) to assess non-stationary and nonlinear signals [159]. This approach provides an extensive time-frequency representation of the signal by extracting frequency data that varies over time. HHT formula is given below:

$$Ht = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x\tau}{t-\tau} d\tau \quad (3.30)$$

3.4 Artificial intelligence and Machine learning

The artificial intelligence (AI) hierarchy arranges ideas, approaches, and technology, supporting the larger field of intelligent systems in a disciplined manner. Fundamentally, artificial intelligence is a broad spectrum of computing techniques allowing machines to replicate human-like intellect including thinking, learning, perception, and decision-making. From basic algorithms to more advanced decision-making systems, the hierarchy clarifies the interactions among several elements of artificial intelligence [160].

3.4.1 Artificial intelligence

Artificial intelligence (AI) is the general term used at the highest level for discipline. AI is any method allowing computers to replicate human intellect. This covers contemporary data-driven methods, logic programming, evolutionary algorithms, expert systems, and rule-based systems. While global artificial intelligence, a hypothetical kind of artificial intelligence that can perceive, learn, and apply intelligence generally, akin to a human being, limited artificial intelligence, which performs specialized tasks such voice recognition or chess, may be separated from one other [161]. Figure 3.11 shows the hierarchy of AI. Usually, it is divided up into many tiers: artificial intelligence, ML, deep learning (DL), and specific models or applications such computer vision (CV) and natural language processing (NLP) [162].

Under artificial intelligence, ML is a subset dedicated to algorithms and statistical models allowing computers to complete tasks without explicit programming for everyone. Rather, ML systems grow over time from data. Supervised learning, in which models are trained on labeled data; unsupervised learning, in which agents learn optimal behavior in a given environment by means of unlabeled data; and reinforcement learning, in which agents learn optimal behavior by means of rewards and penalties in a given environment, can be distinguished from ML [163]. By letting computers forecast, find anomalies, or categorize data depending on training sets, ML methods have transformed artificial intelligence.

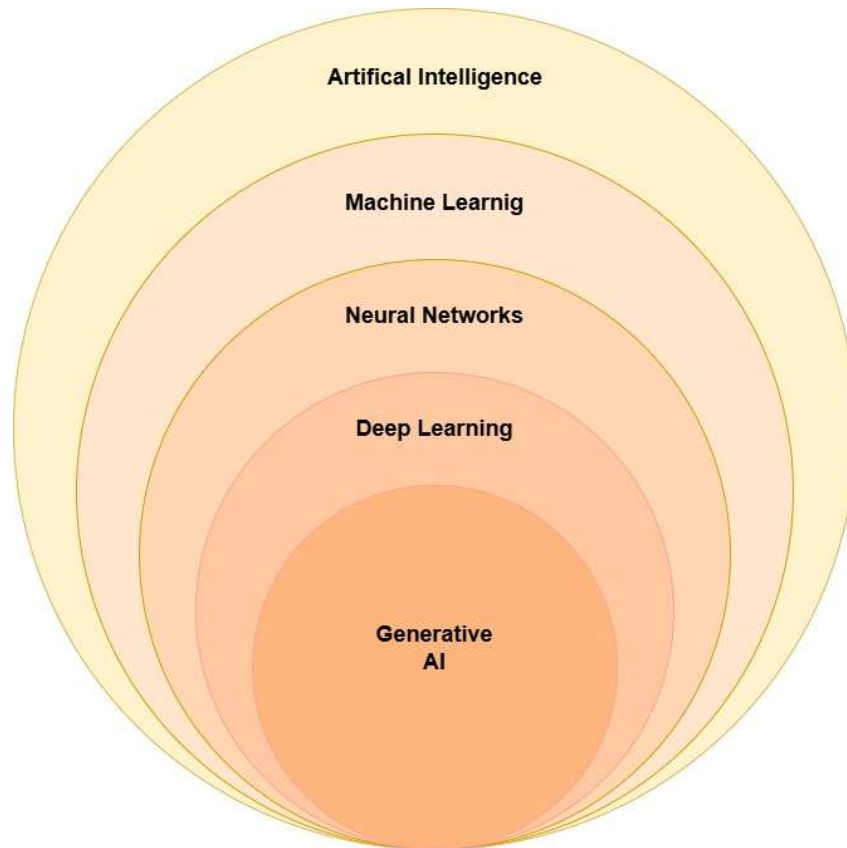


Figure 3.11. Hierarchy of Artificial Intelligence

DL is a highly specialized and potent subset of ML, uses artificial neural networks with several layers, hence the name "deep." Dealing with unstructured data including photos, audio, and natural language, DL is very successful [164]. Modern performance in tasks including face recognition, machine translation, and autonomous driving has made possible by techniques including CNNs, recurrent neural networks (RNNs), and Transformers [165]. Unlike conventional ML, DL learns hierarchical features straight from raw input rather than requiring human feature extraction.

Applications of DL and ML such CV and NLP, models that grasp syntax, semantics, and context, NLP helps robots to comprehend, interpret, and produce human language. It drives sentiment analysis, language translating, and chatbots among other things. Conversely, CV lets robots understand visual data from the environment, execute tasks such scene segmentation, picture categorization, and object recognition [166].

Depending on their application needs, training a rudimentary ML model for predictive maintenance or implementing a whole deep learning pipeline for autonomous navigation, this framework lets researchers and developers concentrate on layers [167]. Making wise choices about data needs, computing resources, and technological stacks in every AI-driven project depends on an awareness of this hierarchy.

3.4.2 Machine Learning:

ML is primarily categorized into three forms: Supervised Learning, Unsupervised Learning, and Reinforcement Learning, each designed to address distinct types of issues using various methodologies. In supervised learning, the model is trained using a labeled dataset where the input data is associated with the correct output [168]. This kind is ideal for tasks requiring prediction and classification. The prevalent techniques employed are Linear Regression, Logistic Regression, SVM, Decision Trees, Random Forests (RF), and k-Nearest Neighbors (k-NN). These algorithms are frequently utilized in spam detection, fraud prevention, and medical diagnosis, learning the correlation between inputs and outputs [169].

In contrast, unsupervised learning operates on data devoid of labeled outputs. The purpose is to analyze data structures, discern trends, and cluster associated data points. It is particularly beneficial in exploration data analysis, customer segmentation, and anomaly identification. Figure 3.12 describes the types and algorithms included in ML. Key techniques in this category include independent component analysis (ICA), principal component analysis (PCA), k-Means clustering, and hierarchical clustering [170]. Frequently employed as a first phase for supervised learning to get deeper insights into complex datasets, these algorithms identify underlying connections within the data.

Reinforcement learning (RL) is a significant category of ML wherein an agent acquires decision-making skills through interaction with an environment [171]. The agent receives rewards or penalties based on its conduct, adjusting its strategy to optimize overall reward over time. This strategy, informed by behavioral psychology, is applicable in dynamic environments like robotics, gaming, and autonomous vehicles.

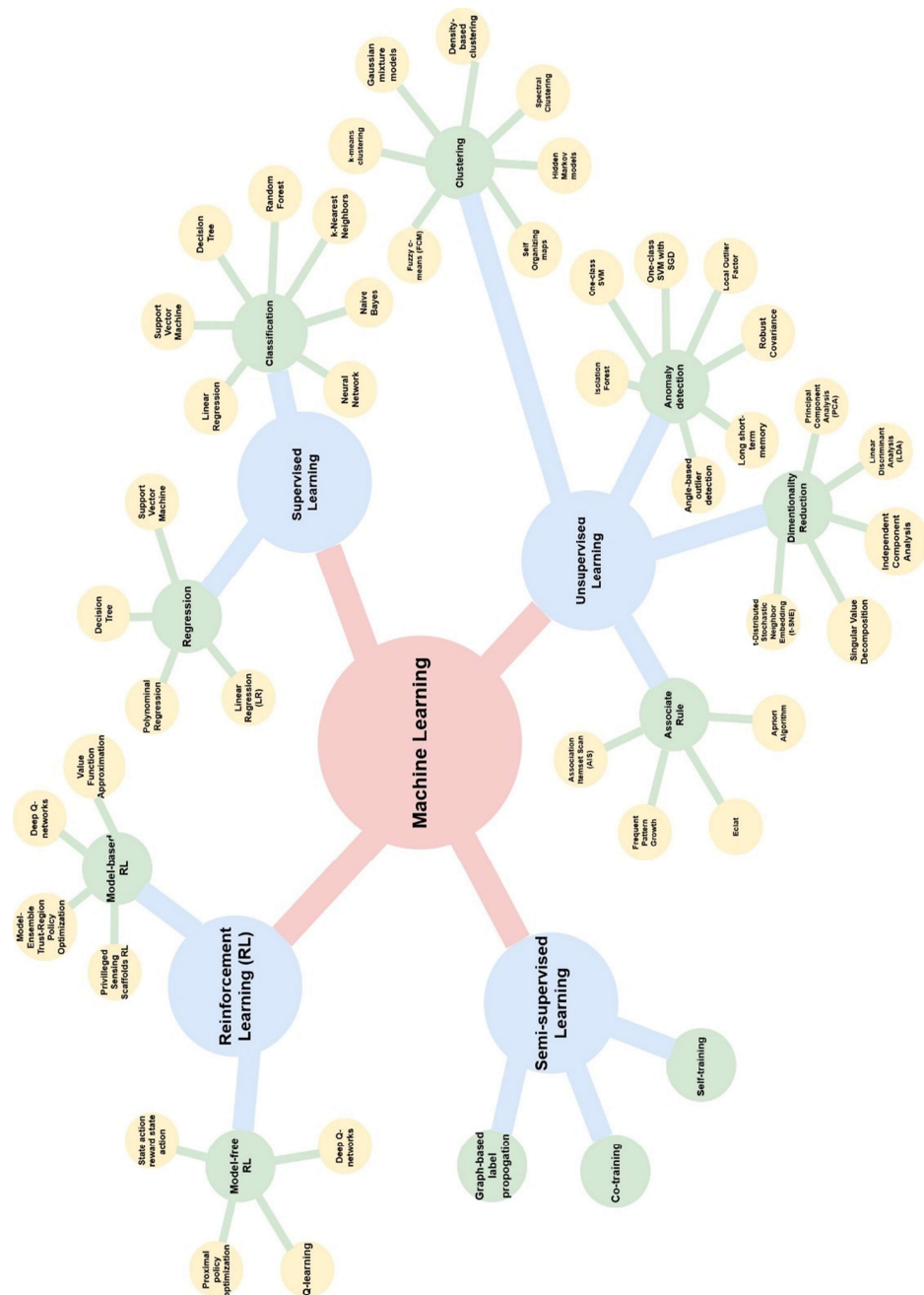


Figure 3.12 Machine Learning Types and algorithms

Prevalent techniques in reinforcement learning include Policy Gradient Methods, Temporal Difference (TD) Learning, Deep Q-Networks (DQN), and Q-Learning [172]. Reinforcement learning is highly successful when a model has to acquire sequences of behaviors that result in long-term success rather than immediate rewards.

Chapter 4

Design and Implementation of Tongue Drive System on FPGA Processor

Advancements in HMI help people with physical challenges regain their freedom. This is possible because of the growing need for AT in healthcare, therapy, and personal mobility. For those with severe SPI or quadriplegia, the TDS provides a hands-free interface. Unlike BCI or eye-tracking systems, which may be intrusive or need continual visual attention, TDS employs tongue motions or signals using sensors like magnetic or speech-based to operate equipment like motorized wheelchairs, mobiles, computers or smart homes, etc. The tongue, controlled by the hypoglossal nerve, is the least effective area in the human body from spinal damage, or accidents. Which offers a dependable and voluntary means for disabled to engage with their surroundings.

Although traditional TDS systems utilize magnetic tracking with an intraoral tongue barbell and multiple magnetic sensors affixed on headgear, this approach is intrusive, expensive, and impracticable for real-time, low-latency control in portable devices. On the other hand, speech-based TDS systems have drawn interest. These systems operate using pre-defined voice commands as control inputs, providing a non-invasive, simple solution that supports the user's cognitive control.

Real-time speech recognition TDS systems rely on the parallel processing design, dependable timing, low-latency operations, and programmable logic capabilities of FPGAs. In contrast to general-purpose CPUs or conventional microcontrollers, FPGAs enable hardware to perform the whole signal processing pipeline, encompassing speech classification, motor control, and audio input, for enhancing power efficiency and accuracy.

The proposed TDS system is implemented on FPGA board- Nexys 4 Artix 7, which has MEMS omni-direction microphone mounted for collecting the PDM audio signals. Commands are processed through blocks of noise filtering, feature extraction, and ML for the activating AD. For actuating the wheelchair L298N motor driver is

used. Two DC motors are connected to the front wheels of the wheelchair. FPGA board will generate the PWM signals according to the recognized command.

The system is deployed on the Xilinx Nexys 4 DDR Artix-7 FPGA, selected for its integrated peripherals, sufficient logic resources, and appropriateness for high-speed data processing. A PDM-to-PCM audio converter, a digital EMI filter, a feature extractor based on the Mel-Frequency Cepstral Coefficient (MFCC), a Support Vector Machine (SVM) classifier for command recognition, and a PWM generator for motor control are among the modules that make up the prototype. The Xilinx Vivado tool is used for synthesis once each module is designed from scratch in Verilog, tested, and validated with test benches.

4.1 Implementation of speech-based TDS

The speech-based TDS system is energy-efficient AD designed to convert spoken commands into directional control signals for a wheelchair. The Xilinx Nexys 4 DDR FPGA board, with an Artix-7 FPGA processor is used, which prioritizes real-time computation, low latency, and hardware-level parallelism. The fundamental processing pipeline is constructed using Verilog HDL modules derived from MATLAB models to provide reliable algorithms and efficient hardware performance.

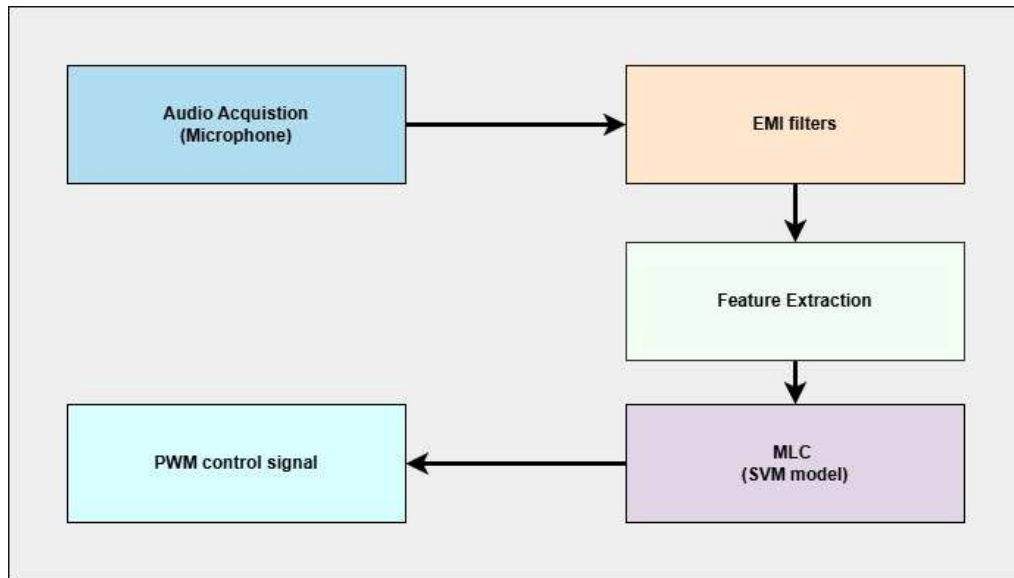


Figure 4.1 Block diagram of speech-based TDS design

The modules which included for designing TDS are audio acquisition, EMI filter, feature extraction, SVM classification for command recognition and wheelchair actuation. These are arranged in sequence, facilitating the conversion of raw audio into motor control signals. Figure 4.1 illustrates the TDS block diagram of speech-based commands. This design enables debugging, and hardware optimization without impacting the remainder of the system.

4.2 Audio Acquisition and preprocessing

The TDS system for wheelchair operation employs a MEMS-based digital microphone ADMP421 integrated on the Xilinx Nexys 4 DDR FPGA board for audio input acquisition. The microphone provides superior real-time audio input in Pulse Density Modulation (PDM) format, which the system analyzes and converts into usable Pulse Code Modulation (PCM) data for subsequent analysis. This section delineates the whole pipeline for audio capture and preprocessing, encompassing hardware interface, PDM-to-PCM conversion via a CIC filter, framing, windowing, DC offset elimination, and amplitude normalization.

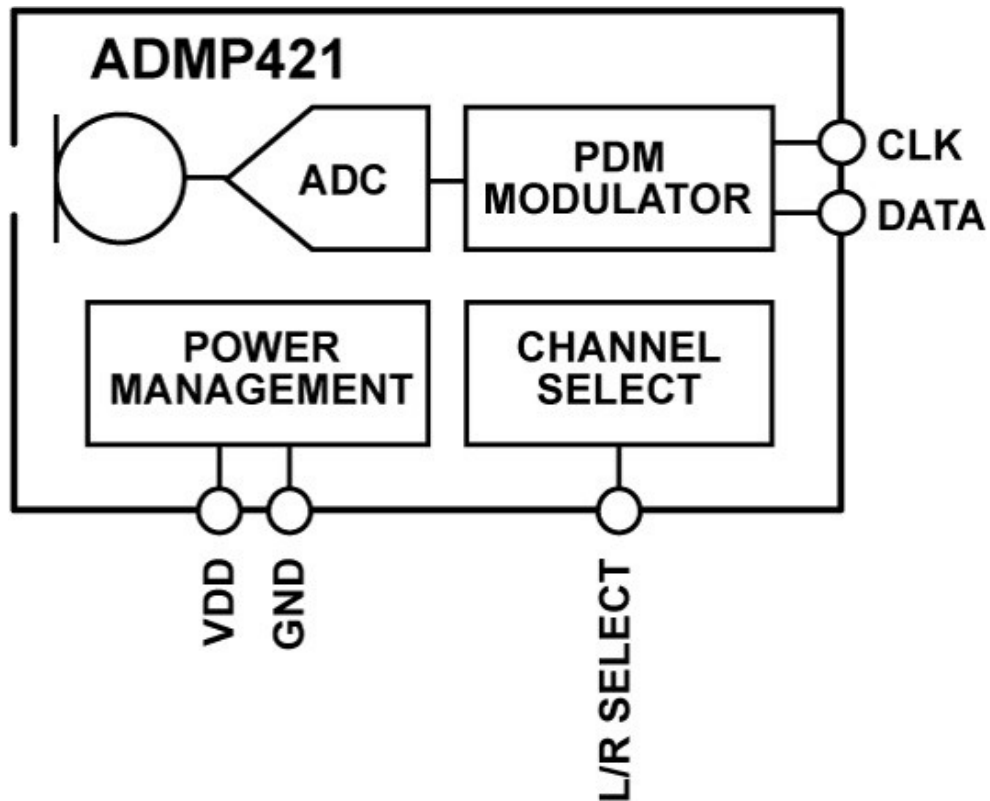


Figure 4.2: MEMS omnidirectional ADMP421 Microphone [173]

4.2.1 ADMP421 microphone

The ADMP421 produces a 1-bit PDM signal as a high-performance digital-output MEMS microphone. The FPGA board interfaces with it by three pins: micData, micClk, and micLRSel. Figure 4.2 illustrates that the micData pin transmits the digital audio stream, micClk synchronizes the sampling process, and micLRSel selects between stereo channels, with just the left channel utilized in this system for mono audio capture. The direct processing of the PDM signal is impeded by its high sample rate. Therefore, decimation and low-pass filtering restore the original analog waveform as a PCM stream. This figure shows the connection between the ADMP421 and the FPGA board. The mic outputs PDM data at a high frequency ~ 3.072 MHz, synchronized via micClk. The micLRSel is fixed to capture only one audio channel.

4.2.2 PDM to PCM Conversion using CIC filter

A Cascaded Integrator-Comb (CIC) filter is employed in Verilog to diminish the sample rate of the PDM stream and derive a pristine audio waveform. CIC filters are advantageous to hardware implementation, involving only adders and delay registers, as a result offering them appropriate for FPGA-based architectures. The filter comprises an integrator section, a down-sampler (decimator), and a comb section. Initial validation occurs in MATLAB to assess decimation precision and signal smoothness, followed by synthesis and implementation via Vivado. The Integrated Logic Analyzer (ILA) facilitates real-time visualization of the PDM and resultant PCM signals. Figure 4.3 describes the first stage of CIC filter. The output we need is 16-bit PCM signal at 48 kHz.

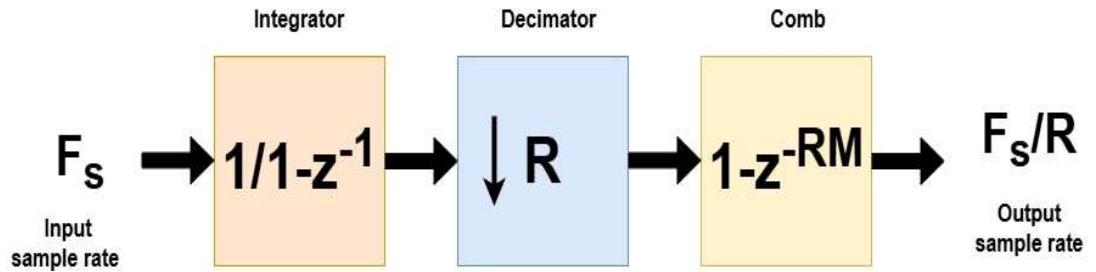


Figure 4.3: single stage CIC filter [179]

4.2.3 Signal Framing and windowing

Speech is a non-stationary signal; it is divided into brief frames when the signal may be regarded as quasi-stationary. In this architecture, each audio frame has 25 milliseconds (ms) of data, with a 10 ms gap between consecutive frames. The overlap guarantees continuity and mitigates information loss at the peripheries of each frame. A Hamming window is utilized for each frame to mitigate edge artifacts prior to transformation into the frequency domain. The window reduces spectral leakage by narrowing the edges of the frame. The Hamming window function mitigates edge discontinuities in each frame, enhancing the outcomes of spectral analysis. This phase is essential for enhancing the precision of features such as Mel Frequency Cepstral Coefficients (MFCC) or Linear Predictive Coding (LPC), which subsequently occur in the processing pipeline.

4.2.4 DC offset removal and amplitude normalization

The recorded signal may have a DC bias due to external noise or technical flaws in the microphone. This bias distorts feature distributions and must be eliminated. The DC offset is eliminated by computing the mean of each frame and subtracting it from every sample. Subsequent to offset correction, the signal is subjected to amplitude normalization to guarantee consistent loudness across samples. This procedure normalizes the frame's peak amplitude, enhancing the classifier's generalization. Fixed-point arithmetic is used for both tasks to work with the FPGA's limits, and saturation logic is used to handle overflows during division.

4.3 EMI Filter Design

In home appliances and systems, electromagnetic interface (EMI) can cause short circuits or malfunctioning. Appliances not sufficiently protected from power run the danger of permanent harm. These cases show how EMI might harm household gadgets [174]. EMI can cause delays in performance, irreversible damage to equipment or electrical circuits, or data loss affecting many different sectors. Among the main issues EMI can bring about in the industrial sector are the following: In the military and medical domains, EMI can much more drastically compromise mission-critical operations [175]. In electronic equipment, a powerful electromagnetic wave can produce undesired electric currents that can cause inadvertent operations or interference

with desired operations. If an electronic gadget absorbs too much outside energy, it might be destroyed. Even if the energy provided from the outside is somewhat tiny, at places where radio waves for broadcasting and communication are weak it might produce interference with reception, aberrant sound, or video distortion. Electromagnetic noise interference is the name given to an electromagnetic wave that causes interference; electromagnetic noise [176] is the name given to an electromagnetic wave that causes interference. Noise signals disturb wireless devices. EMI can influence cell phones and other mobile devices, therefore affecting their reception or performance. Devices that link to Bluetooth or the Internet could have dependability issues in daily use [177]. Small disruptions brought on by EMI to computers run the risk of causing poor performance or monitor flickering. On the other hand, a significant disturbance like a power surge could cause a system crash and irreversible hard disk data loss [178].

4.3.1 Passive EMI

Passive EMI filters are commonly employed by utilizing passive components. To mitigate system-generated EMI, several passive filters are designed to eliminate EMI. Discrete passive EMI filters, predicated on the notion of insertion loss, are the most often employed passive EMI filters for mitigating EMI [179]. The EMI equivalent circuit has been designed using seven inductors (L), five resistors (R), and five capacitors (C) to lower the noise signal. The filter consists of three terminals and two feedback capacitors. Figure 4.4 illustrates the passive EMI filter.

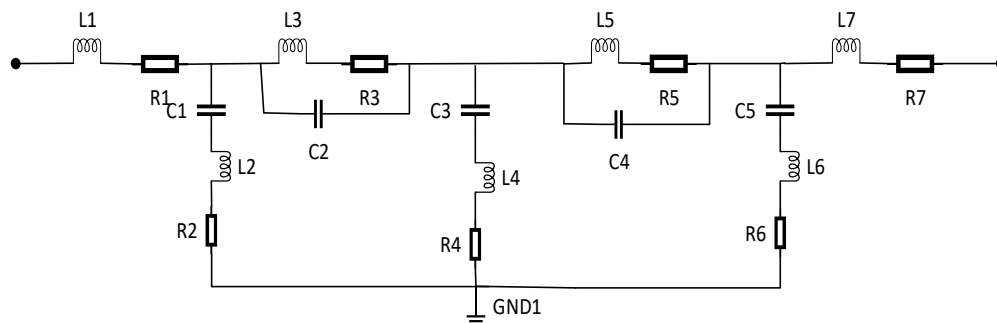


Figure 4.4 Passive EMI filter

The cut-off frequency for filters is determined by the inverse of load resistance and components, as indicated in equation (4.1). The relationship between the phase of the input and output signals of the EMI filter is termed phase response, determined using the phase shift equation (eq. 4.2) and the amplitude response of the filter described in frequency bands as indicated in eq. (4.3), where $\omega=2\pi$.

$$f_{cutoff} = \frac{1}{2\pi RC} \quad (4.1)$$

$$\phi(\omega) = -\tan^{-1}\left(\frac{\omega}{\omega_0}\right) \quad (4.2)$$

$$G(\omega) = |H(e^{j\omega T})| \quad (4.3)$$

4.3.2 Active EMI filter

Active filters utilizing active transistor-based operational amplifiers and passive components deliver output within the specified bandwidth range. Active filters constructed with higher-order designs offer a pronounced roll-off with flat passbands or stopbands, in contrast to the fundamental design of active filters. These higher-order filters are integrated into System on Chips (SoCs) or other Application-Specific Integrated Circuits (ASICs) based on their uses. Active filters are designed in common or differential mode to attenuate noise in certain circuits.

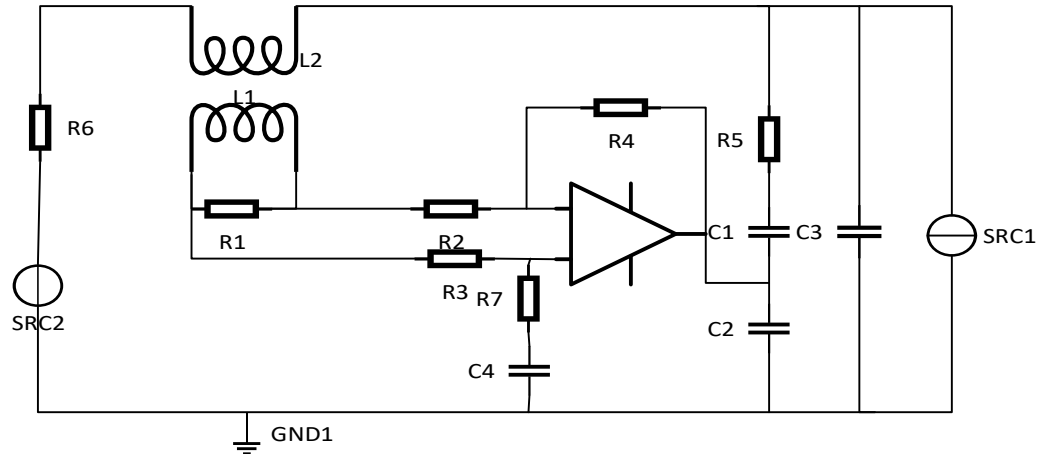


Figure 4.5 Active EMI filter circuit

The filter comprises an operational amplifier, a common-mode choke, and a pair of capacitors interconnected. The reactive components are incorporated into the feedback loop to alter the amplifier's transfer function, hence ensuring the stability of the active EMI filter circuit. The active EMI filter is presented in Figure 4.5. the values of R, L and C values chosen such that which satisfies the cut-off frequency.

4.4 Machine Learning Classifiers

Machine Learning Classifier (MLC) are critical in TDS when speech is used as a command modality gathered from a microphone. Effective classification algorithms are necessary to accurately differentiate between intended commands, as speech signals are highly variable due to variations in intonation, speed, accents, and background noise. Classifiers can map speech input to tongue-drive control commands by extracting feature extractions. Enabling individuals with severe motor disabilities to interact with their environment through speech-based control, this guarantees the robust and accurate operation of TDS for assistive applications. The implementation of classifiers not only improves user safety and usability, but also reduces false activations, thereby increasing system accuracy.

The development of MLCs involves a sequence of critical stages. The first step is to collect data, which includes recording speech commands. After that, the data is processed to remove noise, normalize the values, and extract important features for classification. After that, the data is divided into training and testing sets so that the system can be tested on one set and learn from the other. It is critical to select the appropriate features and method. The model is evaluated using metrics after training to optimize its performance. These steps guarantee the classifier's reliability and efficiency.

4.4.1 Data Collections

A custom speech dataset was created for this research, which includes eight command words: back, call, down, help, left, right, straight, and up. A total of 175 participants, both male and female speakers from different Indian states contributed to the dataset. A deliberate effort was made to ensure that the dataset accurately represents the real-world speech variations by including a variety of regional accents, genders, and voice characteristics.

The dataset was generated by 175 participants who recorded eight distinct audio clips, each audio clip contained a speech command word. These recordings were collected over WhatsApp, which provides a useful and convenient means of collecting information from people who are in other places. The dataset contains 1,400 audio recordings with a range of accents and genders that are evenly split over eight command types. Because of the different accents, the dataset is ideal for training MLCs that work well with new speakers in practical settings.

4.4.2 Dataset Organization and Labeling

The dataset was systematically organized for machine learning models to classify the speech samples. Each of the 175 participants provided eight audio clips, one for each of the following predetermined command words: back, call, down, help, left, right, straight, and up. The audio files were organized into folders according to the corresponding command, which made data loading easier and ensured consistency during evaluation and training.

Audio samples were labeled with the corresponding command words to facilitate supervised learning. Accurate mapping of input audio signals to their corresponding commands was made possible by these labels, which were essential for the classifier's training. Maintaining class balance is essential in dataset preparation, as imbalanced datasets might skew classifiers towards majority classes, hence diminishing accuracy for underrepresented categories. With each participant providing one recording of every command, total of 1,400 audio samples uniformly dispersed across eight classes. This equitable framework enables the classifier to adapt all commands uniformly, which improves the performance.

4.4.3 Data Preprocessing using AUDACITY software

The background noise, amplitude, and duration of the raw audio recordings were subject to variability due to their collection in a variety of environments using mobile devices. In order to prepare high-quality datasets for machine learning, it is essential to have tools that can trim, normalize, reduce noise, and convert formats. Audacity provides these tools. To address these inconsistencies, preprocessing was carried out using AUDACITY software [180] an open-source, free digital audio recording and editing program, widely used in research and professional audio processing.

The audio files are undergone preprocessing using AUDACITY software before extracting features for classifying the speech commands. Preprocessing steps applied to the audio files in AUDACITY mentioned below:

- ❖ Tempo Adjustment: The speed of the audio clips of command words was uniformly standardized across all the audio clips without modifying the pitch.
- ❖ Normalization: To ensure the uniform loudness across all the audio clips, amplitude levels are normalized.
- ❖ Trimming and resizing: Starting and ending silences were eliminated, and all clips were standardized in duration and size for consistency.
- ❖ Format Conversion: All the audio files are changed to .wav format, by employing mono channel with a 48000 Hz sampling rate and signed 16-bit PCM encoding, which is standard for speech classification.

The quality of the dataset was significantly improved by applying above preprocessing procedures. In order to provide proper feature extraction for training the classifier, standardizing all data to the same format and parameters, and reducing unnecessary variability is must.

4.4.4 Feature Extraction for Machine Learning Classifier

Feature extraction was essential in transforming the high-dimensional and speaker-dependent characteristics of raw audio signals into concise, significant representations. In this thesis selected Mel-Frequency Cepstral Coefficients (MFCCs) as the principal features because they effectively represent the spectral characteristics of sound in a manner similar to the human auditory system.

The procedure commenced by segmenting each audio stream into overlapping frames to examine short-term speech attributes. A Fast Fourier Transform (FFT) was utilized to transmute these frames into the frequency domain. Frequencies were subsequently mapped onto the Mel scale to represent perceptual significance, followed by a logarithmic adjustment to simulate human loudness perception. A Discrete Cosine Transform (DCT) was ultimately employed to produce decorrelate coefficients. The resultant feature vector comprised the starting 13 MFCCs, energy, together with their first and second derivatives, for establishing a robust and consistent feature matrix for each audio clip. This methodology guaranteed that the derived characteristics were

resilient to noise and invariant to speakers, rendering them appropriate for classifier training.

4.4.5 Implementation of Machine Learning Classifiers

MLCs are essential to the TDS, enabling it to read spoken commands. MLCs analyze the characteristics of speech inputs, guaranteeing that each spoken command is precisely aligned with a predetermined control output. MLCs serve as the decision-making element of the system, connecting speech input to control functionality.

In this thesis, I designed and compared of four classifiers: Naïve Bayes (NB), Random Forest (RF), Stochastic Gradient Descent (SGD), and Support Vector Machine (SVM) to evaluate their efficacy in identifying speech. All classifiers were trained and tested on the same dataset, and calculated the performance metrics of the classifier accuracy, precision, recall, specificity, F1-score and Receiver Operator Characteristic (ROC). This comparison analysis helped me to identify best classifier for the TDS, balancing classification accuracy and computational efficiency.

Naive Bayes is a simple yet effective probabilistic classifier, works on the Bayes theorem. Which calculates the likelihood of each class. For many real-world datasets, NB performs well while simplifying computations by assuming that characteristics are conditionally independent. NB requires little processing resources, trains fast, and works particularly well with high-dimensional data. NB frequently produces good results in jobs like text categorization, spam detection, and medical diagnosis despite its simplicity.

Support Vector Machine (SVM) is an effective classification tool, identifies the separating optimal hyperplane of different classes in a high-dimensional feature space. SVM performs exceptionally well for both binary and multi-class classification, especially in situations when classes cannot be separated linearly. Even in difficult acoustic environments, SVM can effectively recognize tongue instructions from speech signals by handling non-linear patterns using various kernel functions, such as linear, polynomial, or RBF. It is a great option for applications demanding high accuracy and robustness because of its powerful generalization capabilities.

Random Forest (RF) is a flexible machine learning technique, constructs several decision trees (DT) and aggregates their predictions to increase accuracy and decrease overfitting. It provides insightful information on the significance of features and

performs exceptionally well when dealing with complicated, noisy, or correlated data. It is widely utilized in domains like financial forecasting, bioinformatics, and image classification since it is effective and useful.

Stochastic Gradient Descent (SGD) is one of the popular optimization technique for training linear classifiers. It is scalable to big datasets and computationally economical because it updates the model parameters incrementally for every training sample. SGD is perfect for online learning or situations needing frequent model updates because of its flexibility with new dataset.

Implementation of classifiers workflow starts with data preparation, normalization, feature extraction, and dataset division into training and testing sets. Every classifier had a distinct strategy: RF classifier is a combination of decision trees with majority voting, SGD iteratively updated a linear classifier using stochastic gradient optimization, NB used probabilistic correlations between features and classes, and SVM used linear and RBF kernels to find optimal hyperplanes. Accuracy, precision, specificity, recall, and F1-score metrics are calculated for each classifier on the test set. SVM achieved better performance than the other classifiers, demonstrating robustness to non-linear correlations and high generalization.

4.4.6 Multi-Class Support Vector Machine classifier

The SVM is a binary classifier that uses an ideal hyperplane to divide data into two classes. For multi-class SVM classifier, two approaches are present one is one-vs-all (OvA) and other is one-vs-one (OvO). The multi-class classifier of OvA trains a different SVM for each class, treating the positive class as the one that is positive and the others as negative. In contrast, OvO generates several classifiers for big datasets by training an SVM for each pair of classes. The one-vs-all multi-class SVM classifier is used for our implementation due to its ease of use and reduced processing requirements, particularly for datasets with a reasonable number of classes. Normalized characteristics were used to train each OvA classifier. With this method, SVM can effectively handle multiclass issues while retaining its excellent accuracy and generalization.

Multi-class SVM classifiers classify audio signals according to eight pre-defined categories, recognized as classes. Systems of multi-class classifiers are commonly employed in the recognition of speech commands, the detection of audio,

and the classification of music genres. The SVM classifier is designed for eight classes, 1400 audio clips are used to design and test the SVM multiclass classifier. The classifier has 8 words, which means eight classes. The predicted and true values of classes are mentioned in the matrix. 20 clips are given as testing samples/ dataset to the SVM module.

All the audio clips are adjusted and saved in .wav format using Audacity software. Standardizing the audio format ensures consistency across all data, making it easier to process and analyze. With all clips at the same frequency and duration, it reduces variability that could affect model training and classification accuracy. Additionally, using a common format like .wav guarantees compatibility with various audio processing tools and libraries.

4.5 PWM signal generation

The speech-controlled wheelchair employs a sophisticated system architecture orchestrated by a Nexys 4 Artix 7 FPGA. The process begins with Signal Acquisition (Speech Input), where the user's speech commands are captured by on board microphone. To ensure signal integrity, EMI Filters mitigate electromagnetic interference, providing a clean audio signal to feature extraction module. Command and their action

Table 4.1 Command and their action

Command	Action
back	Move backward
call	Move forward and Activate buzzer (optional)
down	Decrease speed
help	Stop and Activate buzzer (optional)
left	Turn left
right	Turn right
straight	Move forward
up	Increase speed

Feature Extraction algorithms analyze the filtered speech, identifying key acoustic characteristics that distinguish different commands. These extracted features are then fed into the MLC (SVM Model) block, which implements a Support Vector Machine to classify the spoken commands. The SVM model translates the speech patterns into specific control actions, such as forward, backward, left, or right. Table 4.1 explains the command and their actions to be done.

The PWM signals are generated by PWM block to control the motor speed and direction of L298N motor driver. The output of the FPGA board is PWM signals, which are fed to motor driver L298N. These PWM signals drive the H-Bridge Motor Driver - L298N, which regulates the power delivered to the DC Motors (Left and Right Wheels), thereby controlling the movement of the wheelchair.

Table 4.2 Control signal generation according to the commands

Command	Action	Motor 1 (Left Wheel)	Motor 2 (Right Wheel)	Speed (PWM)
back	Move Backward	Reverse (0,1)	Reverse (0,1)	Medium/High
call	Move Forward & Activate Buzzer	Forward (1,0)	Forward (1,0)	Medium/High
down	Decrease Speed	No change	No change	Reduce PWM
help	Stop & Activate Buzzer	Stop (0,0)	Stop (0,0)	0 (stop)
left	Turn Left	Stop (0,0)	Forward (1,0)	Medium
right	Turn Right	Forward (1,0)	Stop (0,0)	Medium
straight	Move Forward	Forward (1,0)	Forward (1,0)	High
up	Increase Speed	No change	No change	Increase PWM

Table 4.2 shows the control signal generation for each DC motor generate for each command. The system is powered by separate power supplies to accommodate the varying voltage and current requirements of the FPGA, motor driver, and DC motors. Secure electrical connections are maintained using jumper wires and terminal blocks to ensure stable performance.

Feature Extraction analyzes the filtered signal to identify key characteristics necessary for speech recognition. The MLC (SVM model) block utilizes a Support Vector Machine to classify the extracted features, translating them into control commands. Finally, Decision Logic and PWM Control Signal generates appropriate Pulse Width Modulation signals to drive the H-Bridge Motor Driver (L298N), which regulates power to the DC Motors that control the wheelchair's movement.

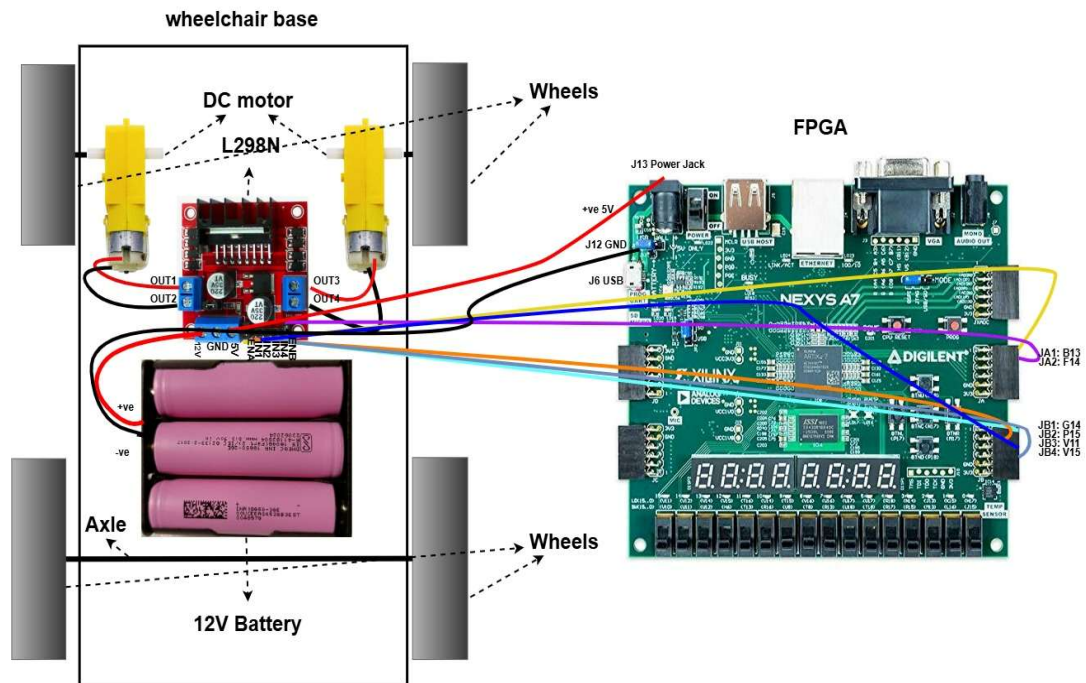


Figure 4.6 Wire connections of FPGA and L298N motor driver to two DC motors of wheelchair

The above figure explains (Figure 4.6) the connections between FPGA ports to H-bridge motor driver to DC motor and power supply. The block diagram of a speech-

controlled system for a wheelchair, employing an FPGA for signal processing and motor control. Initially, the Signal Acquisition block captures speech input, which is then filtered by EMI Filters to reduce electromagnetic interference. Then extract the features for SVM classifier, the classified command then converts to PWM signal for driving the motor. Motor driver gives the control signals to DC motor to speed and moving the wheels.

Chapter 5

Collaborative Case Studies in Biomedical Signal and Image Processing

5.1 Implementation of an Epileptic seizure detection

Epilepsy is one of the most common chronic disorders in the field of neurology, as it affects millions of people every year [181]. Epileptic seizures occur when an abnormal electrical discharge interrupts the brain's normal functioning. These events typically last from a few seconds to several minutes [182]. An individual with this condition will experience loss of awareness, changes in emotions, lack of control of their muscles, and shaking. A seizure could be caused by a drug, a high fever, an injury to the head, or a disease such as epilepsy [183]. Worldwide, there are about 50 million people with epilepsy, and most of them live in countries with low or middle incomes [184]. There is also a small risk of sudden unexpected death among epilepsy patients. It is estimated that about 1% of epilepsy patients die from Sudden unexpected death in epilepsy (SUDEP) [185, 186]. Therefore, tracking or monitoring brain activity might be helpful in identifying seizures [187]. An effective method of monitoring brain activity is collecting electroencephalography (EEG) signals, and these bio-signal waveforms are obtained using electrodes placed on the scalp [188].

Status epilepticus needs to be diagnosed properly and managed accordingly, so EEGs with monitoring should be needed 24 hours a day [189]. The traditional method of diagnosing epilepsy relies on the expert's visual inspection of lengthy EEG recordings. As there are no clear differences between epileptic and non-epileptic seizures in EEG activity, the process is time-consuming and problematic [190]. Manual scanning might provide inaccurate results as the observations and experience of the engaged neurologists affect the outcomes [191]. Reviewing and evaluating the EEG signal for epileptic seizures calls for a skilled specialist; it takes time. Therefore, developing automated seizure detection is helpful to identify seizures without manual inspection and consumes less time compared to traditional analyzing.

In artificial intelligence, deep learning (DL) and machine learning (ML) play a crucial role in this automated era. Particularly in the medical field for diagnosing disease, analyzing tumors in various body parts, and predicting abnormalities, etc. In the healthcare field, ML techniques have been used to diagnose diabetes disease [192], breast cancer [193], type-2 diabetes [194], thyroid disease [195], heart disease [196], neurological disorders like Alzheimer's, Parkinson's, schizophrenia, epilepsy, and more [197]. For detecting and analyzing ML classifiers use medical images, bio-signals, and EEG signals, etc. Artificial intelligence (AI) cannot replace human physicians, but it can produce better solutions to problems that must be addressed, and it can be used to ease the burden of physicians [198]. Thus, machine learning could benefit patients by lowering prices, increasing accuracy, or disseminating experience that is in short supply [199].

In order to automate epilepsy seizure classification, feature extraction techniques are required [200]. Classification accuracy is strongly influenced by feature extraction quality. When it comes to implantable medical devices it is a necessity to select good and high-quality features to ensure that they will function effectively. Extracting the features enhances the system's computational complexity, resilience and distinguish data [201].

The performance measurements of classifiers are determined by several parameters, including dimensionality and pattern class labels [202]. Among all possible classifiers, the Support Vector Machine (SVM) is one of the superior options. One reason SVM [203] demonstrates enhanced accuracy in classification is feature extraction. FPGA-based systems are attractive for applications in image and signal processing. Moreover, FPGA-based systems provide scalability and adaptability. They can be easily adapted for new applications and may be enhanced from existing systems. Consequently, FPGA-based computers represent a cost-effective option for applications such as image and signal processing.

The implementation of Epileptic Seizure Detection (ESD) using MATLAB HDL coder. The module of ESD is designed and analyzed in MATLAB. Fig. 1 shows the block diagram of detection of seizures. Which uses 1/8 compression ratio for compressing the EEG signal. So that the data can be compressed and processed easily

without losing data. Figure 5.1 shows the block diagram of epilepsy detection for software analysis designed in MATLAB.

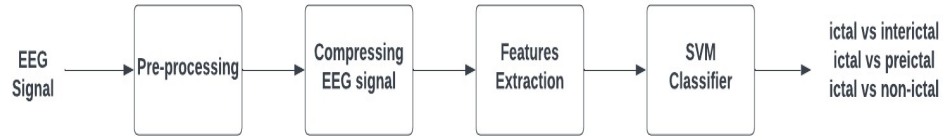


Figure 5.1 Block diagram of epileptic seizure detection

5.1.1 Dataset Description

In this research, two open datasets are used, from Neurology & Sleep Centre, Hauz Khas, New Delhi [204] and the other is Bonn University [205]. The data sampling rate is 200 HZ and was acquired using a Grass Telefactor Comet AS40 Amplification System. A 10-20 electrode placement system was followed during scalp EEG electrode acquisition. Pre-ictal, interictal, and ictal stages were segmented by filtering the signals between 0.5 and 70 Hz. There are a total of 1024 samples in the file, each lasting 5.12 seconds. The dataset from Bonn University was collected using a 128-channel amplifier. The sampling rate of data is 173.61 Hz. The data is divided into 5 datasets (Z, O, N, F, and S) and these EEG signals have 4097 samples per channel with 23.6 seconds of duration.

5.1.2 Compressing EEG signal

Compressive sensing (CS) is capable of compressing EEG signals at a high ratio. Wavelet transform and iterative threshold methods were applied to ECG and EEG signals before being compressed using CS. By using the CS method, we can compress signals. So that the large EEG signals are compressed without loss of the data, and it will be easy to process. In [206], the signals are compressed by 1/2, 1/4 and 1/8. Entropy was calculated for the compressed EEG signals. The results show that the EEG signals are compressed by 1/8 gave better performance than 1/2 and 1/4.

5.1.3 Features extraction

The classification accuracy is improved by selecting proper features from datasets and reducing the size by removing unnecessary or less relevant features. In this paper statistical features are extracted, and performance metrics of classification are calculated. Entropy, mean, standard deviation and variance [207].

Entropy: Entropy can be used to distinguish between different types of EEG signals and can be used to identify different types of emotional states. Moreover, entropy can be utilized to predict mental states and detect anomalies in EEG data.

$$e = -\sum_1^n x^2 - \log x^2 \quad (5.1)$$

Mean: It is also known as the average and is useful for summarizing data quickly and easily. The Arithmetic Mean can be affected by extreme values, so it is important to consider the context of the data before using this statistic.

$$\mu = \frac{1}{N} \sum_N^1 X_n \quad (5.2)$$

Standard Deviation: The standard deviation is a measure of how spread out the values in a data set are. A high standard deviation means that the data points are far from the mean, while a low standard deviation indicates that they are clustered closely around the mean.

$$\sigma_x = \left(\frac{1}{N-1} \right) \sum_{n=1}^N (X_n - \mu_n) \quad (5.3)$$

Variance: Variance is calculated by taking the average of the squared differences from the mean. High variance indicates that the data points are spread out over a wider range, while low variance implies that the data points are closer to the mean and clustered together.

$$Var(X) = E[(X - \mu)^2] \quad (5.4)$$

5.1.4 SVM classifier for Epileptic seizure detection

Support vector machine is perfect and appropriate classifier for data classification. SVM is better than ANN, when compared mathematically [208]. The SVM approach relies on identifying the hyperplane that maximizes the margin between

two classes, hence optimizing their separation. This margin represents the distance from the hyperplane to the nearest data points, known as support vectors. Since SVMs can manage both linear and nonlinear connections, it provides good performance on tiny datasets. SVMs are quite scalable and able to effectively identify big datasets. The hyperplane arranges data points into one of the categories or class. Since SVM can estimate a continuous valued characteristic, it may also be applied in regression issues. Therefore, SVMs also help well in handling outliers. SVMs are better options for real-world use as they can adapt to misclassified data points. And they have low running computational expenses and are rather simple to apply.

Table 5.1 Explanation and Expressions of Performance metrics for SVM Classifier

Performance Metrics	Explanation	Expressions
Accuracy	Shows overall correctness	$\frac{TP + TN}{TP + FP + TN + FN}$
Precision	Shows correctness when the prediction of target.	$\frac{TP}{TP + FP}$
Sensitivity/ Recall/ True Positive Rate	Shows how often it can detect/identify the target/positive class	$\frac{TP}{TP + FN}$
Specificity/ True Negative Rate	Shows how often it can identify the negative class	$\frac{TN}{TN + FP}$
F1 Score	Mean of combined precision and recall	$\frac{(2 * Precision * Recall)}{(Precision + Recall)}$

The above Table 5.1 describes the performance metrics for calculating SVM classifiers with formulas. Where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively.

5.1.5 MATLAB HDL Coder

FPGAs can implement a variety of image processing and signal processing applications. MATLAB is often used to model images and signal processing applications. In this paper, MATLAB HDL coder is used for implementing ESD on FPGAs. This coder can produce HDL codes for RTL implementation in Verilog or VHDL. By designing ESD in MATLAB and mapping them to FPGA CAD tools, which makes the design process significantly faster.

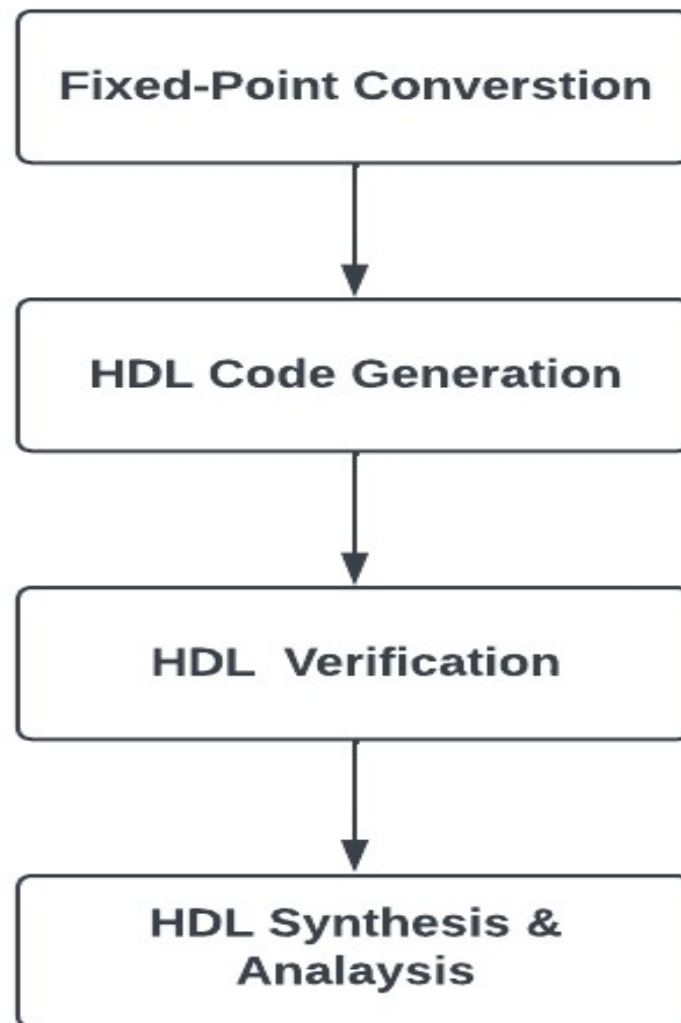


Figure 5.2 Steps involved in HDL Workflow Advisor

Implementations of ESD using MATLAB HDL coder

1. Modules are designed using the MATLAB algorithm.
2. Simulate, debug, and optimize the designs in MATLAB
3. HDL code is generated using MATLAB HDL coder
 - In HDL Workflow Advisor, the first step is to define the input types.
 - The second step is Fixed point conversion.
 - In the third step, must set the targeted FPGA board. Here need to select the synthesis tool, chip family and device package and speed.
 - For HDL code generation, we must set the preferable language VHDL or Verilog, coding styles, clocks & ports, optimization, script options etc.
 - The final step is HDL Verification, generate and run HDL test bench.
4. After verifying the design, synthesis and analysis takes place. Here we have to create the project and synthesis, and Implementation takes place.
 - Altera Cyclone® V SE 5CSEMA5F31C6N device

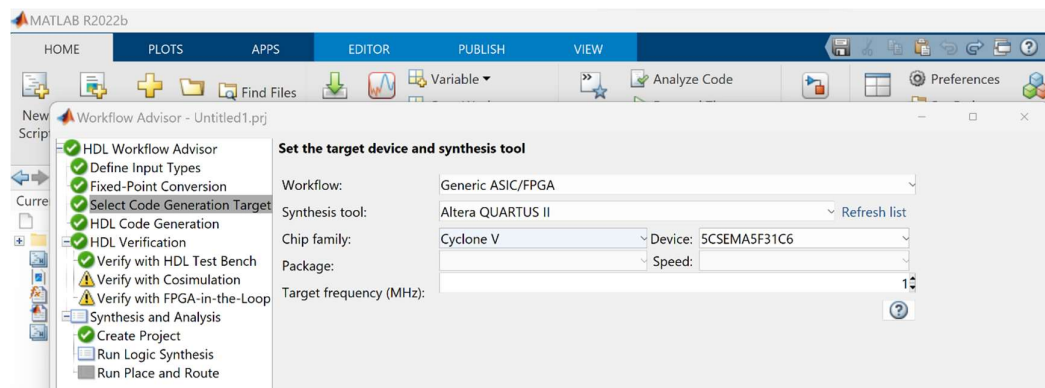


Figure 5.3 workflow advisor of MATLAB HDL Coder

5.1.6 Results and Discussion of Implementation of an Epileptic seizure detection

The performance metrics for classifiers are calculated for two datasets. The accuracy of an algorithm is measured by the true positive rate (TPR)/ sensitivity and false positive rate (FPR)/ specificity. The TPR refers to the ratio of correctly identified

seizure signals, while the FPR is the ratio of incorrectly classified non-seizure signals. Table 5.2 explains the performance metrics of SVM classifier using New Delhi dataset.

The SVM classifier showed an accuracy of 98.50%, a precision of 98%, specificity is of 98% and a recall of 96%. The F1 score for the classifier was also high at 97%. Table 5.3 shows the Bonn university dataset performance metrics.

Table 5.2 Performance Metrics of dataset of Neurology and sleep center, New Delhi

Performance Metrics	Ictal vs interictal	Ictal vs preictal	Ictal vs non-ictal
Accuracy	98.50	97	97
Precision	98	95	95.50
Sensitivity/ Recall	96	100	95
Specificity	98	95	96.12
F1 Score	97	97.43	95

Table 5.3 Performance Metrics of dataset of Bonn University

Performance Metrics	Normal	Inter-ictal	Ictal
Accuracy	97.05	92	96
Precision	93.33	95.45	90
Sensitivity/ Recall	96	98.35	94
Specificity	95	97	98.43
F1 Score	94.65	96.87	92

Table 5.4 Utilization summary of seizure detection modules for New Delhi dataset

Recourses	Pre-Processing	Compressing EEG signal	Feature Extraction	Classifier
LUTs (32070)	754	1531	377	1495
Block Memory (4065280)	353	464	473	1670
DSP (87)	16	18	21	23
User I/O (457)	42	57	68	72

Table 5.5 Utilization summary of seizure detection modules for Bonn University

Recourses	Pre-Processing	Compressing EEG signal	Feature Extraction	Classifier
LUTs (32070)	793	1876	400	1598
Block Memory (4065280)	386	562	473	1763
DSP (87)	12	18	32	25
User I/O (457)	38	56	72	63

Table 5.6 Comparison of utilization of FPGA implementation of seizure detection

Hardware Blocks	dataset	FPGA	LUT	DSP
DWT, SampEn, Threshold algorithm [209]	Bonn University	Virtex 6	2275	-
Feature selection module (Indira et al. 2021) [210]	Bonn University	Virtex 5	1096	-
MSSFs (Sukriti et al., 2021a) MPSoC [211]	Bonn University	Zynq Ultra Scale +	58,678	1160
Pre-processing Compressing EEG signal Feature Extraction SVM	Neurology and sleep center, New Delhi	Cyclone V SE 5CSEMA5F31C6N	754 1531 377 1495	16 18 21 23
Pre-processing Compressing EEG signal Feature Extraction SVM	Bonn University	Cyclone V SE 5CSEMA5F31C6N	793 1876 400 1598	12 18 32 25

Table 5.4 shows the FPGA utilization summary of the modules on INTEL FPGA board using ModelSim simulator. The SVM classifier was found to be highly efficient and accurate, and the FPGA utilization summary showed that the modules ran at a low resource utilization. This indicates that the SVM classifier can be used effectively in real-world applications. Table 5.5 shows the Bonn university dataset utilization.

The results show that the SVM classifier performs well on the New Delhi dataset. The FPGA utilization summary of the modules showed that the classifier utilizes a small number of resources, ensuring efficient implementation.

A comparison of the previous research papers with utilization summary of FPGAs to proposed FPGA epilepsy seizure detection presented in Table 5.6. The results of the comparison showed that the proposed system is more efficient in terms of FPGA utilization and is more suitable for epilepsy implementation. This suggests that the SVM classifier is a suitable choice for epilepsy seizure detection.

5.1.7 Conclusion of an Epileptic seizure detection implementation

The proposed work presents a methodology for utilizing MATLAB HDL Coder in various applications on FPGA. The MATLAB program is utilized for the simulation of planning and analysis in the detection of epileptic seizures. The module was further converted into HDL language via HDL Coder. The MATLAB HDL Coder facilitates a reduction in the development time for FPGA-based projects. FPGA-based systems are therefore an advantageous option for image and signal processing applications. It provides an affordable and expedient method for implementing digital systems on FPGA. It also assists in reducing development periods and swiftly producing complex digital designs. This enables rapid market introduction of innovations and products.

The classification strategy used in Epileptic seizure detection implementation adopted in speech-based TDS. The epileptic seizure detection was developed using MATLAB HDL Coder on an Intel FPGA platform. Which helped an initial hardware validation of the HDL approach, later influenced the final Artix-7 based TDS implementation.

5.2 Classifying the Tongue Contours for Assistive Devices

The development of tongue movement has several applications which may be used for therapy or rehabilitation and accessing ADs. The main motto for classifying tongue contours is to understand the tongue motions during articulation [212]. This study offers substantial insights into several interconnected research domains, such as speech disorders resulting from diseases, the biomechanics of tongue modeling that elucidates speech processing, swallowing, and other tongue functions, and Second Language Acquisition (SLA), encompassing the cognitive processes involved in learning additional languages. Nearly all instruments are unreachable by the tongue, complicating the assessment of tongue function. In speech research, ultrasound imaging is the predominant technology employed to assess tongue movements during articulation, as it captures picture sequences that document these movements in real-time.

Ultrasound imaging is non-invasive, cost-effective, and capable of accurately capturing the motions of the tongue. It serves as an essential instrument for evaluating the efficacy of speech therapies, since it enables the monitoring of alterations in tongue movement [213, 214].

In recent years, several studies have been conducted on the creation of speech-controlled ADs. A speech-centric methodology is employed in the design and execution of assistive technologies. This study [215] discusses several TDS and compares low-power TDS. Advanced machine learning algorithms enhance the accuracy of these systems over time by assimilating user interactions. These methods are advantageous for those with impairments who struggle to use assistive gadgets with their hands. Moreover, these technologies can provide more natural and intuitive interactions with other devices [216].

The categorization of tongue contours can be beneficial in the domain of AT and healthcare. Individuals with speech or movement difficulties may get advantages from assistive communication devices created by tongue contour categorization. The gadgets can discern the user's intent through tongue shapes and motions. Individuals recovering from speech difficulties and accidents may get advantages from the feedback offered by a speech rehabilitation program. Analyzing tongue motions can

furnish therapists with insights into progress and aid in formulating targeted therapeutic activities for certain articulatory actions.

5.2.1 Tongue Contours Classification

Ultrasound offers several benefits when it comes to classifying the shapes of the tongue. In this work, the tongue contour motions for ADs are classified using a hybrid technique. Figure 5.4 illustrates the block diagram of the hybrid technique for categorizing tongue contours from ultrasonography for ADs.

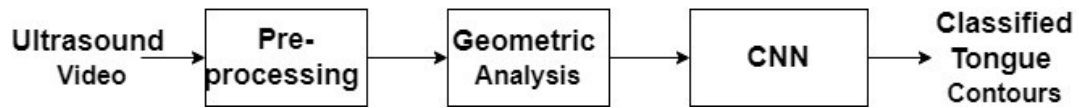


Figure 5.4 Classification of tongue contours using hybrid approaches

5.2.2 Pre-processing

Pre-processing includes Frame Extraction, Noise Reduction, Normalization, and Smoothing. The efficacy of a noise reduction method is dependent upon the nature of the noise and the specific requirements of the ultrasonic imaging assignment.

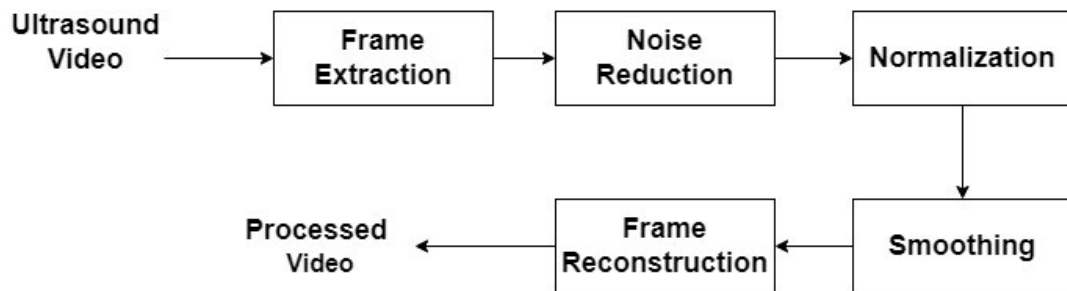


Figure 5.5 Pre-processing blocks for ultrasound video of tongue contours

Figure 5.5 illustrates the pre-processing components of the ultrasound video depicting tongue contours. The initial stage in this process is to extract frames from the ultrasound video. Effective noise reduction methods improve contour quality and boost the accuracy of future analyses. Noise reduction is an essential phase in picture pre-processing, particularly for ultrasound images characterized by widespread speckle noise. Utilizing methods such as SRAD and NLM filtering may substantially improve

the quality of ultrasound recordings, rendering them more appropriate for further analysis and diagnostic applications. These strategies facilitate the preservation of significant features and borders while diminishing extraneous noise, resulting in clearer and more interpretable pictures.

Dataset: The ultrasound tongue contour dataset which is used in this work is available publicly [217]. The dataset is with and without audio of ultrasound videos are available.

Speckle Reducing Anisotropic Diffusion (SRAD): SRAD is a specialized kind of anisotropic diffusion employed to analyze ultrasound pictures, effectively diminishing speckle noise while maintaining intricate features, including edges. Speckle noise is a granular artifact intrinsic to ultrasonic imaging, resulting from the coherent characteristics of wave reflections [218].

The SRAD technique employs a partial differential equation (PDE) to execute anisotropic diffusion. The temporal development of picture $I(x,y,t)$ is described by the following partial differential equation (PDE):

$$\frac{\partial I}{\partial t} = \nabla \cdot (c(x, y, t) \nabla I) \quad (5.5)$$

Here, $\nabla \cdot$ is the divergence operator, ∇I is the gradient of the image, and $c(x, y, t)$ is the diffusion coefficient which controls the rate of diffusion based on local image characteristics.

The diffusion coefficient (eq. 2) $c(x, y, t)$ in SRAD is designed to reduce speckle noise while preserving edges. It is defined as:

$$c(x, y, t) = \exp \left(- \left(\frac{\nabla_{norm}^2(I)}{I^2 + \nabla_{norm}(I^2) + k^2} \right) \right) \quad (5.6)$$

Where $\nabla_{norm}(I)$ is the normalized gradient magnitude of the image, I is the image intensity at a specific point and I is a constant that controls the sensitivity of the diffusion coefficient.

$$\nabla_{norm}(I) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad (5.7)$$

Speckle Reducing Anisotropic Diffusion (SRAD) is an effective method for mitigating speckle noise in ultrasound pictures while maintaining edge integrity. SRAD can considerably improve the quality of ultrasound images, thereby simplifying the process of analyzing and extracting meaningful information, by meticulously adjusting the parameters.

Non-Local Means (NLM) filtering: NLM filtering is a sophisticated method for denoising that utilizes the redundancy present in the picture. It operates by comparing pixel patches instead of individual pixels and averaging pixels depending on the similarity of their surrounding patches [219].

$$I_{NLM}(p) = \frac{\sum_{q \in \Omega} w(p,q)I(q)}{\sum_{q \in \Omega} w(p,q)} \quad (5.8)$$

Where, I is the noisy image, Ω is the search window around pixel p . $w(p, q)$ is the weight that measures the similarity between patches centered at p and q . $I(q)$ is the intensity of pixel q .

$$w(p, q) = \exp\left(\frac{-\|P(p) - P(q)\|^2}{h^2}\right) \quad (5.9)$$

Where $P(p)$ and $P(q)$ are patches (small windows) centered around pixels p and q , respectively. $\|P(p) - P(q)\|^2$ is the Euclidean distance between patches $P(p)$ and $P(q)$. h is the filtering parameter that controls the decay of the exponential function.

Normalization: The pixels in a picture undergo normalization, wherein their intensity levels are evaluated on a uniform scale. This procedure improves contrast and guarantees uniformity across many frames. Normalization can render faint details more discernible by boosting contrast. Guarantees that the frames possess a consistent intensity range, essential for later processing stages.

Smoothing: Smoothing diminishes noise and small fluctuations in a picture while maintaining significant structural elements such as edges. A prevalent smoothing method that use a Gaussian kernel to provide weights to adjacent pixels, prioritizing

those in closer proximity. Accurately calibrated smoothing may diminish noise while maintaining critical features such as edges and curves.

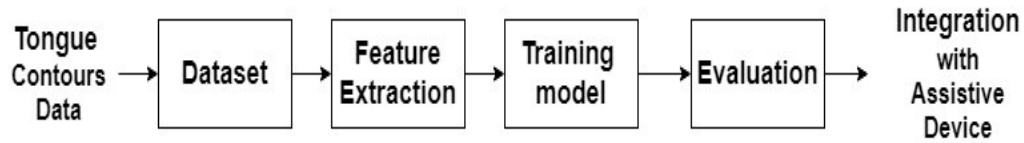


Figure 5.6 Block diagram of classifying the tongue contours

5.2.3 Geometric Analysis

In order to comprehend tongue morphology, movement, and speech-related dynamics, it is imperative to conduct geometric analysis of tongue contours from ultrasound recordings. Examine the morphology of the tongue contours utilizing geometric parameters including area, perimeter, eccentricity, circularity, and others. Apply geometric shapes (e.g., ellipses) to the contours and examine the fitting parameters. Monitor the movement of tongue contours over time to examine dynamics including velocity, trajectory, and acceleration [220]. Compute optical flow or employ tracking techniques to monitor contour points over successive frames [221]. Identify and monitor certain landmarks on the tongue's shape to assess deformation or movement patterns.

5.2.4 Convolution Neural Network (CNN)

In the domain of computer vision, the application of convolutional neural networks (CNN) has undergone a revolution in recent years. and have emerged as the preferred option for several image-related activities [221]. CNNs are very proficient in autonomously identifying hierarchical characteristics from unprocessed data. CNNs have attained significant success across several applications, establishing themselves as essential instruments in contemporary machine learning and artificial intelligence. CNNs may partition pictures into distinct areas or objects, designating each pixel with a label that denotes its category. CNNs are employed in medical imaging for illness diagnosis, tumor detection, and organ segmentation [223]. Figure 5.7 illustrates the convolution of tongue contours via ultrasound imagery.

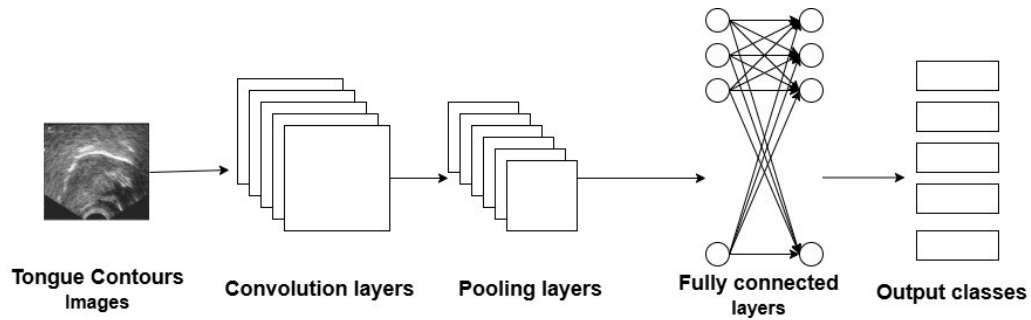


Figure 5.7 Classification of tongue contours using Convolution Neural Network (CNN)[224]

Convolution layers: Convolutional layers constitute fundamental components of Convolutional Neural Networks (CNNs). Feature extraction is achieved by the application of learnable filters (kernels) to input data. Filters convolve over the input data, calculating dot products to generate feature maps.

Pooling Layers: Pooling layers reduce the dimensionality of feature maps by aggregating local information. Alongside max pooling, average pooling is a prevalent technique utilized to diminish spatial dimensions while maintaining essential attributes.

Fully connected layers: Feature maps are flattened and input into one or more fully linked layers subsequent to convolutions and pooling operations. These layers function to consolidate the characteristics acquired from the convolutional layers and generate final predictions. They function as a classifier, integrating and assigning weights to the diverse characteristics to provide the output, whether it a class label, a regression result, or an alternative prediction format. Classification is performed based on characteristics acquired in fully linked layers.

5.2.5 Performance Metrics

The performance measures of noise reduction for ultrasound tongue contours, geometric analysis, and classification methods are discussed. Two strategies for noise reduction are employed.

Mean Squared Error (MSE): MSE is frequently employed to assess the efficacy of prediction models. In ultrasound imaging, particularly for segmentation, registration, or enhancement tasks, MSE significantly contributes to evaluating the quality of these procedures [225].

$$\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [I(i, j) - K(i, i)]^2 \quad (5.10)$$

Peak Signal-to-noise ratio (PSNR): In the domain of image reconstruction, PSNR has emerged as a prominent performance statistic, particularly in ultrasound imaging. PSNR measures the disparity between an original (reference) image and its compressed or reconstructed counterpart, often utilized in image processing, compression, and enhancement contexts [226].

$$10 \log_{10} \left(\frac{MAX^2}{MSE} \right) \quad (5.11)$$

Structural Similarity Index (SSIM): Analyze two images for similarity by comparing their structural content, luminance, and contrast. Unlike Peak Signal-to-Noise Ratio (PSNR), the purpose of SSIM is to mimic the way humans perceive the quality of images on a pixel-by-pixel basis. This makes it particularly useful for assessing ultrasound images of excellent quality, which often require fine details and structures for accurate diagnostics [227].

$$\frac{(2\mu_x\mu_y+C1)(2\sigma_x\sigma_y+C2)}{(\mu_x^2+\mu_y^2+C1)(\sigma_x^2+\sigma_y^2+C2)} \quad (5.12)$$

Extended Performance Index (EPI): EPI is a comprehensive performance metric specifically designed for ultrasound imaging. It assesses various aspects of ultrasound image quality to provide a holistic view of the performance of ultrasound systems and algorithms. EPI combines multiple metrics to evaluate different dimensions of ultrasound image quality, including resolution, contrast, noise, and artifact levels. This multi-faceted approach helps in understanding the overall performance of ultrasound imaging systems [228].

$$\frac{\sum_{ij} \min(\nabla I(i, j), \nabla K(i, j))}{\sum_{ij} \nabla I(i, j)} \quad (5.13)$$

Contrast-to-Noise Ratio (CNR): Assessing the ability to distinguish between different tissue types or structures in the presence of noise. It is a measure of how well a feature (such as a lesion) can be detected against the background tissue [229].

$$\frac{|\mu_{signal} - \mu_{background}|}{\sqrt{\sigma_{signal}^2 + \sigma_{background}^2}} \quad (5.14)$$

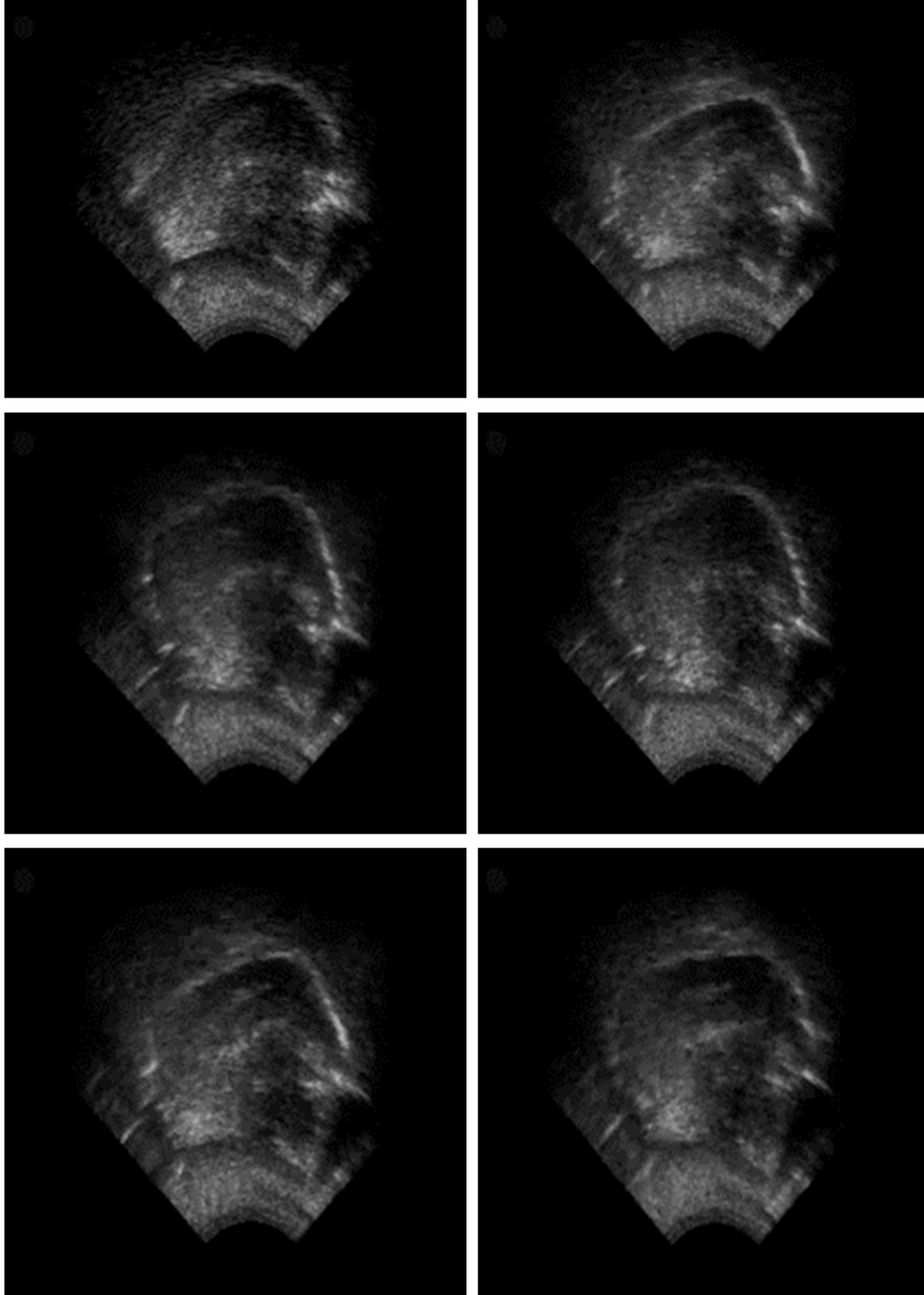


Figure 5.8 Frame extraction of tongue contours from ultrasound video

In the first stage, the noise of the ultrasound was removed using SRAD and NLM technique. Which reduces the speckle noise and enhances the quality of ultrasound videos making them more suitable for further analysis. Analyzing tongue contours in ultrasound videos involves tracking and quantifying the shape of the tongue over time.

5.2.6 Results of Tongue Contours Classification for assistive Devices

The extraction of tongue contours has several applications in the medical domain. The categorization of tongue contours is beneficial for language processing, rehabilitation, monitoring patient speech performance, and AT, among other applications. The main purpose of classifying tongue contours to activate AD for the disabled people who has issues with speaking properly like stuttering problem. Speech-based TDS works well for the people who does not have an issue with speaking, by classifying tongue contours from ultrasound images solves the problem of activating AD even though disabled have stuttering issue. The Figure 5.8 shows the frames of tongue contours extracted from ultrasound video. The ultrasound video of tongue contours has around 10 seconds. To extract the tongue contours from the ultrasound video pre-processing techniques are to eliminate noise.

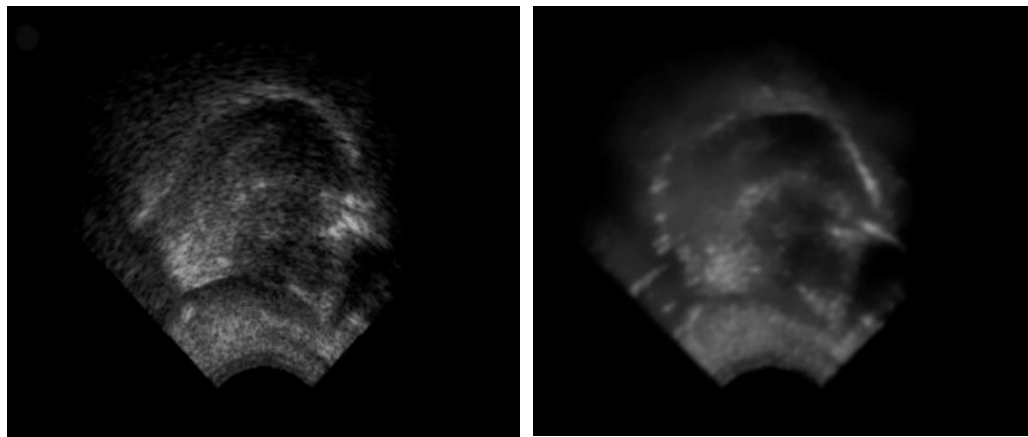


Figure 5.9 Images of input and output ultrasound of tongue contours

In hybrid approaches we merged two techniques to enhance quality and process the frames without noise. Here we used two techniques to remove the speckle noise, one is SRAD, and the other is NLM. The results of ultrasound tongue frame noise reduction showed in Figure 5.9.

Table 5.7 Metrics of SRAD and NLM noise reduction

Performance Metrics	SRAD	NLM
MSE	23.50	21.20
PSNR	35.00	32.00
SSIM	0.83	0.80
EPI	0.71	0.69
CNR	2.85	2.02

Performance metrics of noise reduction are presented in Table 5.7. The calculation of MSE, PSNR, SSIM and NLM were calculated for noise reduction techniques SRAD and NLM.

```
>> geo_feature_extraction
Contour Length: 52.28 pixels
Contour Area: 65.00 pixels^2
Mean Angle: 42.91 degrees
Mean Curvature: 0.55444
>> geo_feature_extraction
Contour Length: 26.14 pixels
Contour Area: 13.00 pixels^2
Mean Angle: 45.00 degrees
Mean Curvature: 0.45807
```

Figure 5.10 Command window of geometric feature extraction

The Figure 5.10 shows the area, length, mean angle and curvature of ultrasound tongue contours image results in MATLAB command window. Geometric analysis of ultrasound tongue is calculated for better extraction of tongue contours. Perimeter and area of tongue over a time is calculated and Figure 5.11 shows the graph of area and perimeter over a time. Classifier's performance metrics - Accuracy, precision, specificity, recall and F1 score formulae and CNN performance are presented in Table 5.8. The present work can be extended by implementing on FPGA. The classification of tongue contours using ultrasound video designed in MATLAB and the program is converted to HDL language using MATLAB HDL coder and Vivado 2022.2 is interfaced with MATLAB for FPGA implementation.

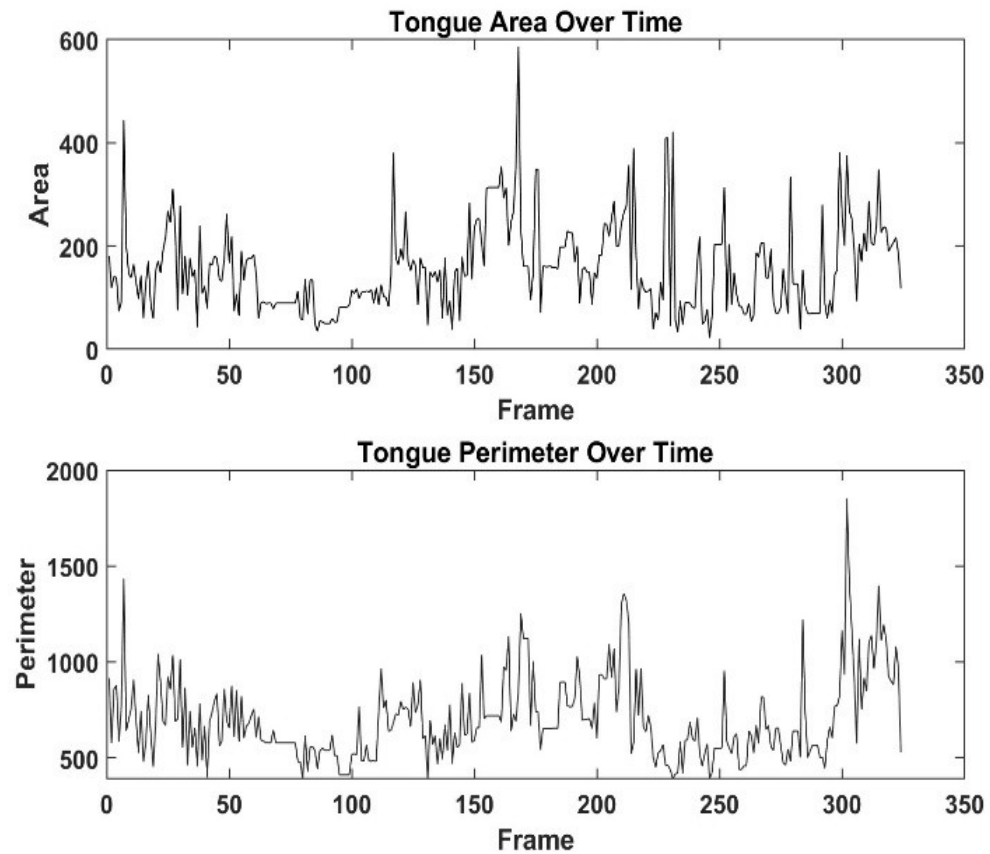


Figure 5.11 Graph of tongue area and perimeter over time

Table 5.8 Performance metrics of classifier using CNN

Metrics	Formula	Performance (in %)
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	87.80
Precision	$\frac{TP}{TP + FP}$	70.58
Specificity	$\frac{TN}{TN + FP}$	84.08
Sensitivity/ Recall	$\frac{TP}{TP + FN}$	70.40
F1 Score	$\frac{2 \times Precision \times Recall}{Precision + Recall}$	77.49

5.2.7 Conclusion of Classification of Tongue Contours for Assistive Devices

A hybrid methodology is proposed for the design of the classifier. A hybrid approach utilizing SRAD and NLM techniques is employed to diminish noise in ultrasound videos. The efficacy of CNNs is assessed for tongue contour categorization. The CNN attains an accuracy of 87.80 for tongue contours. The specificity and sensitivity of the CNN are 84.08% and 70.40%, respectively. The F1 score of CNN is 77.49%. Incorporating data augmentation strategies might boost the performance of the CNN model by increasing the variety of training samples. Moreover, exploring deeper network designs or incorporating transfer learning from pre-trained models may enhance the classifier's accuracy and robustness.

Disabled people with speech problems, such as stuttering, using speech-based TDS may lack reliability. The tongue area over time reflects changes in the area across consecutive frames, documenting volumetric movement during speech. Similarly, the perimeter of the tongue over time indicates boundary deformation, delineating shape variation during speech production. These temporal trajectories encapsulate unique tongue movement patterns, which are then analyzed using a CNN-based architecture

for assisted command word. This method facilitates the differentiation of intended speech, even in the presence of acoustic degradation, and establishes a foundation for future research on TDS.

This chapter presents the case studies of EEG-based epileptic seizure detection and ultrasound-based tongue contour extraction to illustrate the adaptability of the proposed machine learning architecture and assistive technology across diverse biological datasets. The EEG implementation studies the MATLAB-to-FPGA HDL procedure on an Intel FPGA platform, facilitating preliminary hardware validation of the fixed-point classifier implementation. The ultrasound-based CNN framework facilitated the extraction of dynamic tongue descriptors, such as tongue area and perimeter across time, providing an alternative supportive method for individuals with speech articulation difficulties for assistive device classification.

While these experiments are distinct from the speech-based TDS, they substantially enhanced the refinement of signal processing, classifier optimization, and hardware deployment methodologies utilized in the Artix-7 implementation. The examination of tongue contour outlines a prospective path for multimodal assistive control, complementing speech-based input for individuals with stuttering or diminished vocal clarity. These collaborative studies affirm the generalizability of the proposed methodology and enhance the overall design of the FPGA-based assistive system.

Chapter 6

Results and Discussion

6.1 Machine Learning classifiers

The First and main objective the TDS implementation on FPGA based on speech is deciding the best classifier for TDS. Different binary classifications for speech is designed in MATLAB and compared the performance metrics of classifiers. The better metrics classifier chooses for TDS implementation for eight commands of speech signal Classifier.

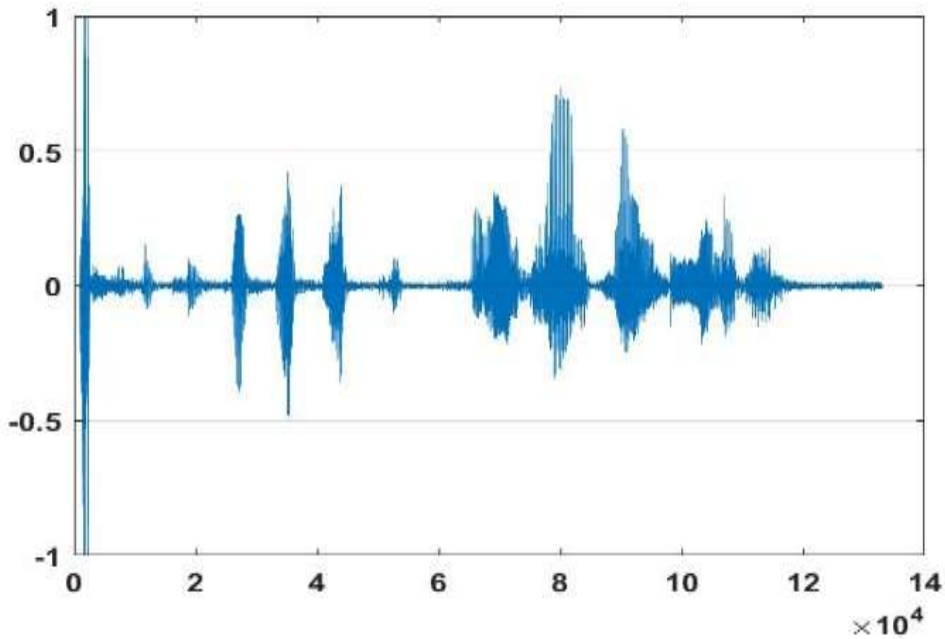


Figure 6.1 Raw audio signal

MLC are engineered to identify speech from a raw audio stream. This study used a 3-second audio stream of around 60 kilobytes as the signal input. Audio signals are represented as Waveform Audio Files (WAV). Figure 6.1 presents an ordinal audio signal, utilized as input for all MLC models to compute performance measures. Classes denote the training dataset; the audio signals are categorized into classes for data training. The audio signals are stored in .wav format, with all files consolidated in a single folder for input into machine learning algorithms for prediction.

The comparison table for different MLCs are presented in Table 6.1. These classifiers are binary classifiers implemented for same dataset. While calculating the metrics of classifier, which mentioned in the methodology chapter. SVM shows 91% of accuracy, 0.6985, 0.7884, 0.6597 and 0.7407 of precision, sensitivity, specificity and F1 score, respectively. For NB, 84.227 % of accuracy, 0.6950, 0.7832, 0.6354, and 0.7365 of precision, sensitivity, specificity and F1 score, respectively. For SGD and RF, 93 and 87 percentage of accuracy was shown. Precision for SGD and RF is 0.7130 and 0.6371; Sensitivity-0.6916 and 0.6402; specificity-0.7216 and 0.6354; F1 score-0.7021 and 0.6387 for both SGD and RF classifiers. Which shows clearly that SVM has a better performance compared to all the audio classifiers. Even though SGD has better accuracy than SVM, it has low sensitivity compared to SVM. Bar chart is presented in Figure 6.2.

Table 6.1 Comparison of Performance metrics for different classifiers

Performance Metrics	Machine Learning Classifiers			
	<i>SVM</i>	<i>NB</i>	<i>SGD</i>	<i>RF</i>
Accuracy (%)	91	84.227	93	87
Precision	0.6985	0.6950	0.7130	0.6371
Sensitivity	0.7884	0.7832	0.6916	0.6402
Specificity	0.6597	0.6354	0.7216	0.6354
F1 score	0.7407	0.7365	0.7021	0.6387

All the performance metrics re calculated for MLC to detect the speech from audio signals. Table 6.1 presents all statistical values of MLC of SVM, NB, SGD and RF.

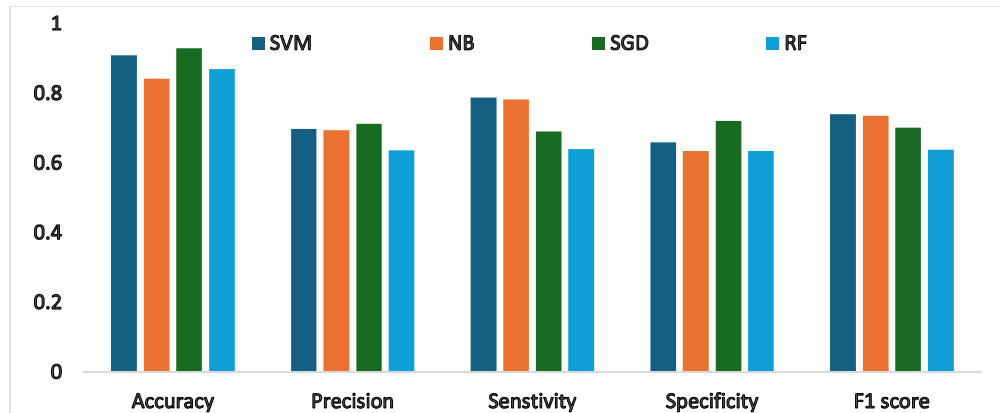


Figure 6.2 Bar chart of comparison of Performance Metrics of MLCs

All performance characteristics are examined (Table 6.1); Stochastic Gradient Descent (SGD) exhibits superior accuracy relative to other classifiers. Enhanced precision will yield accurate figures. Optimal MLCs require great precision, indicating that the precision value approaches 1. The model's specificity should be high for an optimal classifier. RF has reduced sensitivity and elevated specificity. The F1 score for the SVM model is elevated. Figures 6.3, 6.4, 6.5, and 6.6 illustrate the ROC curves of the MLCs: SVM, NB, SGD, and RF. Which provides the accuracy metrics for each cut-point value. Analysis of the graphs indicates that SVM has superior accuracy compared to NB, RF, and SGD.

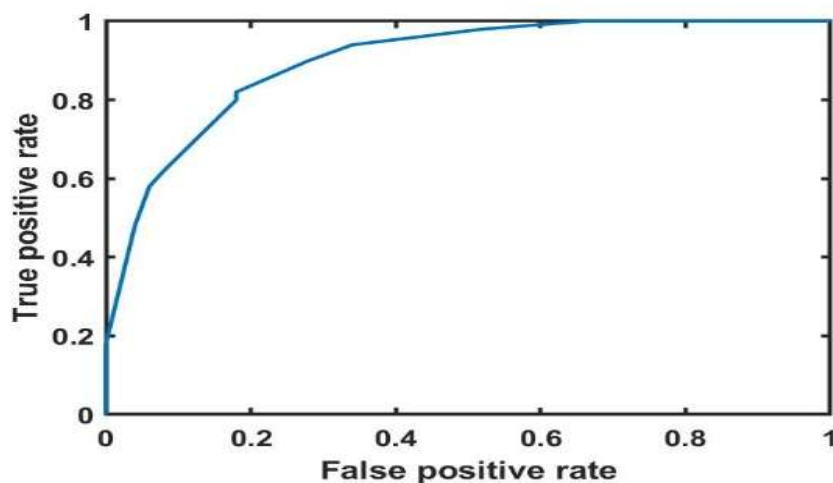


Figure 6.3 ROC diagram of SGD

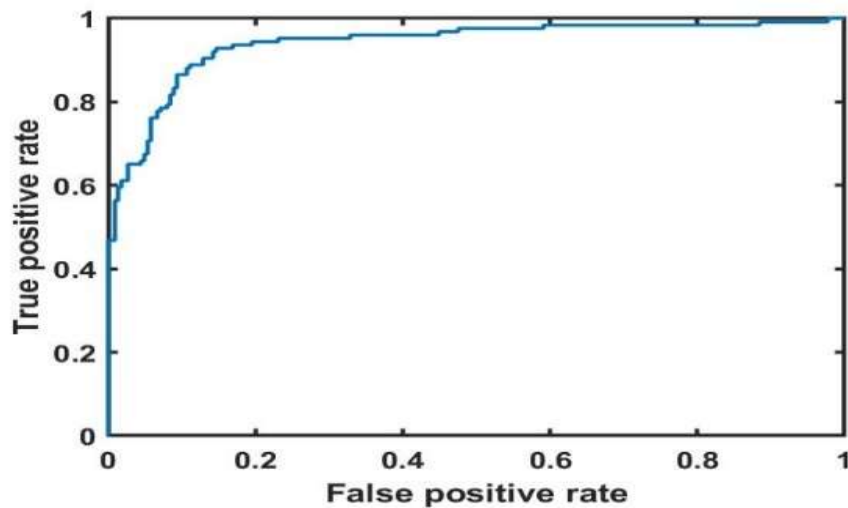


Figure 6.4 ROC diagram of NB

MLCs are engineered to identify speech from audio signals. These classifiers may utilize Tongue Drive methods, security locks that unlock specific individuals, detection of particular sounds from instruments or animals, among several other automatic speech recognition methods. Four classifiers are developed utilizing machine learning methods for voice detection. The MATLAB tool is utilized for designing classifiers. The input audio signal s , lasting 3 seconds, is sent to the model, followed by an evaluation of the outcomes and a comparison of the performance measures. SGD exhibits a high accuracy rate of 93%. SVM has superior sensitivity and F1 score in comparison to NB, RF, and SGD classifiers. The Support Vector Machine (SVM) has the highest Receiver Operating Characteristic (ROC) area for the speech detection classifier, whereas the random forest classifier demonstrates a lower ROC area. The ROC provides accuracy for each cut-point; a superior ROC indicates enhanced performance. The future development of classifiers for audio signals aims to enhance accuracy for extensive data sets; it is essential to utilize a substantial training dataset to get elevated precision. The domain of audio categorization will expand across several domains, whereas the majority of studies now focus on natural language processing. These auditory categories are utilized in AT due to the tongue being the least efficient organ in the human body. Issuing commands verbally will be simpler than through gestures or signs.

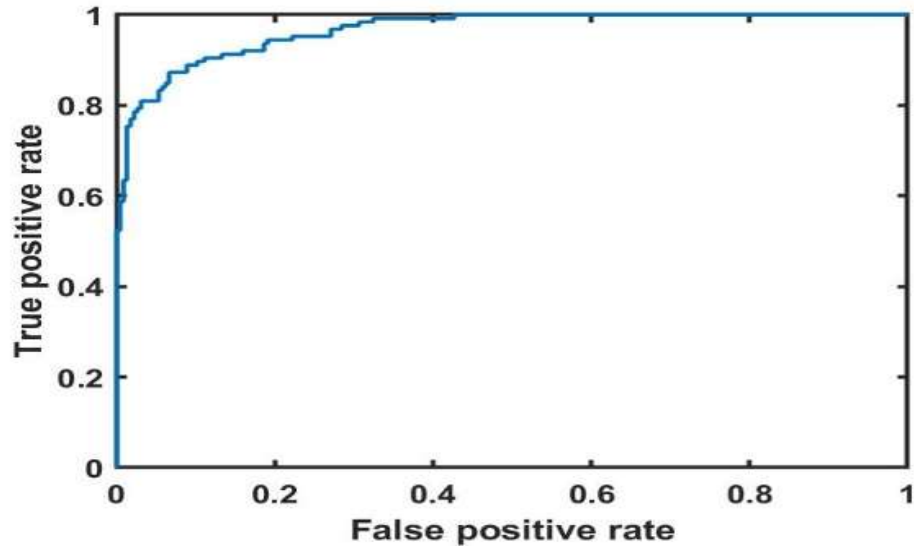


Figure 6.5 ROC diagram of SVM

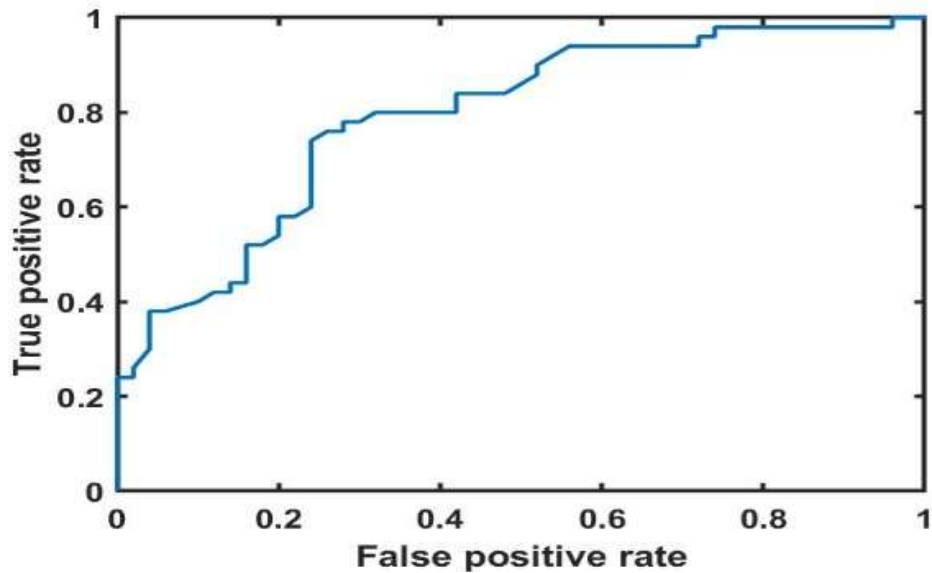


Figure 6.6 ROC diagram of RF

6.2 Active and Passive EMI filters Design

The active and passive EMI filters are specifically intended for audio applications, including a transition range of 20 Hz to 20 kHz. The transition area of a filter influences audio and noise signals differentially, contingent upon the filter's properties and its frequencies. A transition area will progressively diminish noise frequencies as they near the filter's stopband boundary, provided the noise signal encompasses frequencies inside the stopband. The attenuation of the output signal

decreases noise levels. The settings are selected to minimize the noise level. The cutoff frequencies of the active and passive filters are created in MATLAB R2020a, and the amplitude and phase responses of the circuits across frequencies are acquired. The cut-off frequency, insertion loss, return loss, and gain of the filters are computed based on the passive parameter values, which are determined in accordance with the data sampling frequency.

Figure 6.7, 6.8, 6.9 and 6.10 illustrates the amplitude and frequency response of the EMI filters for Stage I and Stage II. The cut-off frequency of the Stage-I EMI filter is 4.84 kHz. The EMI filter's cut-off frequency is 1.014 kHz. Table I illustrates the phase and amplitude response across frequencies when the values of the passive components for Stage I and Stage II are specified.

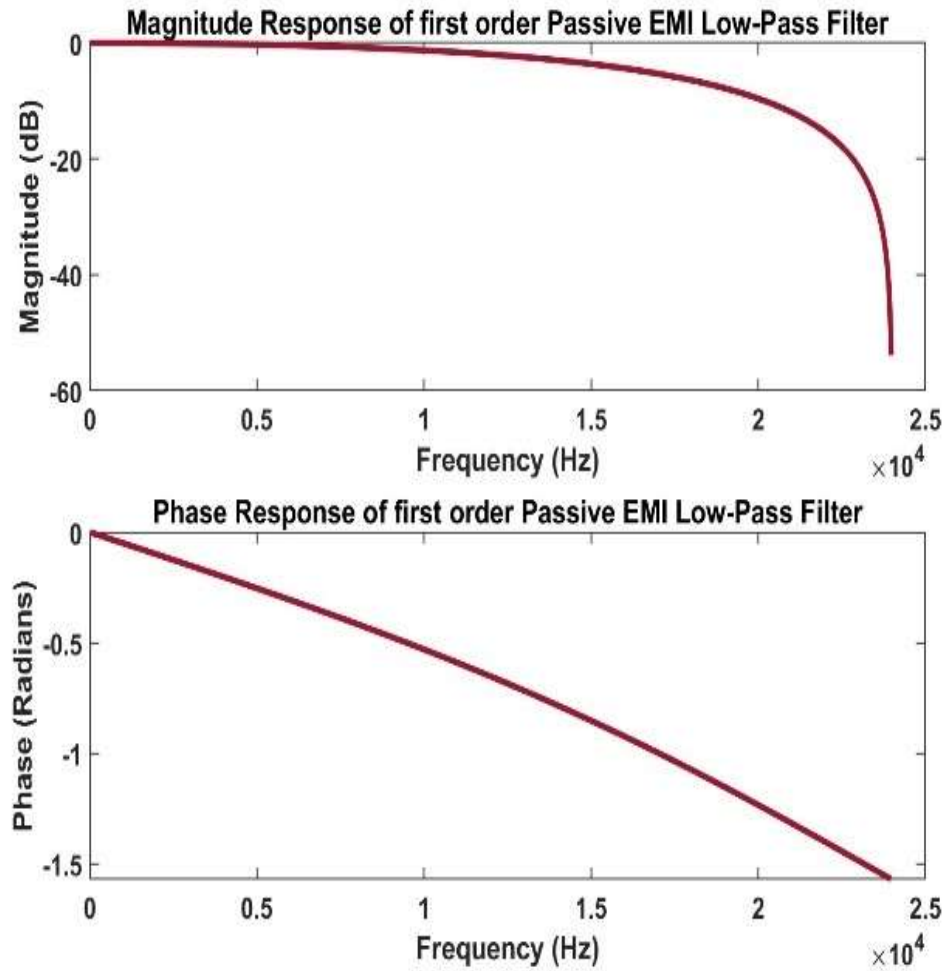


Figure 6.7 Stage-I Passive EMI low-pass filter

EMI filters are significantly influenced by resistance (R), capacitance (C), and inductance (L), which dictate the filter's frequency response, attenuation characteristics, and overall efficacy in eliminating undesirable EMI signals. Signals via a filter are often diminished by elevated resistance levels. Increasing the capacitance of a low-pass or band-pass filter reduces the cutoff frequency, enabling the transmission of lower frequencies.

The gradient of the filter's frequency response curve is dictated by capacitors. The roll-off slope is more pronounced with greater capacitance values. High-pass and band-pass filters have reduced cutoff frequencies with increased inductance values. Consequently, higher frequencies can traverse such filters. The frequency response curves of filters are affected by inductors. The roll off characteristics are more pronounced with elevated inductance levels.

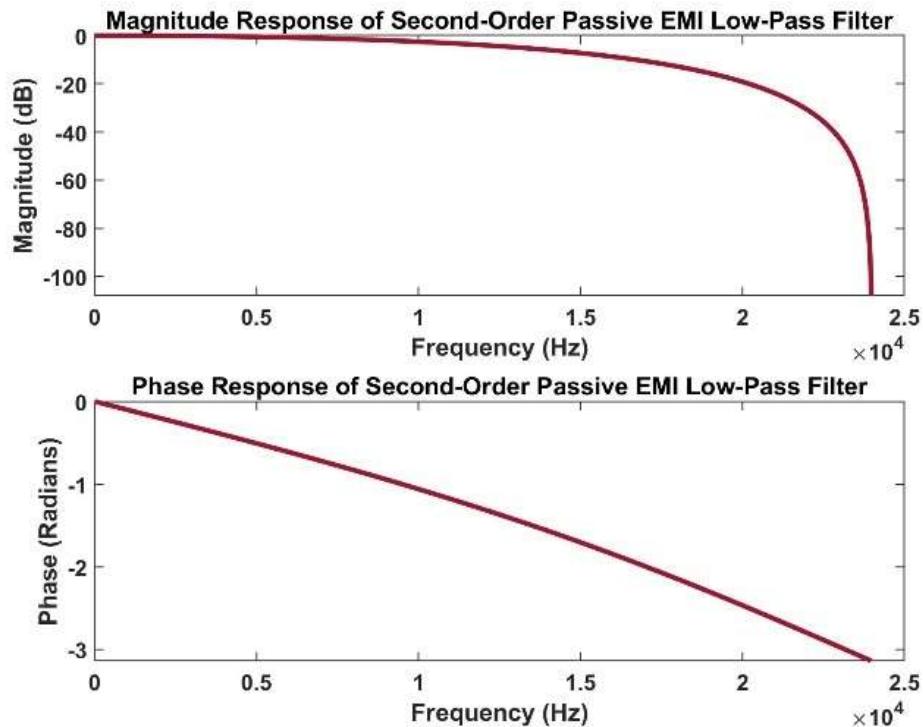


Figure 6.8 Stage-II Passive EMI low-pass filter

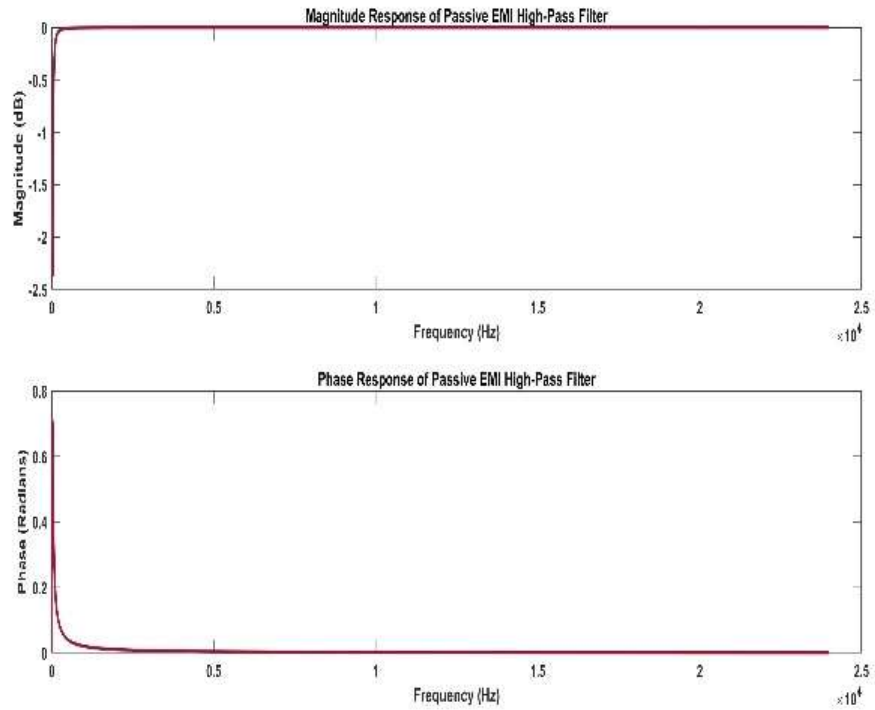


Figure 6.9 Stage I passive EMI High-pass filter response

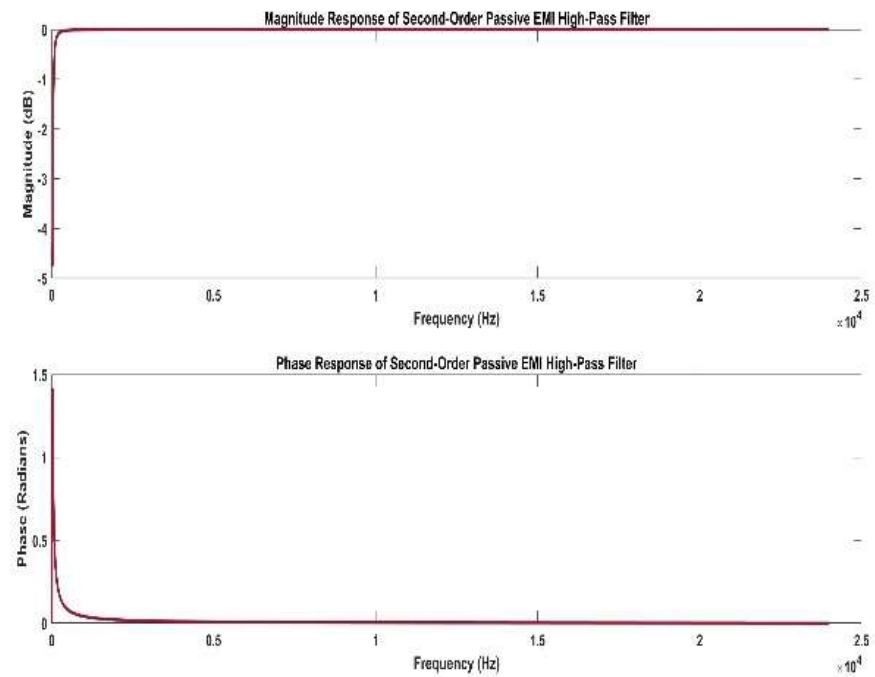


Figure 6.10 Stage-II Passive EMI filter response

The amplitude and phase response throughout the frequency of the active EMI filter, both high pass and low pass (figure 6.11, was derived when the values of the passive components were selected based on the sampling frequency, as specified in table II. Figure 6.12 illustrates the output of simple passive low and high pass filters.

The filters' response is contingent upon the parameter value. The signal will be filtered based on the parameters when a noisy input signal is provided. Two distinct noise signals are presented to filters. Figure 6.13 illustrates the amplitude of the noisy signal with time, alongside the filtered signal in both the time and frequency domains. The active EMI filter is constructed with an amplifier and passive components. The MATLAB program is utilized to design and optimize the filter, which is subsequently turned into HDL code for further implementation.

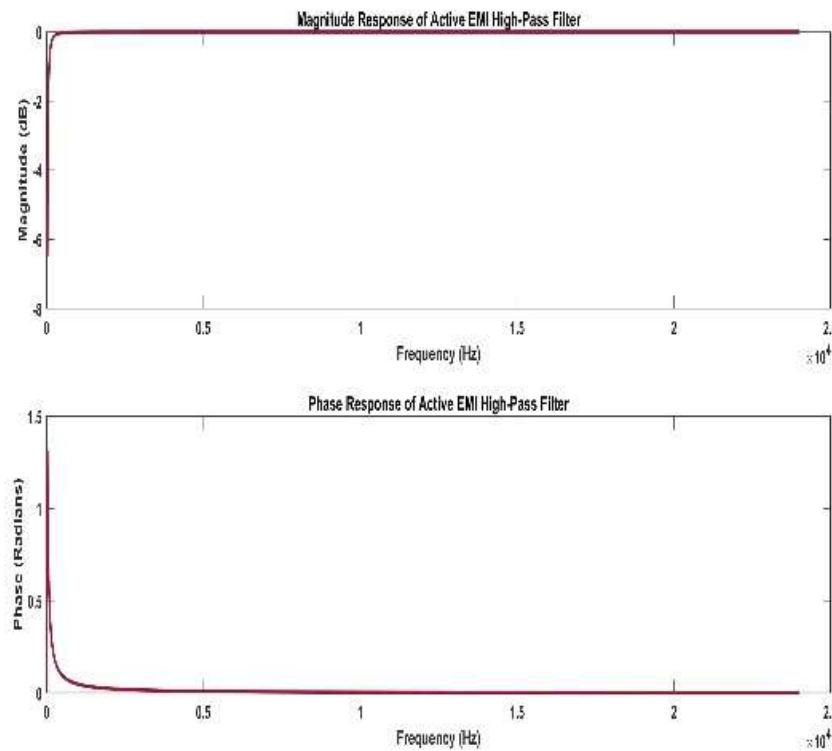


Figure 6.11 Response of Active EMI high pass filter

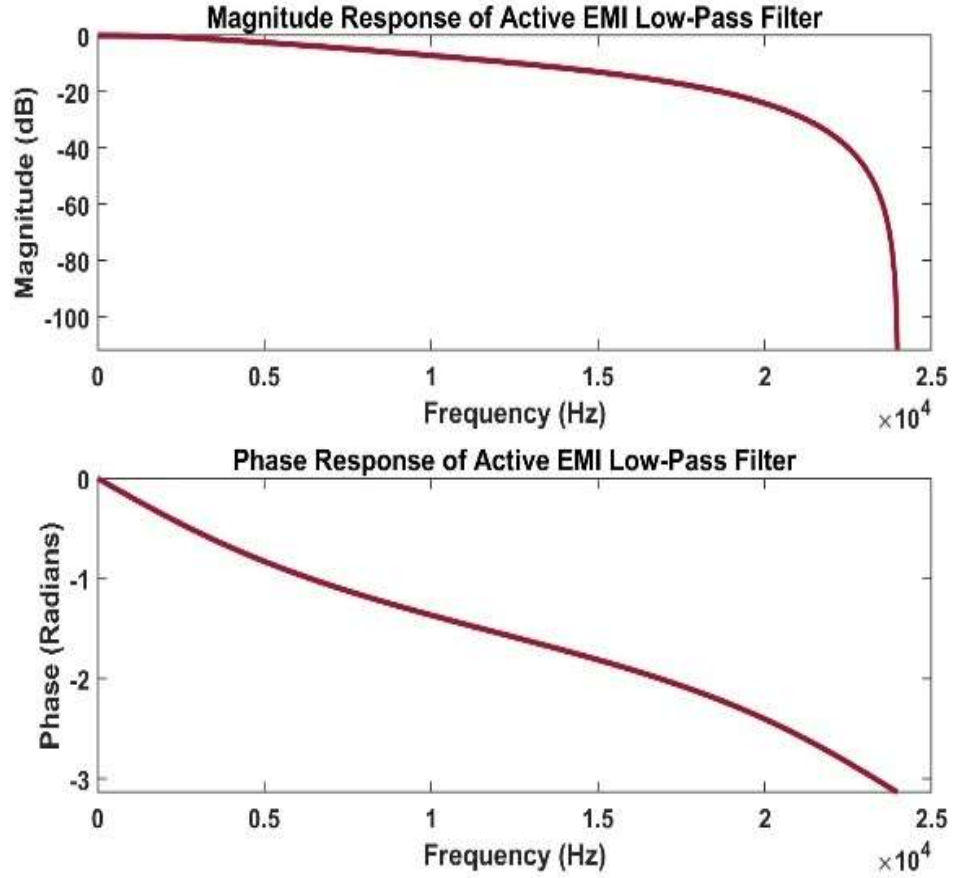


Figure 6.12 Response of Active EMI low-pass filter

The primary performance metric for noise reduction is insertion loss, measured in decibels. The equation shown below eq. (6.1) is used to compute the cut-off frequency for each filter, which will be employed for data sampling in communication. The cut-off frequency equation (6.2) of the filter is derived using passive components. The return loss of the EMI filter is determined using equation (6.3).

$$\text{insetion loss} = 20 \log_{10} \frac{\text{unfiltered signal amplitude}}{\text{filtered signal amplitude}} \quad (6.1)$$

$$\text{Cutoff Frequency } (f_c) = \frac{1}{2\pi RC} \quad (6.2)$$

$$\text{Return Loss (dB)} = 20 \log_{10} \left(\frac{\text{amplitude of the incident signal}}{\text{amplituded of the reflected signal}} \right) \quad (6.3)$$

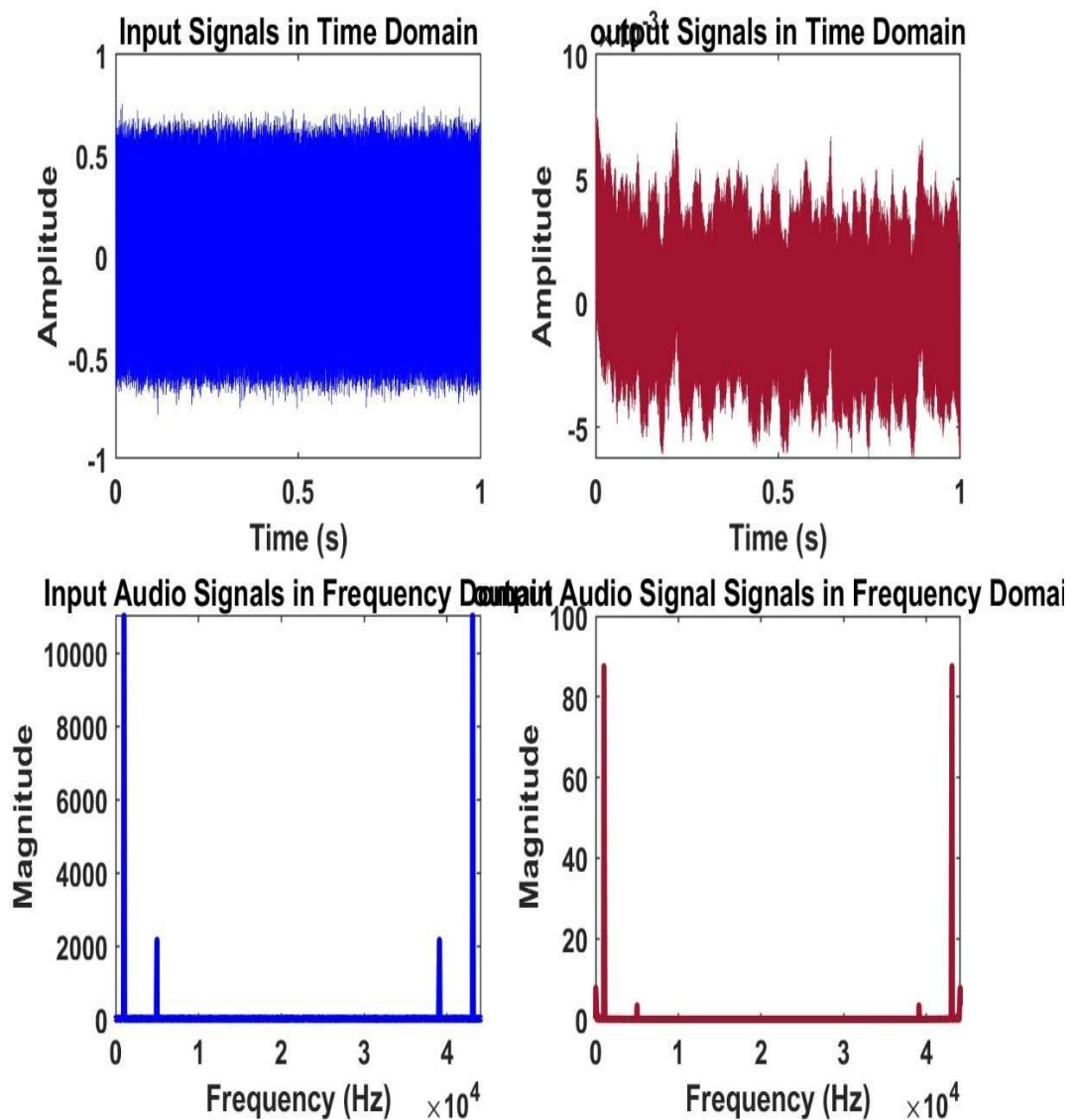


Figure 6.13 The output response of active EMI filters time and Frequency domain

Power that is reflected back to the source via an EMI filter, rather than being absorbed or transmitted, is termed return loss. Return loss generally denotes the efficacy with which a line or device aligns with its load, commonly referred to as impedance. Poor return loss in the filter may lead to issues with signal integrity, electromagnetic interference (EMI), and electromagnetic compatibility (EMC). To enhance filter performance, return loss measurements assist in diagnosing and refining the filter design.

Table 6.2 Parameters comparison of Active and Passive EMI filters

EMI Filter	Insertion Loss (dB)	Return Loss (dB)	Gain
Active	1.5836	20	0.7359
Passive	1.9382	20	0.912

The above table (table 6.2) presents the performance of active and passive EMI filters. The performance of the filters is determined using equations. Insertion loss and return loss are critical performance metrics for EMI filters. Insertion loss and return loss are quantified differently based on the filter's characteristics. Return loss is quantified as the power reflected towards the source due to impedance mismatches. Insertion loss quantifies the reduction of signal power as it traverses the filter. The insertion losses of the active and passive EMI filters (1.5836 dB and 1.9382 dB, respectively) are quite low, enabling significant attenuation of EMI signals while preserving the intended signal. The return loss of the EMI filters is substantial, indicating that a majority of the signal is sent down the targeted channel while a little amount is reflected.

EMI filters and other electronic equipment experience thermal noise, an inescapable attribute of electronics. While it may increase the noise floor and impair signal quality, a judiciously chosen component and design might alleviate its impact. Enhancing the performance and safeguarding the integrity of EMI filters may be achieved by comprehending and addressing thermal noise. Thermal noise or thermal interference from active components and resistors generates Interference in active and passive EMI filters is determined by the following equation (6.4), where T represents absolute temperature (Kelvin), k denotes Boltzmann's constant (1.38×10^{-23} J/K), R signifies resistance (Ω), and B is the bandwidth in Hz. The thermal noise of both active and passive EMI filters is roughly 190 nV.

$$V_{n,therm} = \sqrt{4kTRB} \quad (6.4)$$

The transmission of an electromagnetic wave across a material can disclose the medium's texture. By filtering, we may ascertain the medium's response to various

frequencies. Electromagnetic waves are utilized to examine surfaces in these systems, and filtering can be applied to distinguish textures. By employing filters, one may ascertain surface roughness and material qualities through the analysis of reflection patterns and scattering. Characteristics of the emitter is feasible to ascertain the attributes of any signal-emitting equipment, such as a speaker, antenna, or other signal sources, by analysis of their transmitted signals.

Signals are filtered to extract specific information with filters. Electronic gadgets frequently encounter disruption from outside electromagnetic impulses. EMI filters diminish these signals to prevent interference with their operation. The polarization of an electromagnetic wave can influence the efficacy of these filters. EMI filters may be selectively refined by the combination of polarizers and wave plates. Modern EMI filters include electronic control methods to automatically modify their filtering settings according to varying polarization states.

The model developed in MATLAB is executed on a hardware platform known as a Field Programmable Gate Array (FPGA) board. FPGA boards comprise input/output (I/O) blocks, configurable logic blocks (CLBs), programmable switches, memory, registers, and many peripheral devices. FPGAs facilitate the execution of intricate algorithms for real-time signal and image processing. A MATLAB HDL Coder may transform an EMI filter design for audio signals into a digital format suitable for FPGA implementation. HDL code generated from MATLAB is utilized for FPGA-based hardware development. Xilinx Vivado serves as a synthesis tool. The HDL code is derived from MATLAB and executed on an FPGA board with Vivado software. The filter is implemented on the Artix 7 Nexys4 (or xa7a100t csg324 -1) FPGA board, utilizing constraint files to enable the integrated clock cycle before the filter implementation.

Filter design is executed with constraint files and .vhd files derived from the .m file. The filter is emulated via Vivado 2020.1, coded in VHDL. Table IV presents the input and output values of the design. The simulation outcome of the active filter is illustrated in Figure 6.14. Figure 6.15 illustrates the power report of the executed filter design. The aggregate on-chip power of the suggested architecture is 0.146 W, including dynamic power of 0.049 W and static power of 0.097 W, respectively.

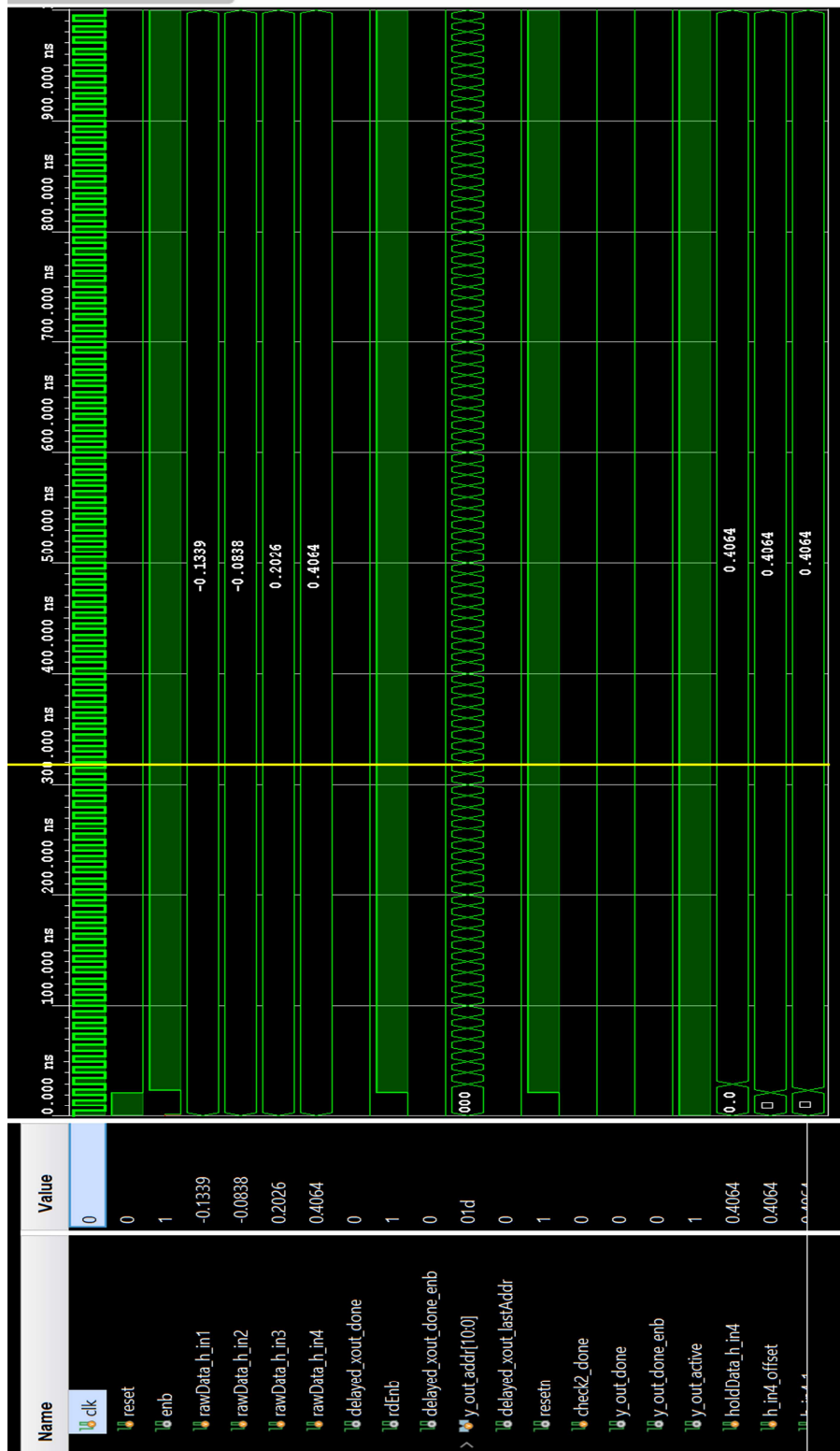


Figure 6.14 Output waveform of filter design

Power analysis from Implemented netlist. Activity derived from constraints files, simulation files or vectorless analysis.

Total On-Chip Power: 0.146 W
Design Power Budget: Not Specified
Power Budget Margin: N/A
Junction Temperature: 25.7°C
 Thermal Margin: 59.3°C (12.9 W)
 Effective θ_{JA} : 4.6°C/W
 Power supplied to off-chip devices: 0 W
 Confidence level: Low

[Launch Power Constraint Advisor](#) to find and fix invalid switching activity

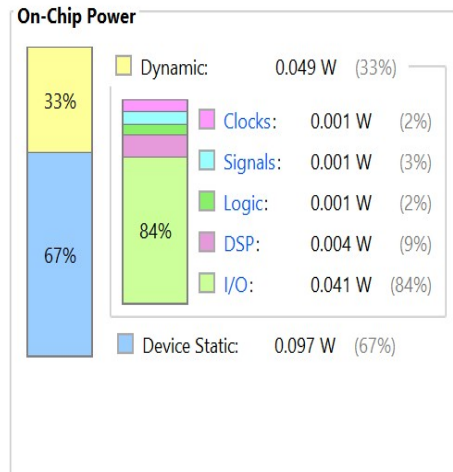


Figure 6.15 Power summary of filter design on FPGA

The results of the active filter simulation was derived from the test digital signal. The filter was subsequently synthesized and deployed, resulting in the generation of a bitstream file for implementation on the Artix 7 Nexys 4 series FPGA board. Figures 6.16 and 6.17 illustrate the device packaging and schematic of the filter design. The resultant pictures of the filter LEDs are displayed (6.18 & 6.19).

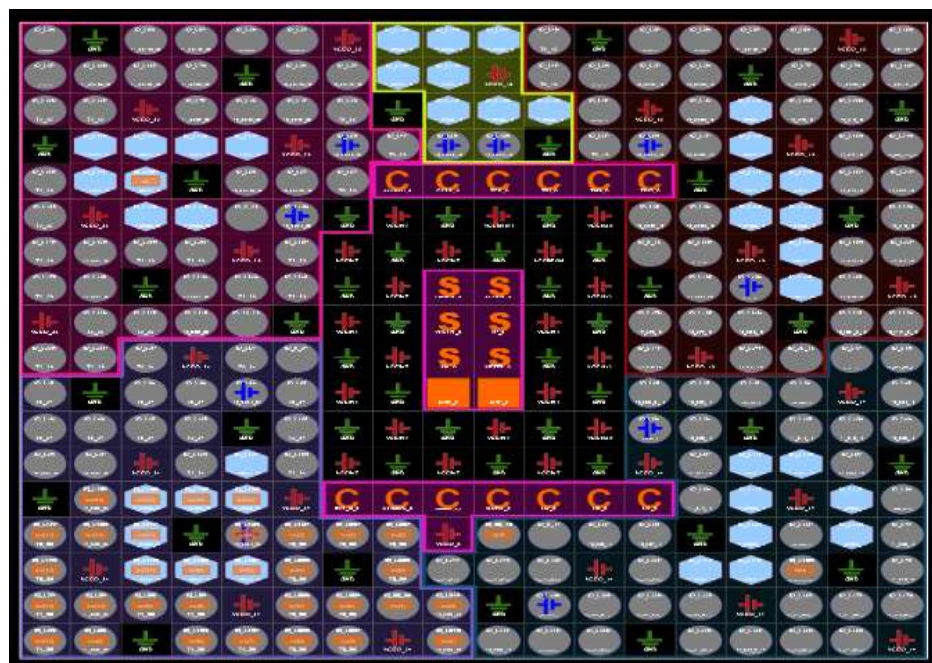


Figure 6.16 Device package

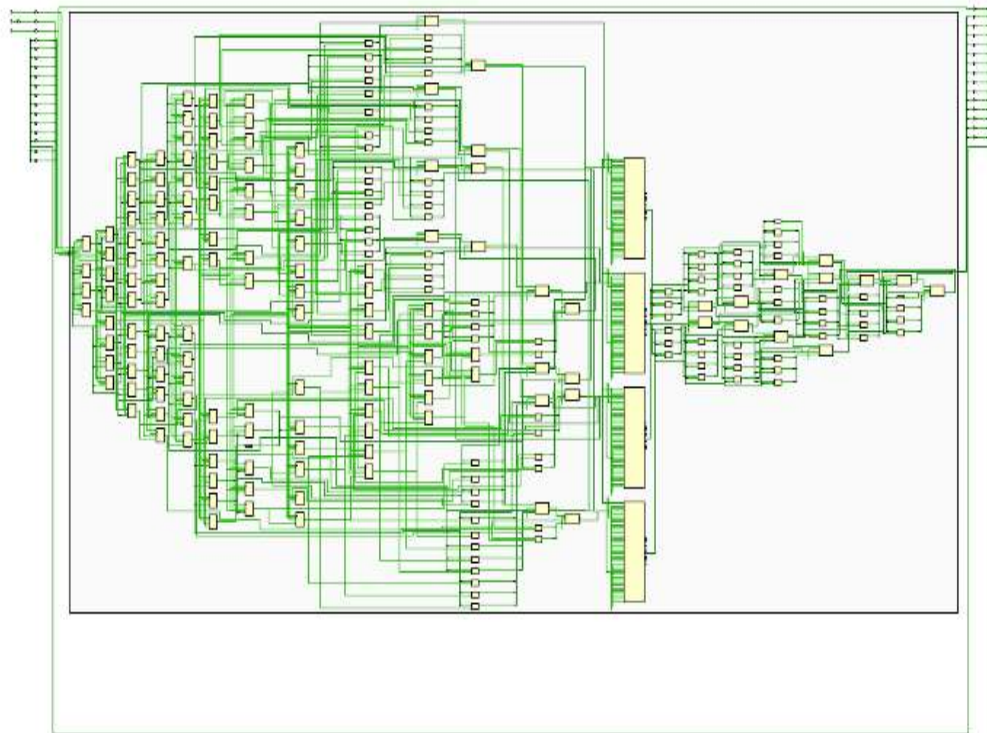


Figure 6.17 The schematic of filter design

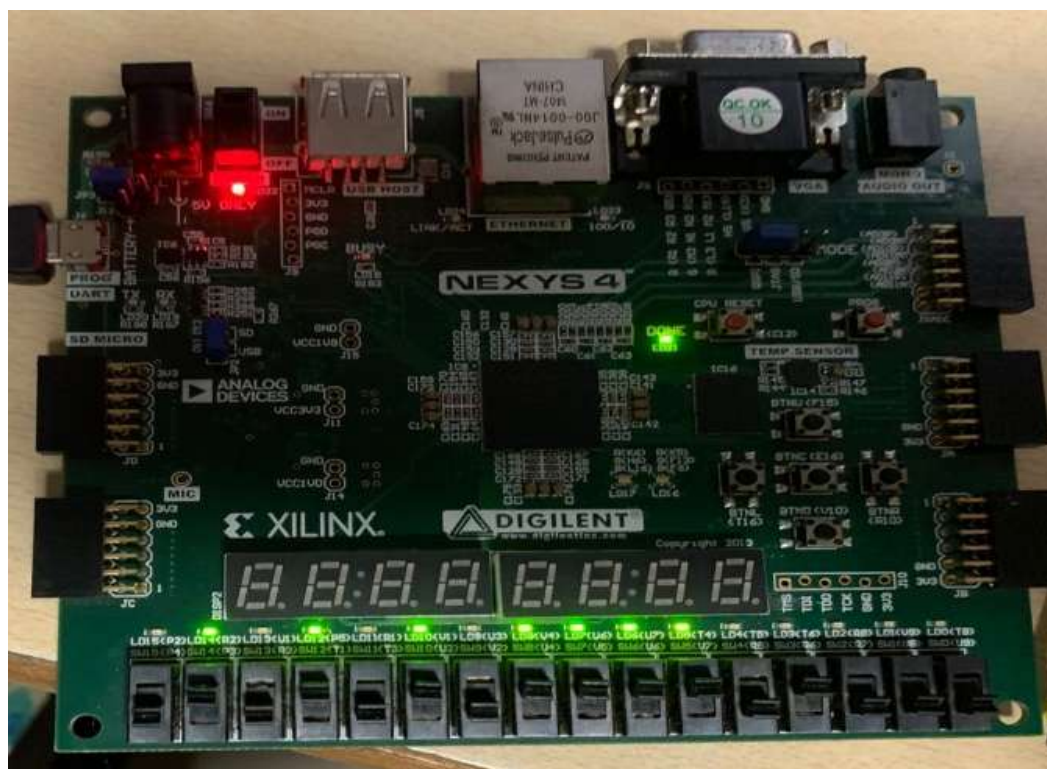


Figure 6.18 The output image of FPGA – delayed_xout

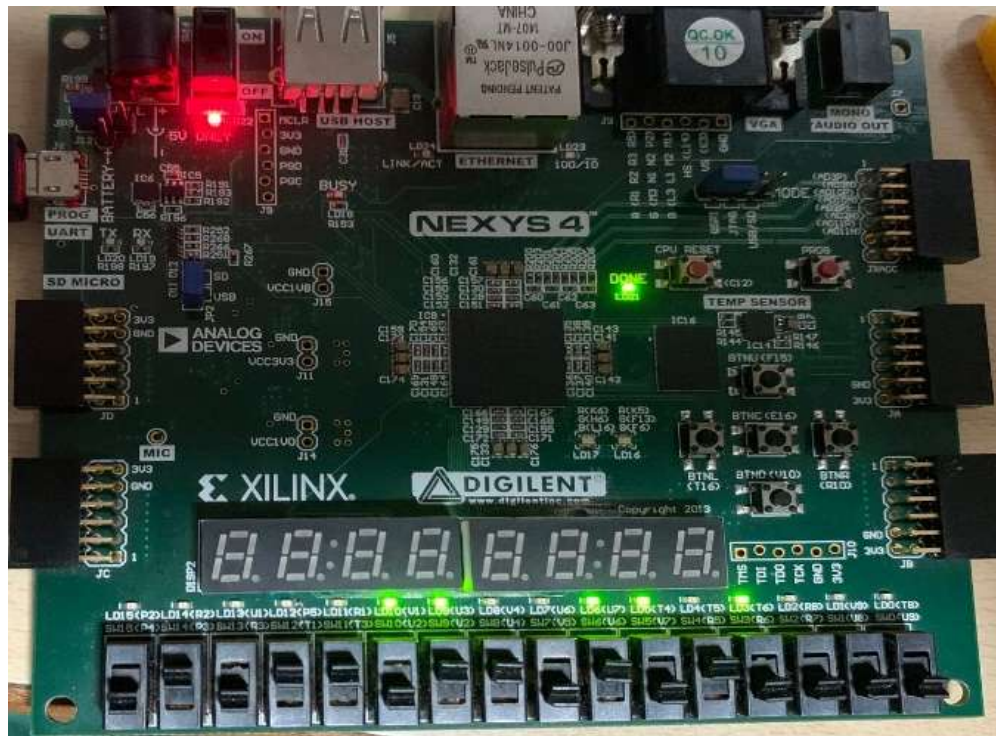


Figure 6.19 The output image of yout

Tables 6.3 and 6.4 present the utilization report of filter implementation. The utilization summary indicates that 98 LUTs, 112 flip-flops, 4 DSPs, 32 I/Os, and a slice count of 41 are employed for the implementation of the EMI filter. The EMI filters developed in MATLAB are transformed into fixed-point HDL code for simulation and evaluation on an FPGA board utilizing a digital signal. The resource use of EMI filter implementation on FPGA is presented in Table V, detailing the number of LUTs, flip-flops (FF), DSP blocks, and I/Os. The EMI filter will be further implemented for an analog signal utilizing an omnidirectional microphone included into the specified Xilinx Nexys 4 Artix 7 FPGA board.

The undesirable EMI signal impairs the device's functioning. Eliminating the noise signal will prevent performance degradation of the system. EMI filters assist in eliminating noise from transmissions. Active and passive EMI filters are engineered for noise reduction through software and hardware integration. EMI filters are engineered and deployed on the Artix 7 FPGA board. The filter is constructed with MATLAB R2020a and executed with Vivado 2020. 1.

Table 6.3 Utilization summary of FPGA implementation

Resource	Utilisation	Available	Utilisation (%)
LUT	98	63400	0.15
FF	112	126800	0.09
DSP	4	240	1.67
IO	32	210	15.24

Table 6.4 Slice and block counts

Design module	Slice LUTS	Slice Registers	Slice	LUT as logic	Block RAM tile	Blocked IOB	Bonded IPADs	BUFIO
Filter Top_module	98	112	41	98	112	4	32	1

Performance metrics are computed to guarantee device efficacy. The insertion loss of the active filter is 1.5836, whereas that of the passive filter is 1.9382. The return loss of the active and passive EMI filters is 20 dB and 20 dB, respectively. Attained minimal insertion loss while preserving elevated return loss for efficient noise reduction. The thermal noise or electromagnetic interference filter is 190 nV. The gains of active and passive filters are 0.7359 and 0.912, respectively. The EMI filter will be further implemented for an analog signal utilizing an omnidirectional microphone. The HDL code is produced via MATLAB. The aggregate on-chip power of the design is 0.146W. The dynamic power of the design is 0.049 W, while the static power of the

device is 0.097 W. The utilization data indicates that 98 LUTs, 112 flip-flops, 4 DSPs, and 32 I/Os are employed in the system architecture. This technique demonstrates reduced consumption of slice registers and LUTs compared to prior EMI filter implementations. In high-volume production contexts, when hardware expenses are paramount, a reduction in LUTs and slice registers might result in a more compact FPGA device, hence yielding cost reductions.

6.3 One vs All SVM 8x8 multi-class classifier

SVM multi-class classifier is designed for eight words to control the wheelchair. The performance metrics are calculated for SVM multiclass classifier. Obtained 8x8 confusion matrix and ROC of SVM multiclass classifier. For real-time audio signals on board microphone is used and utilization summaries are calculated. For SVM implementation MATLAB HDL coder is used to convert the MATLAB code to HDL language- Verilog. Comparative analysis of implementation of different blocks are mentioned. Each objective runtime and performance metrics of each class is calculated. SVM classifiers classify audio signals according to eight pre-defined categories, recognized as classes. Systems of multi-class classifiers are commonly employed in the recognition of speech commands, the detection of audio events, and the classification of music genres

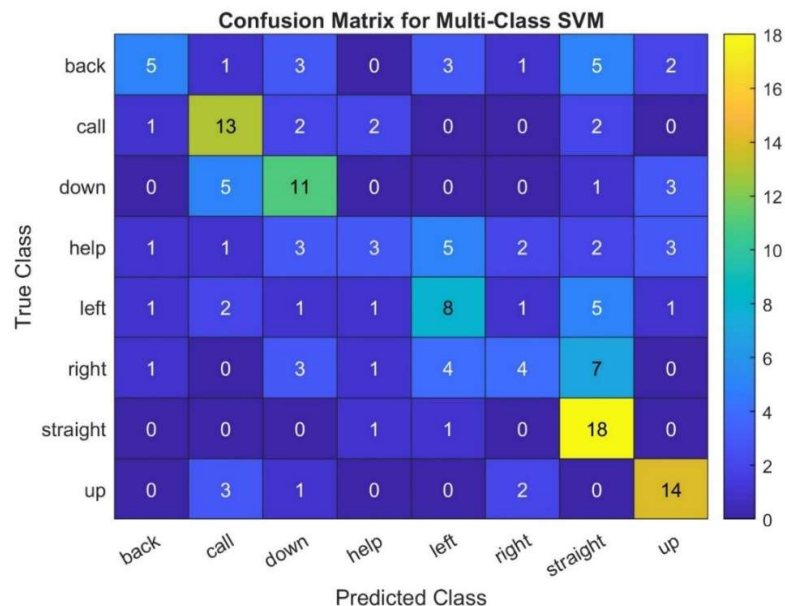


Figure 6.20 Confusion matrix for 20 testing sets for 800 audio clips

The SVM classifier is designed for eight classes, 800 audio clips are used to design and test the SVM multiclass classifier. The classifier has 8 words, which means eight classes. The predicted and true values of classes are mentioned in the matrix. 20 clips are given as testing samples/ dataset to the SVM module. Figure 6.20 shows the first trail confusion matrix of classifier, when dataset of the eight classes considering 800 audio clips.

All the audio clips are adjusted and saved in .wav format using Audacity software. Standardizing the audio format ensures consistency across all data, making it easier to process and analyze. With all clips at the same frequency and duration, it reduces variability that could affect model training and classification accuracy. Additionally, using a common format like .wav guarantees compatibility with various audio processing tools and libraries.

Iter	Eval	Objective	Objective	BestSoFar	BestSoFar	Coding	BoxConstraint	KernelScale
	result		runtime	(observed)	(estim.)			
1	Best	0.68906	13.918	0.68906	0.68906	onevsone	1.7231	0.012978
2	Best	0.425	9.1988	0.425	0.44048	onevsone	669.64	82.578
3	Accept	0.46719	3.5193	0.425	0.42538	onevsall	1.8026	394.87
4	Accept	0.48438	26.732	0.425	0.42501	onevsone	0.0037126	4.988
5	Accept	0.66563	19.447	0.425	0.42502	onevsall	1.7494	0.56183
6	Accept	0.51563	28.03	0.425	0.42502	onevsone	0.18294	682.53
7	Accept	0.51094	33.233	0.425	0.47919	onevsone	0.051291	32.702
8	Accept	0.5625	11.815	0.425	0.48648	onevsall	1.896	880.01
9	Accept	0.43594	34.99	0.425	0.43075	onevsone	879.37	81.731
10	Accept	0.51406	37.588	0.425	0.43101	onevsone	3.2914	141.79
11	Accept	0.45625	12.483	0.425	0.43098	onevsall	5.3028	233.27
12	Accept	0.58125	11.525	0.425	0.42801	onevsall	0.016047	194.86
13	Accept	0.43438	31.628	0.425	0.42926	onevsone	137.73	62.541
14	Accept	0.43125	32.163	0.425	0.42915	onevsone	260.48	118.62
15	Accept	0.45625	14.806	0.425	0.42947	onevsall	619.66	259.2
16	Accept	0.4375	24.059	0.425	0.42953	onevsall	895.21	83.514
17	Accept	0.44375	31.782	0.425	0.42956	onevsall	953.72	28.49
18	Accept	0.44375	14.072	0.425	0.42951	onevsall	46.989	86.782
19	Accept	0.45625	35.18	0.425	0.4293	onevsone	578.15	12.978

Figure 6.21 Runtime of each objective

```

Overall Accuracy: 85.00%
Class: back - Precision: 0.87, Sensitivity: 1.00, Specificity: 0.98, F1 Score: 0.93
Class: call - Precision: 0.88, Sensitivity: 0.70, Specificity: 0.99, F1 Score: 0.78
Class: down - Precision: 0.81, Sensitivity: 0.85, Specificity: 0.97, F1 Score: 0.83
Class: help - Precision: 0.67, Sensitivity: 0.70, Specificity: 0.95, F1 Score: 0.68
Class: left - Precision: 0.95, Sensitivity: 1.00, Specificity: 0.99, F1 Score: 0.98
Class: right - Precision: 0.86, Sensitivity: 0.95, Specificity: 0.98, F1 Score: 0.90
Class: straight - Precision: 0.88, Sensitivity: 0.75, Specificity: 0.99, F1 Score: 0.81
Class: up - Precision: 0.89, Sensitivity: 0.85, Specificity: 0.99, F1 Score: 0.87
>>

```

Figure 6.22 Overall Accuracy and each class performance metrics

Figure 6.21 shows the runtime of each objective in one vs all 8x8 multi class SVM for 20 test samples for each class, when the 1400 audio samples are considered. Figure 6.22 shows the overall accuracy of 8x8 multi-class SVM classifier, which is 85%, and also mentioned each class performance metrics. Class “back” has 0.87-precision, sensitivity- 1, specificity-0.98 and F1 score 0.93; for “call” class, it is 0.88, 0.70, 0.99, and 0.78 of precision, sensitivity, specificity and F1 score respectively. For class “down”, 0.81,0.85, 0.97, and 0.83 of precision, sensitivity, specific and F1 score. Class “help” has 0.67-precision, sensitivity-0.70, specificity-0.95 and F1 score 0.68; for “left” class, it is 0.95, 1, 0.99, and 0.98 of precision, sensitivity, specificity and F1 score respectively. For class “right”, 0.86, 0.95, 0.98, and 0.90 of precision, sensitivity, specific and F1 score. For class “straight” and “up”, precision is 0.88 and 0.89, sensitivity is 0.75 and 0.85, specificity is 0.99 and 0.99 and F1 score is 0.81 and 0.87 respectively. The people has a different accent from all the state of India, still I got the overall accuracy of 85%.

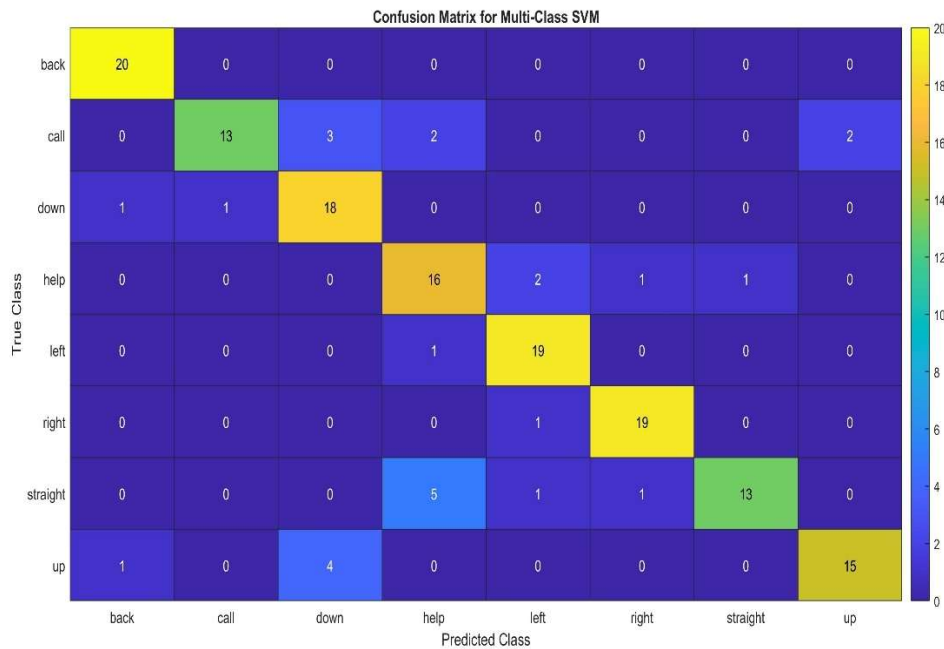


Figure 6.23 SVM confusion matrix of one vs all SVM multiclass classifier

Above figure 6.23, explains an 8x8 confusion matrix has rows and columns corresponding to the 8 different classes (words), 1400 audio clips are considered as dataset. 280 audio clips set as testing and 1120 .wav files are given to train the model. When 20% data is feed as testing to the multiclass classifier, class “back” detected all the words. Class “left” and “right” also detected almost. Each entry in the matrix represents the number of instances where the classifier predicted one class when the true class was another. Specifically: Rows correspond to the actual (true) classes. and Columns correspond to the predicted classes. Each element at position (i, j) in the matrix tells you how many instances of class i (actual class) were predicted as class j (predicted class). For an 8x8 matrix, you have 8 possible classes for both the actual and predicted values, so the matrix will contain 64 entries in total. Figure 6.24 shows the observed and estimated feasible points presented in command prompt.

```
Best observed feasible point:
      Coding      BoxConstraint      KernelScale
      _____      _____      _____
onevsone      669.64      82.578

Observed objective function value = 0.425
Estimated objective function value = 0.4383
Function evaluation time = 18.7268

Best estimated feasible point (according to models):
      Coding      BoxConstraint      KernelScale
      _____      _____      _____
onevsall      0.63508      18.837

Estimated objective function value = 0.44025
Estimated function evaluation time = 11.101|
```

Figure 6.24 Evaluation time of one vs all 8x8 multiclass classifier

In an 8x8 multi-class classifier, performance metrics such as accuracy, precision, recall, and F1-score provide insight into how the classifier distinguishes between the 8 different classes (e.g., words). Accuracy is the overall percentage of correct predictions across all classes and is calculated by dividing the sum of true positives (diagonal elements) by the total number of predictions. Precision measures how many of the instances predicted as a certain class were actually correct, focusing

on minimizing false positives for that class. Recall (or sensitivity) evaluates how well the classifier identifies all true instances of a class, minimizing false negatives. The F1-score is the harmonic mean of precision and recall, providing a balanced measure when precision and recall differ. For each class, these metrics are computed individually (using a One-vs-Rest approach) and can also be averaged across all classes using macro-averaging (treating each class equally) or micro-averaging (weighted by the number of instances per class). These metrics offer a comprehensive evaluation of the classifier's ability to distinguish between all 8 words in the dataset, revealing strengths and weaknesses in both general accuracy and class-specific performance.

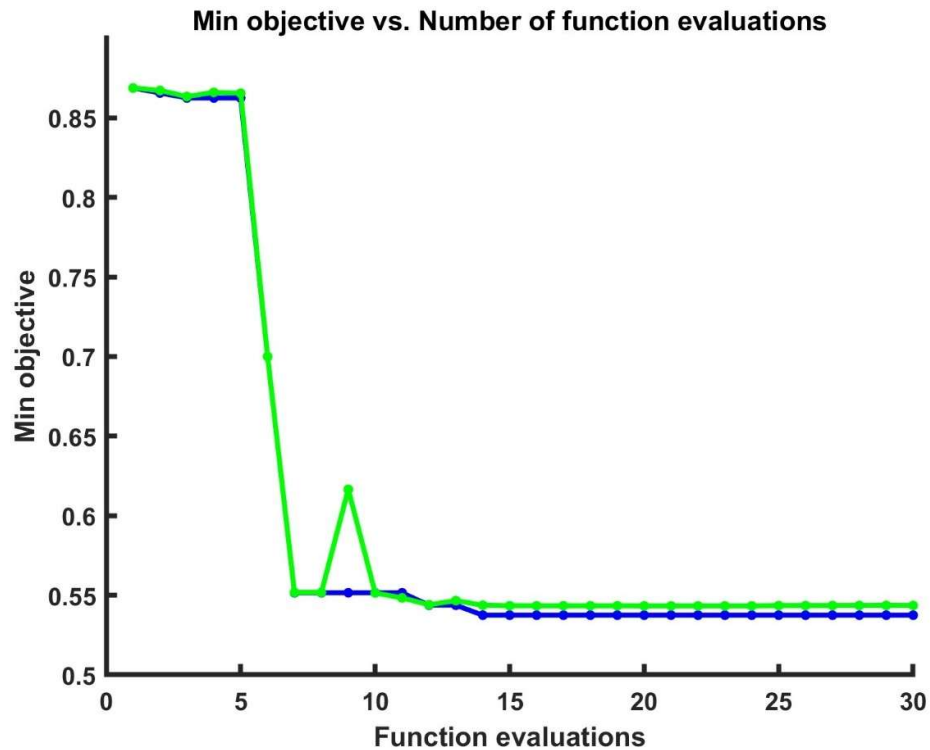


Figure 6.25 Minimum objective vs Number of function Evaluation graph

ROC curves are commonly used in medical diagnostics to evaluate the performance of tests and determine optimal thresholds for disease detection. In machine learning, they are used to assess the discriminatory power of classification models.

Figure 6.25 shows the graph of evaluation to minimum objectives. Additionally, ROC curves are applied in credit scoring to analyze the ability of models to distinguish between good and bad credit risks. ROC curves are constructed by plotting TPR against FPR for each class and varying the decision threshold (figure 6.26). As a result of different thresholds, different pairs of TPR and FPR will result, which is plotted to form the ROC curve.

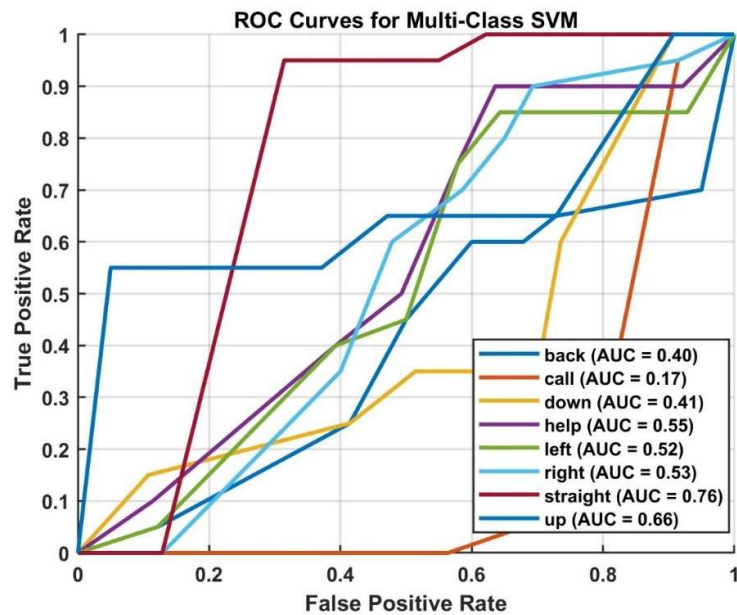


Figure 6.26 ROC curve for 8x8 metaclassifier

6.4 VIVADO Results

The Vivado design suite provides essential insights into the operation and resource use of each functional block of the TDS built on the Xilinx Nexys 4 Artix-7 FPGA through waveform simulations, RTL schematics, and utilization summaries. These results ensure the system's functional accuracy, optimal synthesis, and compliance with hardware constraints at every block of the TDS architecture, from audio capture to PWM control.

6.4.1 Audio acquisition

The waveform output displays a bitstream at the input (micData) and clock (micClk) more refined PCM signal at the output. The Audio Acquisition Block

employs a CIC filter to convert PDM data from the ADMP421 microphone into PCM. Figure 6.27 shows the waveform of PDM data, PCM data, microphone clock (micClk). It ensures that the CIC filter is appropriately decimating and filtering the signal. Figure 6.28 shows the schematic diagram of PDM to PCM conversion with logic analyzer. The utilization report for this block indicates little LUT and register usage, with no consumption of DSP or BRAM, hence confirming the hardware efficiency of CIC filters.

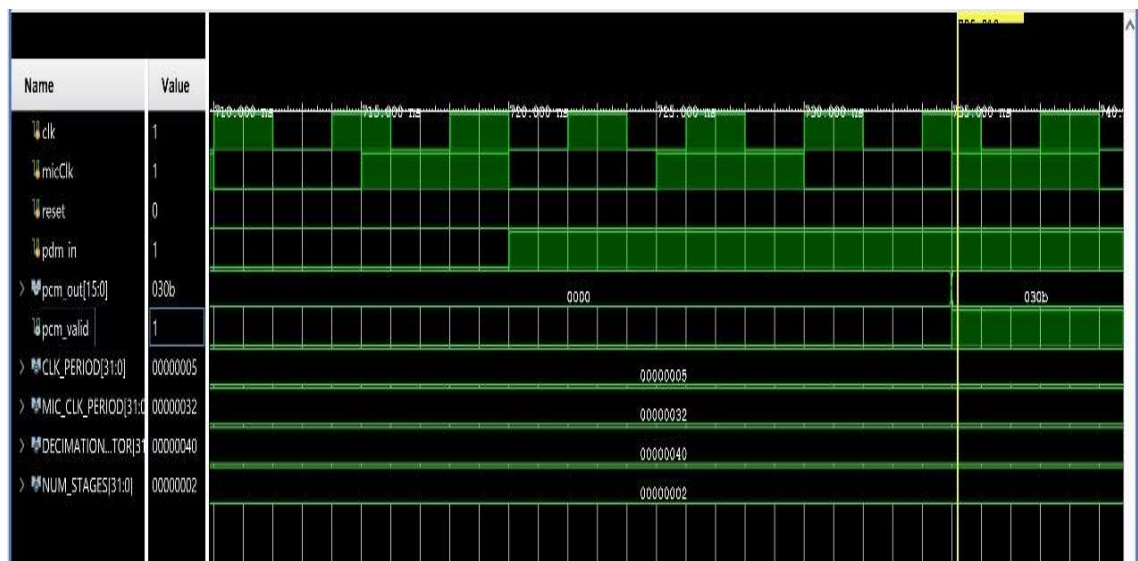


Figure 6.27 Audio acquisition block with CIC filter and PDM-to-PCM Converter

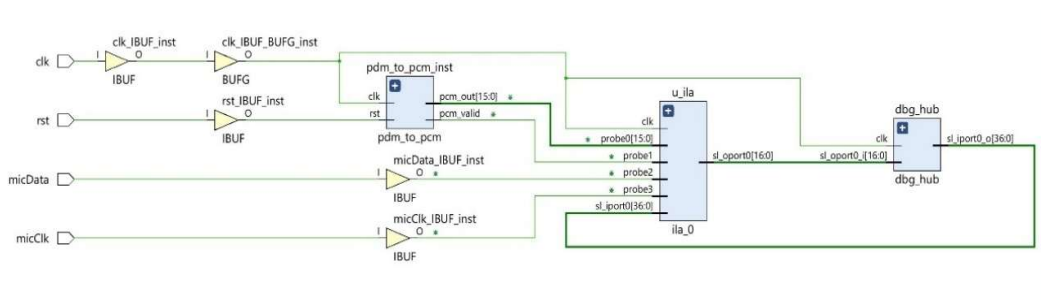


Figure 6.28 Schematic PDM to PCM conversion

The power utilization of audio acquisition is given below figure 6.29. It shows 13% dynamic power and 87% static power.

Summary

Power analysis from Implemented netlist. Activity derived from constraints files, simulation files or vectorless analysis.

Total On-Chip Power: 0.097 W
Design Power Budget: Not Specified
Power Budget Margin: N/A
Junction Temperature: 25.4°C
Thermal Margin: 59.6°C (12.9 W)
Effective θ_{JA} : 4.6°C/W
Power supplied to off-chip devices: 0 W
Confidence level: Medium
[Launch Power Constraint Advisor](#) to find and fix invalid switching activity

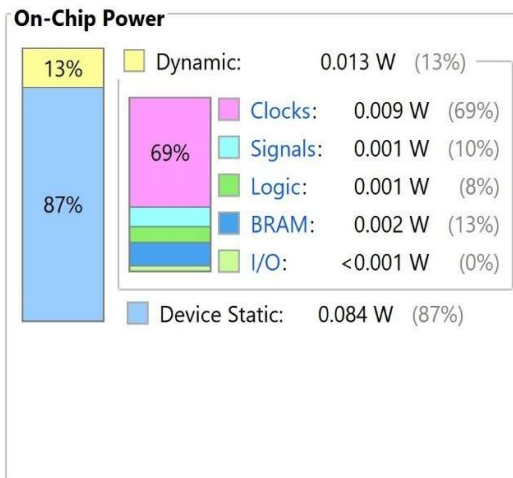


Figure 6.29 Audio acquisition power summary

6.4.2 Passive Low-Pass EMI filter

The EMI filter block, Passive low-pass EMI filter is designed for the system. Figure 6.30 demonstrates noise attenuation in the filtered output (pcmData_out) and input signal (pcmData_in) and clock signal. The RTL diagram confirms the appropriate FIR layout, displaying several tap delay elements connected to coefficient multipliers and accumulators. Utilizing HDL-generated filters from MATLAB may result in a substantial use of DSP slices for MAC operations, along with a little increase in registers and LUTs, contingent upon the filter length and precision. Figure 6.31 and 6.32 shows the schematic and power utilization of EMI filter.



Figure 6.30 EMI waveforms

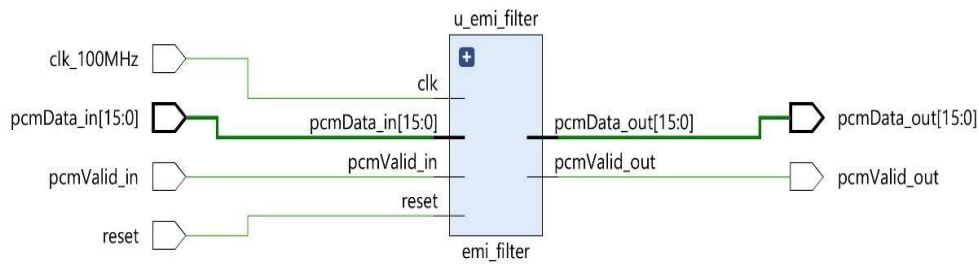


Figure 6.31 Schematic of EMI filter

Summary

Power analysis from Implemented netlist. Activity derived from constraints files, simulation files or vectorless analysis.

Total On-Chip Power: **0.182 W**
Design Power Budget: **Not Specified**
Power Budget Margin: **N/A**
Junction Temperature: **25.8°C**
Thermal Margin: 59.2°C (12.8 W)
Effective θ_{JA} : 4.6°C/W
Power supplied to off-chip devices: 0 W
Confidence level: **Low**
[Launch Power Constraint Advisor](#) to find and fix invalid switching activity

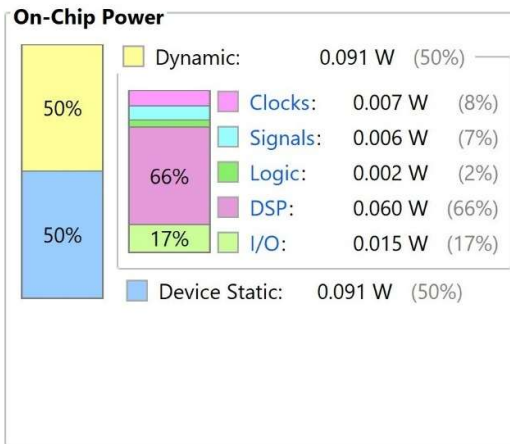


Figure 6.32 Power utilization summary for EMI filter

6.4.3 Feature Extraction

The Feature Extraction block produces properties such as spectral entropy, zero-crossing rate (ZCR), and MFCC from filtered audio. The simulated waveforms exhibit periodic feature creation of each frame, every 10 milliseconds, mirroring the dynamic characteristics of the speech signal. The RTL diagram illustrates modules for each computation, RMS through squaring and summation, ZCR via sign bit comparison, and entropy through logarithmic calculations. This block significantly utilizes DSPs for mathematical operations and consumes a moderate quantity of registers and LUTs as anticipated. Frame-level data can also be buffered with BRAMs.

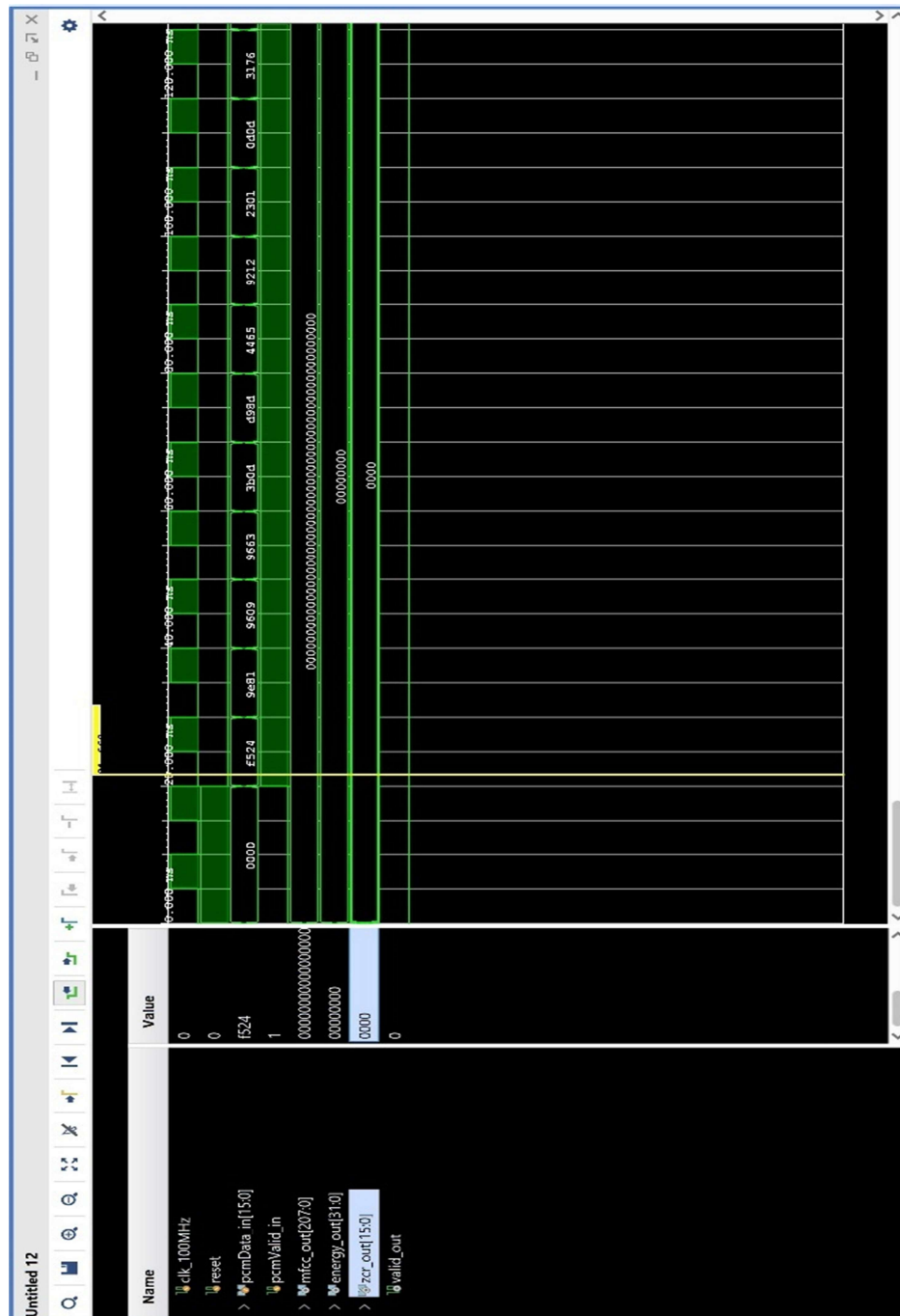


Figure 6.33 Waveforms of Feature extraction

Figure 6.33, 6.34 and 6.35 explains the waveforms of different feature extractions and Power summary and schematic for classifying the speech command.

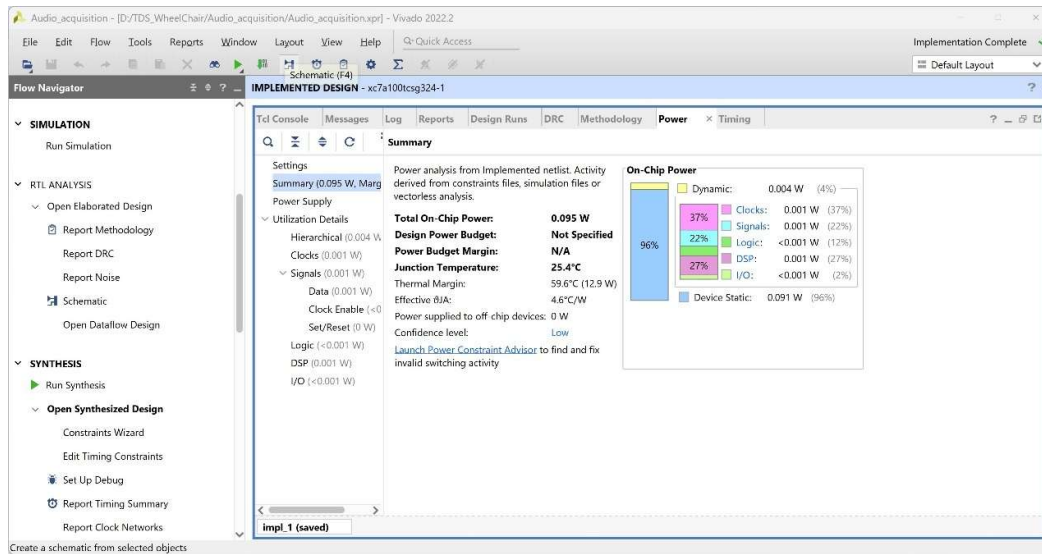


Figure 6.34 Power utilization summary of feature extractions

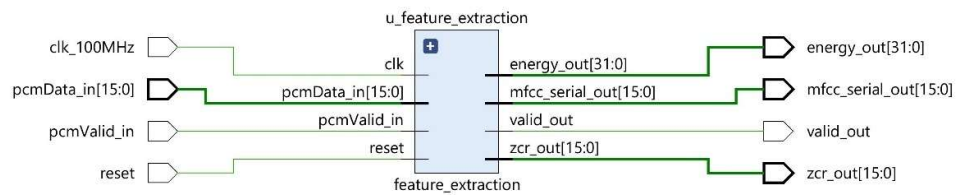


Figure 6.35 schematic of feature extraction

6.4.4 Multiclass one vs all SVM classifier

The waveform output for the SVM classifier, utilizing a One-vs-All technique, typically consists of 3-bit values and is presented once every frame, yielding discrete classification results. Figure 6.36 illustrates the waveforms generated from SVM module. This block is resource-intensive because to the extensive use of DSP slices for dot product computations, particularly when feature dimensions or class counts increase. LUTs and registers are employed with moderate frequency; weight vectors may be retained in BRAM. Figure 6.37 shows power utilization.

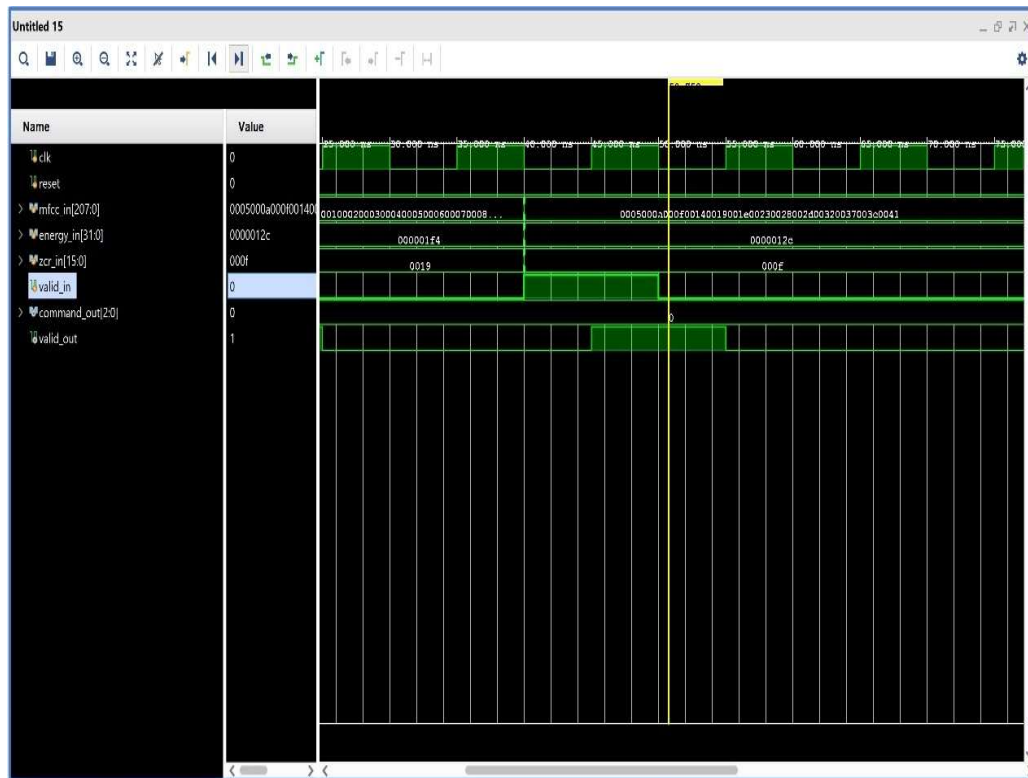


Figure 6.36 waveforms generated for SVM classifier

Summary

Power analysis from Implemented netlist. Activity derived from constraints files, simulation files or vectorless analysis.

Total On-Chip Power: 0.354 W
Design Power Budget: Not Specified
Power Budget Margin: N/A
Junction Temperature: 26.6°C
 Thermal Margin: 58.4°C (12.7 W)
 Effective θ_{JA} : 4.6°C/W
 Power supplied to off-chip devices: 0 W
 Confidence level: Low
[Launch Power Constraint Advisor](#) to find and fix invalid switching activity

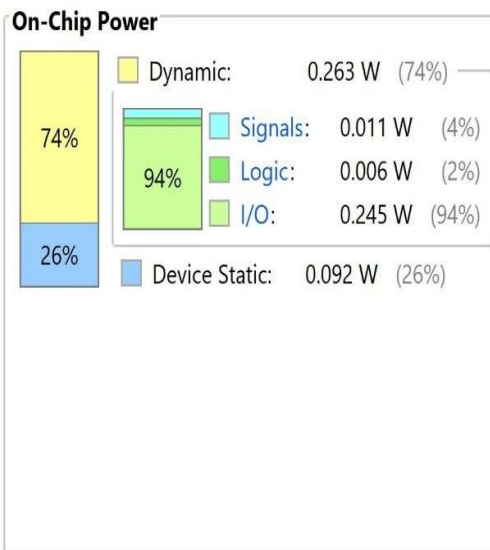


Figure 6.37 power summary for SVM classifier

6.4.5 PWM generation

The TDS included into the Xilinx Nexys 4 Artix-7 FPGA generates PWM signals specifically to control the L298N motor driver, which operates the wheelchair's wheels based on classified spoken commands. The L298N is a dual H-bridge motor driver that receives two PWM signals for each motor channel, with two logic inputs to regulate direction and speed. The PWM generating block is tailored to deliver command-specific PWM signals that precisely satisfy the control requirements of the L298N. Figure 6.38 Generates the signals according to the command for In1, In2, In3 and In4 control signals and two enable signals to activate the DC motor.

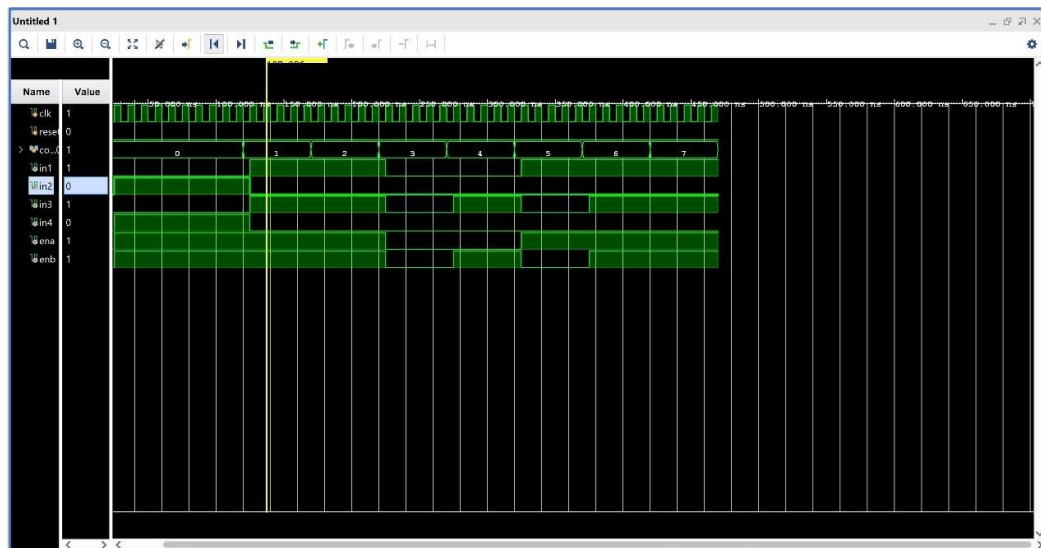


Figure 6.38 Control signals are generated for L298N motor driver.

Chapter 7

Conclusion and Future Scope

In recent years, the development of technology made easy for the disabled and able people to work smartly and making them more independent. Especially, for the disabled people it will be difficult to depend on others for mobility and daily necessities. As accidents happen, aging of people and medical issues like Spinal cord Injuries (SPI) increased the number of people made it difficult to do normal work and mobility. The technology and Devices which helps the abled and disabled people known as AT and AD. Tongue is the least effected organ in the human body. Hence, developing the AT using tongue movement or speech is a best choice for the disabled people. Using magnetic sensors, ultrasound, capacitive, electromyographic, and optical sensors are a bit expensive, and maximum of sensors has to be pierced or mounted inside the tongue. Which makes uncomfortable to use, unhygienic – needs to clean time to time, and saliva makes the AD bit inaccurate. The best way is to assist the device by avoiding these issues are using microphone sensors. Which is cheap, no need of piercing and accuracy also be not affected.

The design and implementation of a speech-based TDS for wheelchair control are significant advancements in the pursuit of independent living and mobility for individuals with severe motor impairments. This thesis effectively showcased the viability of combining real-time speech processing with FPGA technology to operate a motorized wheelchair using Speech commands or instructions. The system utilized the onboard MEMS microphone (ADMP421) for audio input, so that we don't need to use external microphone sensors, followed by a comprehensive pipeline of preprocessing, filters, feature extraction, and classification of commands using Support vector machine (SVM) model, then converting recognized commands into Pulse Width Modulated (PWM) signals for the wheelchair using L298N motor driver.

The dataset is collected from 175 people of different genders, a total of 1400 audio clips are collected for eight commands. The audio clips are normalized and trimmed using AUDACITY application. These audio clips are converted to .wav files. These audio clips are mono channel, signed 16-bit PCM with 48000 sample rate.

Utilizing the benefits of FPGAs low latency, and reconfigurability for the AT to attain real-time performance appropriate for actual assistive applications. The speech-based architecture comprised several phases: audio acquisition by ADMP421 microphone, CIC filter is used for converting PDM-to-PCM signal, Electromagnetic Interference (EMI) filtering for noise reduction, feature extraction employing spectral and temporal methods, and classification with a One-vs-All SVM classifier. Each phase was evaluated separately and included into a primary module that effectively operated the wheelchair by eight specified commands: "up," "down," "left," "right," "back," "straight," "call," and "help."

EMI filters are designed and implemented using both active and passive components for noise cancellation in MATLAB. Verilog code is developed using MATLAB HDL coder application from MATLAB function. Calculated the insertion loss, gain and return loss of both active and passive EMI filters. The results showed that Insertion Loss of 1.5836 dB and 1.9382 dB, for active and passive EMI filters, respectively. Return loss for both active and passive filters are 20 dB. 0.7359 gain for active filter and 0.912 for passive filters. The top level of EMI filters for lowpass passive filter utilized 98 LUTs, and 112 slice registers.

The classification of audio signals is designed for all possible machine learning algorithms. The algorithms are Stochastic Gradient Descent (SGD), Support Vector Machine (SVM), Random Forest (RF) and Naïve Bayes (NB) are designed and compared the performance. SVM generated 91% accuracy and other algorithms NB, RF and SGD, 84.223%, 87% and 93%, respectively. F1 score and sensitivity for SVM is high compared to other algorithms, which are 0.7407 and 0.7884, respectively. By comparing all the performance metrics SVM gave better performance.

SVM classifier is designed for eight commands by using one vs All method using 85% of the data to train, which means 1120 and 20% of data to test the SVM algorithm. Performance metrics of one vs all SVM eight class classifier is calculated. ROC is drawn for all eight classes. SVM module is converted to HDL language using MATLAB HDL coder. The FPGA implementation shown exceptional efficacy in managing real-time processing limitations while maintaining precision and resilience in noisy conditions. The design emphasized power optimization and hardware economy, rendering it suitable for portable, battery-operated applications such as smart

wheelchairs. Moreover, the implementation of a hardware-based approach guaranteed predictable timing, which is crucial for key applications concerning user safety and accessibility. It can provide significant enhancements in the quality of life for individuals dependent on assistive mobility devices. The flexible and adaptable characteristics of the speech-based TDS render it especially advantageous for users with diverse physical restrictions and various language competencies.

Tongue Drive System on FPGA processors using speech, by employing optimal MLCs, which led to reduced power consumption, lower noise levels, and improved accuracy. The primary objective of the thesis is classification of speech for AD and implement on FPGA, however, the international research collaborations provided significant technical depth and concepts. The FPGA implementation of EEG seizure detection enhanced understanding of hardware-efficient machine learning, while CNN-based tongue contour extraction provided valuable insights into speech analysis. This collaboration did not establish distinct objectives, it enhanced the practical relevance and technical robustness of the proposed system.

To enhance the accuracy and reliability of the speech-based Tongue Drive System on FPGA, additional input signals, such as tongue movements or EEG, may be incorporated. Future development may concentrate on developing TDS on more configurable and advanced FPGA to reduce more power consumption and fast data processing, privacy, easier reconfiguration and access for different users. This will facilitate the portability and wearability of the system. Adaptive learning techniques can enhance the system efficacy in interacting with each user progressively. Overall, these modifications can enhance their efficacy as a real-time AD for individuals with speech or mobility impairments.

Future research may improve current international collaboration by integrating multimodal techniques, such as EEG, ultrasound tongue imaging, and speech signals, into a integrated FPGA-based TDS. International collaboration remains crucial for evaluating and enhancing these systems for application in practical AT. The current implementation is a laboratory prototype utilizing an FPGA development platform with an integrated microphone. For the future development, the design will be translated into a small embedded module, with the processing chip mounted into the wheelchair to provide useful real-time assistance device.

References

- [1] World Health Organization: WHO, “Disability,” Mar. 07, 2023. <https://www.who.int/news-room/fact-sheets/detail/disability-and-health>
- [2] S. Alper and S. Raharinirina, “Assistive Technology for Individuals with Disabilities: A Review and Synthesis of the Literature,” *Journal of Special Education Technology*, vol. 21, no. 2, pp. 47–64, Mar. 2006, doi: 10.1177/016264340602100204.
- [3] S. Singh, M. Tomar, V. S. Mahendra, and Amaltas, “Rapid Review of Assistive Technologies for Persons with Disability in India,” 2020. [Online]. Available: https://assets.publishing.service.gov.uk/media/5ed6321bd3bf7f4606f1e15b/DFID-AT_Report.pdf
- [4] M. Grimmer, R. Riener, C. J. Walsh, et al., “Mobility related physical and functional losses due to aging and disease – a motivation for lower limb exoskeletons,” *J. Neuroeng. Rehabil.*, vol. 16, p. 2, 2019, doi: 10.1186/s12984-018-0458-8.
- [5] G. Shi, S. Ke, and A. Banozic, “The role of assistive technology in advancing sustainable development goals,” *Front. Polit. Sci.*, vol. 4, p. 859272, 2022, doi: 10.3389/fpos.2022.859272.
- [6] L. Nierling, M. Maia, L. Hennen, G. Wolbring, T. Bratan, P. Kukk, J. Cas, L. Capari, J. Krieger-Lamina, and E. Mordini, “Assistive technologies for people with disabilities – Part II: Current and emerging technologies,” 2018, doi: 10.2861/567013.
- [7] A. M. Alshabeb and O. Alharbi, “Critical analysis of the benefits and drawbacks of assistive technology with special needs”, *ASSRJ*, vol. 6, no. 8, pp. 210–215, Aug. 2019.
- [8] E. M. Orellano-Colón, F. M. Morales, Z. Sotelo, N. Picado, E. J. Castro, M. Torres, M. Rivero, N. Varas, and J. Jutai, “Development of an assistive technology intervention for community older adults,” *Phys. Occup. Ther. Geriatr.*, vol. 35, no. 2, pp. 49–66, 2017, doi: 10.1080/02703181.2017.1283656.

- [9] J. Mensah-Gourmel, M. Thépot, J. W. Gorter, M. Bourgain, C. Kandalaf, A. Chatelin, G. Letellier, S. Brochard, and C. Pons, “Assistive products and technology to facilitate activities and participation for children with disabilities,” *Int. J. Environ. Res. Public Health*, vol. 20, no. 3, p. 2086, Jan. 2023, doi: 10.3390/ijerph20032086.
- [10] D. A. Oladele, E. D. Markus, and A. M. Abu-Mahfouz, “Adaptability of assistive mobility devices and the role of the Internet of Medical Things: Comprehensive review,” *JMIR Rehabil. Assist. Technol.*, vol. 8, no. 4, p. e29610, Nov. 2021, doi: 10.2196/29610.
- [11] K. Manjari, M. Verma, and G. Singal, “A survey on Assistive Technology for visually impaired,” *Internet of Things*, vol. 11, p. 100188, Mar. 2020, doi: 10.1016/j.iot.2020.100188.
- [12] S. Zgonec and D. Bogataj, “Assistive technologies supporting the independence of elderly adults with intellectual disability: Literature review and research agenda,” *IFAC-PapersOnLine*, vol. 55, no. 39, pp. 129–134, Jan. 2022, doi: 10.1016/j.ifacol.2022.12.023.
- [13] P. Muthu, Y. Tan, S. Latha, S. Dhanalakshmi, K. W. Lai, and X. Wu, “Discernment on assistive technology for the care and support requirements of older adults and differently-abled individuals,” *Front. Public Health*, vol. 10, p. 1030656, 2023, doi: 10.3389/fpubh.2022.1030656.
- [14] B. Dicianno, R. Cooper, and J. Coltellar, “Joystick control for powered mobility: Current state of technology and future directions,” *Phys. Med. Rehabil. Clin. N. Am.*, vol. 21, pp. 79–86, 2010, doi: 10.1016/j.pmr.2009.07.013.
- [15] Forgrave, Karen E. “Assistive Technology: Empowering Students with Learning Disabilities.” *The Clearing House*, vol. 75, no. 3, 2002, pp. 122–26. JSTOR, <http://www.jstor.org/stable/30189719>. Accessed 18 Feb. 2026.
- [16] J. Kim, H. Park, and M. Ghovanloo, “Tongue-operated assistive technology with access to common smartphone applications via Bluetooth link,” in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, 2012, pp. 4054–4057, doi: 10.1109/EMBC.2012.6346857.

- [17] S. K. Fager, T. Sorenson, E. Jakobs, H. Koester, T. Jakobs, and D. R. Beukelman, "Multi-modal access method (eye-tracking + switch-scanning) for individuals with severe motor impairment: A preliminary investigation," *Assist. Technol.*, vol. 35, no. 4, pp. 321–329, Jul. 2023, doi: 10.1080/10400435.2022.2053895.
- [18] D. Faris-Cole and R. Lewis, "Exploring speech recognition technology: Children with learning and emotional/behavioral disorders," *Learning Disabilities: A Multidisciplinary Journal*, vol. 11, no. 1, pp. 3–12, 2001.
- [19] N. Gistituati, R. Refnaldi, and S. Syaifullah, "Using visual aids to develop students' speaking skill," *Proceedings of the 1st International Conference on Innovation in Education (ICoIE 2018)*, Jan. 2019, doi: 10.2991/icoie-18.2019.59.
- [20] J. Alkadri and J. Jutai, "Cognitive impairment and assistive devices: Outcomes and adverse effects," *J. Rehabil. Assist. Technol. Eng.*, vol. 3, p. 2055668316668146, Oct. 2016, doi: 10.1177/2055668316668146.
- [21] S. M. U. Sankar, D. Dhinakaran, T. Kavya, S. Priyanka and P. P. Oviya, "A Way for Smart Home Technology for Disabled and Elderly People," *2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, Uttarakhand, India, 2023, pp. 369-373, doi: 10.1109/ICIDCA56705.2023.10099817.
- [22] R. Simpson, H. H. Koester, and E. Lopresti, "Research in computer access assessment and intervention," *Phys. Med. Rehabil. Clin. N. Am.*, vol. 21, no. 1, pp. 15–32, Feb. 2010, doi: 10.1016/j.pmr.2009.07.006.
- [23] J. M. Fernández-Batanero, M. Montenegro-Rueda, J. Fernández-Cerero, and I. García-Martínez, "Assistive technology for the inclusion of students with disabilities: a systematic review," *Educational Technology Research and Development*, vol. 70, no. 5, pp. 1911–1930, Jun. 2022, doi: 10.1007/s11423-022-10127-7.
- [24] G. Burdea, N. Kim, K. Polistico, A. Kadaru, N. Grampurohit, D. Roll, and F. Damiani, "Assistive game controller for artificial intelligence-enhanced telerehabilitation post-stroke," *Assist. Technol.*, vol. 33, no. 3, pp. 117–128, May 2021, doi: 10.1080/10400435.2019.1593260.

- [25] M. Finocchiaro et al., “A Framework for the Evaluation of Human Machine Interfaces of Robot-Assisted Colonoscopy,” in *IEEE Transactions on Biomedical Engineering*, vol. 71, no. 2, pp. 410-422, Feb. 2024, doi: 10.1109/TBME.2023.3301741.
- [26] Lazcano-Herrera, Alicia Guadalupe, Rita Q. Fuentes-Aguilar, Isaac Chairez, Luz María Alonso-Valerdi, Miguel Gonzalez-Mendoza, and Mariel Alfaro-Ponce, “Review on BCI Virtual Rehabilitation and Remote Technology Based on EEG for Assistive Devices”, *Applied Sciences*, 2022, 12, no. 23: 12253. <https://doi.org/10.3390/app122312253>
- [27] Héctor A. Caltenco, Lotte NS Andreassen Struijk and Björn Breidegard, “TongueWise: Tongue-computer interface software for people with tetraplegia”, *2010 IEEE Annual International Conference of Engineering in Medicine and Biology Society (EMBC)*, pp. 4534-4537, 2010.
- [28] M. Ghovanloo, “Tongue operated assistive technologies,” in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, 2007, pp. 4376–4379, doi: 10.1109/IEMBS.2007.4353307.
- [29] K. Chand et al., “An Optimized Tongue Drive System for Disabled Persons,” *2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Glasgow, United Kingdom, 2021, pp. 1-6, doi: 10.1109/I2MTC50364.2021.9460029.
- [30] J. Kim, H. Park, J. Bruce, D. Rowles, J. Holbrook, B. Nardone, D. West, A. Laumann, E. Roth, and M. Ghovanloo, “Assessment of the tongue-drive system using a computer, a smartphone, and a powered-wheelchair by people with tetraplegia,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2015, doi: 10.1109/TNSRE.2015.2405072.
- [31] X. Huo, C. Cheng, and M. Ghovanloo, “Evaluation of the tongue drive system by individuals with high-level spinal cord injury,” in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, 2009, pp. 555–558, doi: 10.1109/IEMBS.2009.5334555.
- [32] A. Bodini, M. Borghetti, C. Paganelli, E. Sardini and M. Serpelloni, “Low-Power Wireless System to Monitor Tongue Strength Against the Palate,” in

- IEEE Sensors Journal*, vol. 21, no. 4, pp. 5467-5475, 15 Feb.15, 2021, doi: 10.1109/JSEN.2020.3036137.
- [33] H. Park and M. Ghovanloo, "An Arch-Shaped Intraoral Tongue Drive System with Built-in Tongue-Computer Interfacing SoC," *Sensors*, vol. 14, no. 11, pp. 21565–21587, Nov. 2014, doi: 10.3390/s141121565.
 - [34] J. Bajwa, U. Munir, A. Nori, and B. Williams, "Artificial intelligence in healthcare: Transforming the practice of medicine," *Future Healthc. J.*, vol. 8, no. 2, pp. e188–e194, Jul. 2021, doi: 10.7861/fhj.2021-0095.
 - [35] M. Chustecki, "Benefits and risks of AI in health care: Narrative review," *Interact. J. Med. Res.*, vol. 13, p. e53616, Nov. 2024, doi: 10.2196/53616.
 - [36] D. Goswami, A. C. Sajjan, R. Yadav, and N. Hathila, "Artificial intelligence in radiology and pathology: Transforming medical imaging interpretation," in *Medical Science: Recent Advances and Applications*, vol. 4, pp. 79–87, 2025, doi: 10.9734/bpi/msraa/v4/5423.
 - [37] G. J., R. S., H. L. G., V. Ravi, M. Almeshari, and Y. Alzamil, "Electronic health record (EHR) system development for study on EHR data-based early prediction of diabetes using machine learning algorithms," *Open Bioinform. J.*, vol. 16, p. e187503622309010, 2023, doi: 10.2174/18750362-v16-e230906-2023-15.
 - [38] J. N. K. Wah, "The rise of robotics and AI-assisted surgery in modern healthcare," *J. Robot. Surg.*, vol. 19, no. 1, p. 311, Jun. 2025, doi: 10.1007/s11701-025-02485-0.
 - [39] O. Boursalie, R. Samavi, and T. E. Doyle, "Machine learning and mobile health monitoring platforms: A case study on research and implementation challenges," *J. Healthc. Inform. Res.*, vol. 2, no. 1–2, pp. 179–203, May 2018, doi: 10.1007/s41666-018-0021-1.
 - [40] B. S. Reddy, "Advancements in Machine Learning: A comprehensive exploration of methods, applications, and future perspectives," *International Journal of Scientific Research in Engineering & Technology*, pp. 01–04, Feb. 2024, doi: 10.59256/ijsreat.20240401001.

- [41] M. H. Assaf, R. Kumar, K. Sharma, and B. Sharma, "Optimized tongue driven system using artificial intelligence," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, vol. 11, no. 3, pp. 696–710, 2023, doi: 10.1080/21681163.2022.2103451.
- [42] Asci, Francesco, Giovanni Costantini, Pietro Di Leo, Alessandro Zampogna, Giovanni Ruoppolo, Alfredo Berardelli, Giovanni Saggio, and Antonio Suppa. 2020. "Machine-Learning Analysis of Voice Samples Recorded through Smartphones: The Combined Effect of Ageing and Gender", *Sensors*, 20, no. 18: 5022. <https://doi.org/10.3390/s20185022>
- [43] Md. Hossan, S. Memon, and M. Gregory, "A novel approach for MFCC feature extraction," in *Proc. Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, 2011, pp. 1–5, doi: 10.1109/ICSPCS.2010.5709752.
- [44] T. R. Mahesh, S. Singh, V. Srivastav, H. Kalra, T. Rajan, and G. Yadav, "Adaptive modeling for real-time data analysis with machine learning," in *Proc. IEEE SMARTGENCON*, 2023, pp. 1–6, doi: 10.1109/SMARTGENCON60755.2023.10442430.
- [45] J. Kim, H. Park, J. Bruce, D. Rowles, J. Holbrook, B. Nardone, D. P. West, A. E. Laumann, E. Roth, E. Veledar, and M. Ghovanloo, "Qualitative assessment of tongue drive system by people with high-level spinal cord injury," *J. Rehabil. Res. Dev.*, vol. 51, no. 3, pp. 451–465, 2014, doi: 10.1682/JRRD.2013.08.0178.
- [46] P. Yu, "The future prospects of deep learning and neural networks: Artificial intelligence's impact on education," *Appl. Comput. Eng.*, vol. 33, pp. 94–101, 2024, doi: 10.54254/2755-2721/33/20230239.
- [47] H. Lin, "Embedded Artificial Intelligence: Intelligence on Devices," in *Computer*, vol. 56, no. 09, pp. 90-93, Sept. 2023, doi: 10.1109/MC.2023.3280397.
- [48] D. Amuru *et al.*, "AI/ML algorithms and applications in VLSI design and technology," *Integration*, vol. 93, p. 102048, Jun. 2023, doi: 10.1016/j.vlsi.2023.06.002.
- [49] Y. Attaoui, M. Chentouf, Z. Ismaili, and A. Elmourabit, "Machine learning in VLSI design: A comprehensive review," *J. Integr. Circuits Syst.*, vol. 19, pp. 1–14, 2024, doi: 10.29292/jics.v19i2.826.

- [50] A. Biscontini, E. Popovici and A. Temko, "Machine Learning for FPGA Electronic Design Automation," in *IEEE Access*, vol. 12, pp. 182640-182662, 2024, doi: 10.1109/ACCESS.2024.3511345.
- [51] M. B. Altman, W. Wan, A. S. Hosseini, S. Arabi Nowdeh, and M. Alizadeh, "Machine learning algorithms for FPGA implementation in biomedical engineering applications: A review," *Heliyon*, vol. 10, no. 4, p. e26652, Feb. 2024, doi: 10.1016/j.heliyon.2024.e26652.
- [52] H. Lin, "Embedded Artificial Intelligence: Intelligence on Devices," *Computer*, vol. 56, no. 09, pp. 90-93, Sept. 2023, doi: 10.1109/MC.2023.3280397.
- [53] R. Sinha, S. K. Gupta, P. Rajalingam and S. Routray, "Detecting Handwritten digits using fully connected Artificial Neural Network in FPGA which will act like a replacement for GPU," *2023 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, Chennai, India, 2023, pp. 1-5, doi: 10.1109/RAEEUCCI57140.2023.10134313
- [54] T. Zhao, "FPGA-based machine learning: Platforms, applications, design considerations, challenges, and future directions," *Highlights Sci. Eng. Technol.*, vol. 62, pp. 96–101, 2023, doi: 10.54097/hset.v62i.10430.
- [55] P. Goswami and D. Bhatia, "Application of Machine Learning in FPGA EDA Tool Development," in *IEEE Access*, vol. 11, pp. 109564-109580, 2023, doi: 10.1109/ACCESS.2023.3322358.
- [56] Types of AT / Guide to Assistive Technology. (n.d.). Guide to Assistive Technology. <https://mn.gov/admin/at/getting-started/understanding-at/types/>
- [57] World Health Organization: WHO, "Assistive technology," Jan. 02, 2024. <https://www.who.int/news-room/fact-sheets/detail/assistive-technology>
- [58] P. V. Ramudu and P. P. M. Krishna, "Tongue operated robotic wheel chair using tongue drive assistive technology," *International Journal of Engineering Research And*, vol. 2, no. 9, Sep. 2013, [Online]. Available: <https://www.ijert.org/research/tongue-operated-robotic-wheel-chair-using-tongue-drive-assistive-technology-IJERTV2IS90483.pdf>

- [59] M. Baig, S. Masud and M. Awais, "Support Vector Machine based Voice Activity Detection," *2006 International Symposium on Intelligent Signal Processing and Communications*, 2006, pp. 319-322, doi: 10.1109/ISPACS.2006.364896.
- [60] M. M. Abdulghani, K. M. Al-Aubidy, M. M. Ali, and Q. J. Hamarsheh, "Wheelchair neuro fuzzy control and tracking system based on voice recognition," *Sensors*, vol. 20, no. 10, p. 2872, May 2020, doi: 10.3390/s20102872.
- [61] A. Škraba, R. Stojanović, A. Zupan, A. Koložvari, and D. Kofjač, "Speech-controlled cloud-based wheelchair platform for disabled persons," *Microprocessors and Microsystems*, vol. 39, no. 8, pp. 819–828, Oct. 2015, doi: 10.1016/j.micpro.2015.10.004.
- [62] J. Hua, J. Rao, Y. Peng, J. Liu, and J. Tang, "Deep Compressive Sensing on ECG Signals with Modified Inception Block and LSTM," *Entropy*, vol. 24, no. 8, p. 1024, Jul. 2022, doi: 10.3390/e24081024.
- [63] N. Mohan, A. Hosni, and M. Atef, "Neural networks implementations on FPGA for biomedical applications: A review," *SN Computer Science*, vol. 5, 2024, doi: 10.1007/s42979-024-03381-4.
- [64] A. M. A. Hassan and S. Mohsen, "Compression of electrocardiogram signals using compressive sensing technique based on curvelet transform toward medical applications," *Multimedia Tools and Applications*, May 2024, doi: 10.1007/s11042-024-19328-z.
- [65] S. Mohsen, S. G. Scholz, and A. Elkaseer, "Detection of COVID-19 in Chest X-Ray Images Using a CNN Model toward Medical Applications," *Wireless Personal Communications*, vol. 137, no. 1, pp. 69–87, Jul. 2024, doi: 10.1007/s11277-024-11309-7.
- [66] N. S. Mohsen, N. W. M. F. Abdel-Rehim, N. A. Emam, and N. H. M. Kasem, "A convolutional neural network for automatic brain tumor detection," *Proceedings of Engineering and Technology Innovation*, vol. 24, pp. 15–21, Apr. 2023, doi: 10.46604/peti.2023.10307.
- [67] M. M. Zayed, S. Mohsen, A. Alghuried, H. Hijry, and M. Shokair, "IOUT-Oriented an efficient CNN model for modulation schemes recognition in optical

- wireless communication systems,” *IEEE Access*, p. 1, Jan. 2024, doi: 10.1109/access.2024.3515895.
- [68] D. L. Prasanna and S. L. Tripathi, “Machine and deep-learning techniques for text and speech processing,” in *Machine Learning Algorithms for Signal and Image Processing*, D. Ghai, S. L. Tripathi, S. Saxena, M. Chanda, and M. Alazab, Eds., 2022, doi: 10.1002/9781119861850.ch7.
- [69] S. Preetham and P. A. Vijaya, “Microcontroller Based Tongue Drive System for Physically Challenged People “, *International Journal of Artificial Intelligence and Mechatronics*, Volume 2, Issue 2, ISSN 2320 – 5121.
- [70] B. Yousefi, N. X. Huo, and M. Ghovanloo, “Preliminary assessment of Tongue Drive System in medium term usage for computer access and wheelchair control,” *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5766–5769, Aug. 2011, doi: 10.1109/iembs.2011.6091427.
- [71] X. Huo, H. Park, and M. Ghovanloo, “Dual-mode tongue Drive system: using speech and tongue motion to improve computer access for people with disabilities,” *WH '12: Proceedings of the Conference on Wireless Health*, pp. 1–8, Oct. 2012, doi: 10.1145/2448096.2448102.
- [72] P. S. and P. A. Vijaya, “Microcontroller based tongue drive system for physically challenged people,” *International Journal of Artificial Intelligence and Mechatronics*, vol. 2, no. 2, [Online].
https://www.ijaim.org/administrator/components/com_jresearch/files/publications/IJAIM_149_Final.pdf
- [73] J. Kim *et al.*, “Assessment of the Tongue-Drive system using a computer, a smartphone, and a Powered-Wheelchair by people with tetraplegia,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 68–78, Feb. 2015, doi: 10.1109/tnsre.2015.2405072.
- [74] K. Taniguchi, H. Kondo, M. Kurosawa, and A. Nishikawa, “Eearable TEMPO: A Novel, Hands-Free Input Device that Uses the Movement of the Tongue Measured with a Wearable Ear Sensor,” *Sensors*, vol. 18, no. 3, p. 733, Mar. 2018, doi: 10.3390/s18030733.

- [75] A. Botter, T. Vieira, C. Busso, F. Vitali, M. Gazzoni and G. L. Cerone, "Design and Test of an Intraoral Electrode Grid for Tongue High-Density Electromyography," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 2805-2814, 2024, doi: 10.1109/TNSRE.2024.3434360.
- [76] T. Hashimoto *et al.*, "TongueInput: Input Method by Tongue Gestures Using Optical Sensors Embedded in Mouthpiece," *2018 57th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, Nara, Japan, 2018, pp. 1219-1224, doi: 10.23919/SICE.2018.8492690.
- [77] M. Sahadat, A. Alreja, P. Srikrishnan, and M. Ghovanloo, "A multimodal human computer interface combining head movement, speech and tongue motion for people with severe disabilities," in *Proc. 2015 IEEE Biomed. Circuits Syst. Conf.*, 2015, pp. 1–4.
- [78] C. Gomez, J. Oller, and J. Paradells, "Overview and evaluation of bluetooth low energy: An emerging low-power wireless technology," *Sensors*, vol. 12, no. 9, pp. 11 734–11 753, 2012.
- [79] E. B. Sadeghian, X. Huo, and M. Ghovanloo, "Command detection and classification in tongue drive assistive technology," in *Proc. 2011 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2011, pp. 5465–5468.
- [80] X. Huo, J. Wang, and M. Ghovanloo, "Using magneto-inductive sensors to detect tongue position in a wireless assistive technology for people with severe disabilities," in *Proc. IEEE Sensors*, 2007, pp. 732–735.
- [81] S. L. Tripathi, L. P. Dasari, I. Wijayanto, D. Ghai, and M. Mahmud, "Assistive technology for neuro-rehabilitation applications using machine learning techniques," in *Machine Learning Models and Architectures for Biomedical Signal Processing*, Eds. Academic Press, 2025, pp. 179–191, doi: 10.1016/B978-0-443-22158-3.00008-9.
- [82] M. Inoue and Y. Koshikiya, "Enhanced frequency and time domain feature extraction for communication infrastructure type classification using optical fiber sensing," *Opt. Fiber Technol.*, vol. 87, p. 103859, 2024, doi: 10.1016/j.yofte.2024.103859.

- [83] N. Chauhan, T. Isshiki and D. Li, "Speaker Recognition Using LPC, MFCC, ZCR Features with ANN and SVM Classifier for Large Input Database," *2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS)*, Singapore, 2019, pp. 130-133, doi: 10.1109/CCOMS.2019.8821751.
- [84] C. Dromey, S. Nissen, P. Nohr, and S. G. Fletcher, "Measuring tongue movements during speech: Adaptation of a magnetic jaw-tracking system," *Speech Communication*, vol. 48, no. 5, pp. 463–473, 2005, doi: 10.1016/j.specom.2005.
- [85] R. I. Roni, M. M. Rahman and N. Mamun, "Multi-level Feature Extraction for Bangla Speech Emotion Recognition," *2025 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Chittagong, Bangladesh, 2025, pp. 1-6, doi: 10.1109/ECCE64574.2025.11013544.
- [86] A. Jafari, S. Gandhi, S. H. Konuru, W. David Hairston, T. Oates, and T. Mohsenin, "An EEG artifact identification embedded system using ICA and multi-instance learning," *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2017, pp. 1–4.
- [87] P. W. Schönle, K. Gräbe, P. Wenig, J. Höhne, J. Schrader, and B. Conrad, "Electromagnetic articulography: Use of alternating magnetic fields for tracking movements of multiple points inside and outside the vocal tract," *Brain and Language*, vol. 31, no. 1, pp. 26–35, 1987, doi: 10.1016/0093-934x(87)90058-7.
- [88] C. N. Huang et al., "Application of facial electromyography in computer mouse access for people with disabilities," *Disability Rehabilitation*, vol. 28, no. 4, pp. 231–237, 2006.
- [89] G. Krishnamurthy and M. Ghovanloo, "Tongue drive: A tongue operated magnetic sensor based wireless assistive technology for people with severe disabilities," in *Proc. International Symp. Circuits and Systems*, 2006, pp. 5551–5554.
- [90] Ayala-Acevedo, A., & Ghovanloo, M. (2014), "Smartphone-compatible robust classification algorithm for the Tongue Drive System," *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*, doi: 10.1109/biocas.2014.6981670

- [91] M. Diab, “Tongue drive wheelchair,” *International Journal of Computational and Cognition*, vol. 10, pp. 29–34, 2012.
- [92] W. F. Katz, S. Mehta, M. Wood, and J. Wang, “Using electromagnetic articulography with a tongue lateral sensor to discriminate manner of articulation,” *The Journal of the Acoustical Society of America*, vol. 141, no. 1, pp. EL57–EL63, 2017, doi: 10.1121/1.4973907.
- [93] N. Sebkhi, A. Bhavsar, D. V. Anderson, J. Wang, & O. T. Inan, “Inertial measurements for tongue motion tracking based on magnetic localization with orientation compensation,” *IEEE Sensors Journal*, 2022, 21(6), 7964–7971. <https://doi.org/10.1109/jsen.2020.3046469>
- [94] T. Rebernik, J. Jacobi, R. Jonkers, A. Noiray & M. Wieling, “A review of data collection practices using electromagnetic articulography”, *Laboratory Phonology*, 2021, 12(1): 6. doi: <https://doi.org/10.5334/labphon.237>
- [95] R. Barea, L. Boquete, M. Mazo, and E. Lopez, “System for assisted mobility using eye movements based on electrooculography,” *IEEE Transaction on Neural System Rehabilitation Engineering*, vol. 10, no. 4, pp. 209–218, Dec. 2002.
- [96] M. Afify and O. Siohan, “Comments on vocal tract length normalization equals linear transformation in cepstral space,” *IEEE Trans. Audio, Speech, Language Process.*, vol. 15, no. 5, pp. 1731–1732, Jul. 2007, doi: 10.1109/TASL.2007.896653.
- [97] G. S. V. S. Sivaram, S. K. Nemala, N. Mesgarani, and H. Hermansky, “Data-driven and feedback based spectro-temporal features for speech recognition,” *IEEE Signal Process. Lett.*, vol. 17, no. 11, pp. 957–960, Nov. 2010, doi: 10.1109/LSP.2010.2079930.
- [98] Z. Huang, J. Epps, D. Joachim, and V. Sethu, “Natural language processing methods for acoustic and landmark event-based features in speech-based depression detection,” *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 2, pp. 435–448, Feb. 2020, doi: 10.1109/JSTSP.2019.2949419.
- [99] M. Bhattacharjee, S. R. M. Prasanna, and P. Guha, “Speech/music classification using features from spectral peaks,” *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 28, pp. 1549–1559, 2020,

doi: 10.1109/TASLP.2020.2993152.

- [100] D. L. Moore, Z. Hamilton, and R. J. Haddad, "Pico-grid smart home energy management system using mel frequency cepstral coefficients," in *Proc. SoutheastCon*, Huntsville, AL, USA, 2019, pp. 1–6, doi: 10.1109/SoutheastCon42311.2019.9020412.
- [101] D. S. Widyowaty and A. Sunyoto, "Accent recognition by native language using mel-frequency cepstral coefficient and K-Nearest neighbor," in *Proc. 3rd Int. Conf. Inf. Commun. Technol. (ICOIACT)*, Nov. 2020, pp. 314–318, doi: 10.1109/ICOIACT50329.2020.9332026.
- [102] M. Sahidullah and G. Saha, "A novel windowing technique for efficient computation of MFCC for speaker recognition," *IEEE Signal Process. Lett.*, vol. 20, no. 2, pp. 149–152, Feb. 2013, doi: 10.1109/LSP.2012.2235067.
- [103] K.S.Abbirame, A.S.Haripriya, V.Aravind and G.Jamuna, "An electronic gadget to control wheelchair motion using tongue gesture," *SSRG International Journal of Pharmacy and Biomedical Engineering*, vol. 4, no. 1, pp. 1-3, 2017.
Crossref, <https://doi.org/10.14445/23942576/IJPBE-V4I1P101>
- [104] Sasaki, M., Onishi, K., Stefanov, D. *et al.* Tongue interface based on surface EMG signals of suprahyoid muscles. *Robomech J.*, 3, 9 (2016),
doi: <https://doi.org/10.1186/s40648-016-0048-0>
- [105] A. Botter, T. Vieira, C. Busso, F. Vitali, M. Gazzoni, and G. L. Cerone, "Design and test of an intraoral electrode grid for tongue High-Density electromyography," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 32, pp. 2805–2814, Jan. 2024, doi: 10.1109/tnsre.2024.3434360.
- [106] A. Bashatah, B. Mukherjee, A. Rima, S. Patwardhan, P. Otto, R. Sutherland, E. L. King, B. Lancaster, A. Aher, G. Gibson, L. De Marzi, Z. Taghizadeh, S. Acuna, P. V. Chitnis, and S. Sikdar, "Wearable ultrasound system using low-voltage time delay spectrometry for dynamic tissue imaging," *IEEE Transactions on Biomedical Engineering*, vol. 71, no. 11, pp. 3232–3243, Nov. 2024, doi: 10.1109/TBME.2024.3414419.

- [107] S. Dugan, S. R. Li, J. Masterson, H. Woeste, N. Mahalingam, C. Spencer, T. D. Mast, M. A. Riley, and S. E. Boyce, “Tongue part movement trajectories for /r/ using ultrasound,” *Perspectives of the ASHA Special Interest Groups*, vol. 4, no. 6, pp. 1644–1652, Dec. 2019, doi: 10.1044/2019_pers-19-00064.
- [108] M. S. Ribeiro, J. Cleland, A. Eshky, K. Richmond, and S. Renals, “Exploiting ultrasound tongue imaging for the automatic detection of speech articulation errors,” *Speech Commun.*, vol. 128, pp. 24–34, 2021.
- [109] I. Fasel and J. Berry, “Deep Belief Networks for Real-Time Extraction of Tongue Contours from Ultrasound During Speech,” *2010 20th International Conference on Pattern Recognition*, Istanbul, Turkey, 2010, pp. 1493–1496, doi: 10.1109/ICPR.2010.369
- [110] K. Al-Hammuri, F. Gebali, I. Thirumarai Chelvan, and A. Kanan, “Tongue contour tracking and segmentation in lingual ultrasound for speech recognition: A review,” *Diagnostics*, vol. 12, p. 2811, 2022, doi: 10.3390/diagnostics12112811.
- [111] S. Zhang, “A framework for tongue model visualization based on ultrasound images,” *Signal, Image and Video Processing*, vol. 19, 2025, doi: 10.1007/s11760-025-04270-w.
- [112] P. Zhang and H. Li, “Hybrid model of continuous hidden Markov model and multi-layer perceptron in speech recognition,” in *Proc. 2nd Int. Conf. Intelligent Computation Technology and Automation*, 2009, pp. 62–65.
- [113] N. Najkar, F. Razzazi, and H. Sameti, “A novel approach to HMM-based speech recognition systems using particle swarm optimization,” *Mathematical and Computer Modelling*, vol. 52, no. 11–12, pp. 1910–1920, 2010, doi: 10.1016/j.mcm.2010.03.041.
- [114] A. Rudnický, L. Baumeister, K. DeGraaf, and E. Lehmann, “The lexical access component of the CMU continuous speech recognition system,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, 1987, pp. 376–379, doi: 10.1109/ICASSP.1987.1169705.
- [115] S. P. Bayerl, A. Tammewar, K. Riedhammer, and G. Riccardi, “Detecting emotion carriers by combining acoustic and lexical representations,” in *Proc.*

- IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2021, pp. 31–38, doi: 10.1109/ASRU51503.2021.9687893.
- [116] R. Li, J. Yu, C. Jiang, C. Luo and Z. Wang, “A mass-spring tongue model with efficient collision detection and response during speech”, *The 9th International Symposium on Chinese Spoken Language Processing*, pp. 354-358, 2014.
 - [117] D. L. Prasanna and S. L. Tripathi, “A review on tongue based assistive technology, devices and FPGA processors using machine learning module,” *Wireless Pers. Commun.*, vol. 134, pp. 151–170, 2024, doi: 10.1007/s11277-024-10897-8.
 - [118] H. F. Chen, C. Y. Yeh, and K. H. Lin, “A method of using two equivalent negative inductances to reduce parasitic inductances of a three-capacitor EMI filter,” *IEEE Trans. Power Electron.*, vol. 24, no. 12, pp. 2867–2872, Dec. 2009.
 - [119] L. J. Chen and K. H. Lin, “Implementation of a compact EMI filter array for 4G-LTE applications on LTCC,” *IEEE Trans. Compon., Packag., Manuf. Technol.*, vol. 5, no. 6, pp. 713–722, Jun. 2015.
 - [120] A. B. Menicanin et al., “Improved model of T-type LC EMI chip filters using new microstrip test fixture,” *IEEE Trans. Magn.*, vol. 47, no. 10, pp. 3975–3978, Oct. 2011.
 - [121] H. Chen et al., “Series-connected grounding of common-mode EMI filter,” *IEEE Trans. Electromagn. Compat.*, vol. 52, no. 4, pp. 1066–1068, Nov. 2010.
 - [122] H. F. Huang et al., “Techniques for improving the high-frequency performance of the planar CM EMI filter,” *IEEE Trans. Electromagn. Compat.*, vol. 55, no. 5, pp. 901–908, Oct. 2013.
 - [123] G. Ala, G. C. Giaconia, G. Giglia, M. C. D. Piazza, and G. Vitale, “Design and Performance Evaluation of a High Power-Density EMI Filter for PWM Inverter-Fed Induction-Motor Drives,” *IEEE Trans. Ind. Appl.*, vol. 52, no. 3, pp. 2397-2404, May-June 2016.
 - [124] D. Hamza, M. Qiu, and P. K. Jain, “Application and Stability Analysis of a Novel Digital Active EMI Filter Used in a Grid-Tied PV Microinverter Module,” *IEEE Trans. Power Electron.*, vol. 28, no. 6, pp. 2867-2874, Jun 2013.

- [125] W. Chen, X. Yang, and Z. Wang, "An active EMI filtering technique for improving passive filter low-frequency performance," *IEEE Trans. Electromagn. Compat.*, vol. 48, no. 1, pp. 172-177, Feb 2006.
- [126] D. Shin, S. Kim, G. Jeong, J. Park, J. Park, K. J. Han, et al., "Analysis and Design Guide of Active EMI Filter in a Compact Package for Reduction of Common-Mode Conducted Emissions," *IEEE Trans. Electromagn. Compat.*, vol. 57, no. 4, pp. 660-671, Aug 2015.
- [127] X. Wu, D. Xu, Z. Wen, Y. Okuma, and K. Mino, "Design, Modeling, and Improvement of Integrated EMI Filter With Flexible Multilayer Foils," *IEEE Trans. Power Electron.*, vol. 26, no. 5, pp. 1344-1354, May 2011.
- [128] R. Lai, Y. Maillet, F. Wang, S. Wang, R. Burgos, and D. Boroyevich, "An Integrated EMI Choke for Differential-Mode and Common-Mode Noise Suppression," *IEEE Trans. Power Electron.*, vol. 25, no. 3, pp. 539-544, Mar 2010.
- [129] M. Satish Kumar and A. Jhansirani, "Simulation of EMI Filters Using Matlab," *International Journal of Engineering Science Invention*, vol. 6, no. 5, May 2017, pp. 73-76.
- [130] V. Tarateeraseth, K. Y. See, F. G. Canavero, and R. W.-Y. Chang, "Systematic electromagnetic interference filter design based on information from in-circuit impedance measurements," *IEEE Transactions on Electromagnetic Compatibility*, vol. 52, no. 3, pp. 588-597, Aug. 2010.
- [131] B. Kaka and A. Maji, "Performance evaluation of shunt active power filter (SAPF) connected to three phase four wire distribution networks," *2016 IEEE International Telecommunications Energy Conference (INTELEC)*, Austin, TX, USA, 2016, pp. 1-9, doi: 10.1109/INTLEC.2016.7749109.
- [132] A. Richelli, S. Kennedy and J. Redouté, "An EMI-Resistant Common-Mode Cancellation Differential Input Stage in UMC 180 nm CMOS," in *IEEE Transactions on Electromagnetic Compatibility*, vol. 59, no. 6, pp. 2049-2051, Dec. 2017, doi: 10.1109/TEMC.2017.2696166.
- [133] Y. Bai, W. Chen, R. He, D. Zhang and X. Yang, "EMI noise cancelation by optimizing transformer design without need for the traditional Y-

- capacitor,” *2016 IEEE Applied Power Electronics Conference and Exposition (APEC)*, 2016, pp. 766-771, doi: 10.1109/APEC.2016.7467958.
- [134] Saumya Sah, “Voice Controlled Wheelchair Using Feature Extraction Techniques Through MATLAB”, *International Journal for Research in Engineering Application & Management (IJREAM)*, Vol-04, Issue-10, Jan 2019. <https://ijream.org/papers/IJREAMV04I1046048.pdf>
- [135] R. Deshmukh and K. Ghule, “Feature extraction techniques for speech recognition: A review,” *International Journal of Scientific and Engineering Research*, vol. 6, pp. 143–147, 2015.
- [136] R. Hibare and A. Vibhute, “Feature extraction techniques in speech processing: A survey,” *International Journal of Computer Applications*, vol. 107, no. 5, pp. 1–8, Dec. 2014, doi: 10.5120/18744-9997.
- [137] Shing-Tai Pan, Chih-Chin Lai and Bo-Yu Tsai, “The Implementation of Speech Recognition Systems on FPGA-Based Embedded Systems with SOC Architecture”, *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 11, pp. 6161-6175, November 2011.
- [138] H. Tyagi, R. M. Hegde, H. A. Murthy and A. Prabhakar, "Automatic identification of bird calls using Spectral Ensemble Average Voice Prints," *2006 14th European Signal Processing Conference*, Florence, Italy, 2006, pp. 1-5.
- [139] Dwijen Rudrapal, Smita Das, S. Debbarma, N. Kar, N. Debbarma, “Voice Recognition and Authentication as a Proficient Biometric Tool and its Application in Online Exam for P. H People”, *International Journal of Computer Applications* (0975 – 8887), vol. 39, no. 12, pp. 7-12, February 2012.
- [140] V. Gupta and N. Jain, “Frequency analysis of speech signals for Devanagari script and numerals,” *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, vol. 11, pp. 124–129, 2016, doi: 10.9790/2834-110501124129.
- [141] Asm Sayem, “Speech Analysis for Alphabets in Bangla Language:Automatic Speech Recognition,” *International Journal of Engineering Research*, vol. 3, no. 2, pp. 88-93, February 2014.
- [142] D. K., S. X., and B. X., “Optimization of tone recognition via applying linear discriminant analysis in feature extraction,” in *Proc. 3rd Int. Conf. Innovative*

Computing Information and Control (ICICIC), Dalian, China, 2008, pp. 528–531.

- [143] A. Abdusalomov, F. Safarov, M. Rakhimov, B. Turaev, and T. K. Whangbo, “Improved feature parameter extraction from speech signals using machine learning algorithm,” *Sensors*, vol. 22, no. 21, p. 8122, 2022, doi: 10.3390/s22218122.
- [144] P. K. Kurzekar, R. R. Deshmukh, V. B. Waghmare, and P. P. Shrishrimal, “A comparative study of feature extraction techniques for speech recognition system,” *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 12, 2014, doi: 10.15680/IJIRSET.2014.0312034.
- [145] R. Loughran, A. Agapitos, A. Kattan, and A. Brabazon, “Feature selection for speaker verification using genetic programming,” *Evolutionary Intelligence*, vol. 10, no. 1–2, pp. 1–21, 2017, doi: 10.1007/s12065-016-0150-5.
- [146] K. H. Kim, H. K. Kim, J. S. Kim, W. Son, and S. Y. Lee, “A biosignal-based human interface controlling a power-wheelchair for people with motor disabilities,” *ETRI Journal*, vol. 28, no. 1, pp. 111–114, 2006.
- [147] A. A. Bashit and D. Valles, “A Mel-Filterbank and MFCC-based Neural Network Approach to Train the Houston Toad Call Detection System Design,” *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, Vancouver, BC, Canada, 2018, pp. 438–443, doi: 10.1109/IEMCON.2018.8615076.
- [148] J. B. and R. B. N., “Speech recognition using MFCC and DTW,” in *Proc. IEEE Int. Conf. Advances in Electrical Engineering (ICAEE)*, 2014, doi: 10.1109/ICAEE.2014.6838564.
- [149] M. J. M. H., S. A. Al-Haddad, and C. N. Kyun, “A flexible speech recognition system for cerebral palsy disabled,” in *Proc. Int. Conf. Informatics Engineering and Information Science*, 2011, pp. 42–55.
- [150] H. Coskun and T. Yigit, “Artificial intelligence applications on classification of heart sounds,” in *Nature-Inspired Intelligent Techniques for Solving Biomedical Engineering Problems*. Hershey, PA, USA: IGI Global, 2018, pp. 146–183.
- [151] S. B. Dhonde, A. Chaudhari, and S. M. Jagade, “Integration of mel-frequency cepstral coefficients with log energy and temporal derivatives for text-

- independent speaker identification,” in *Proc. Int. Conf. Data Engineering and Communication Technology*, 2017, pp. 791–797.
- [152] I. Idris and M. S. Salam, “Improved speech emotion classification from spectral coefficient optimization,” in *Advances in Machine Learning and Signal Processing*, 2016, pp. 247–257, doi: 10.1007/978-3-319-32213-1_22.
 - [153] N. Dave, “Feature extraction methods LPC, PLP and MFCC in speech recognition,” *International Journal for Advance Research in Engineering and Technology*, vol. 1, no. 6, 2013.
 - [154] Z. Tüfekci and J. N. Gowdy, “Feature extraction using discrete wavelet transform for speech recognition,” in *Proc. IEEE SOUTHEASTCON*, 2000, pp. 116–123, doi: 10.1109/SECON.2000.845444.
 - [155] S. Ayat, M. T. Manzuri-Shalmani, and R. Dianat, “An improved wavelet-based speech enhancement by using speech signal features,” *Comput. Electr. Eng.*, vol. 32, no. 6, pp. 411–425, 2006, doi: 10.1016/j.compeleceng.2006.05.002.
 - [156] A. Alshammari, M. S. A. Alzaidi, O. Alrusaini, A. Alshuhail, R. Allafi, A. Yafoz, R. Alsini, and A. O. Khadidos, “Robust speech perception and classification-driven deep convolutional neural network with natural language processing,” *Alex. Eng. J.*, vol. 123, pp. 358–368, 2025, doi: 10.1016/j.aej.2025.03.046.
 - [157] M. Strake, B. Defraene, K. Fluyt, W. Tirry, and T. Fingscheidt, “Speech enhancement by LSTM-based noise suppression followed by CNN-based speech restoration,” *EURASIP Journal on Advances in Signal Processing*, vol. 2020, no. 1, pp. 1–26, 2020.
 - [158] M. R. Hasan, M. M. Hasan, and M. Z. Hossain, “How many mel-frequency cepstral coefficients to be utilized in speech recognition? A study with the Bengali language,” *Journal of Engineering*, vol. 2021, no. 12, pp. 817–827, 2021.
 - [159] J. Zhao, R. Li, M. Tian, W. An, “Multi-view self-supervised learning and multi-scale feature fusion for automatic speech recognition”, *Neural Process. Lett.*, 56 (4) (2024), p. 168
 - [160] M.N. Khan, A. Altalbe, F. Naseer, Q. Awais, “Telehealth-enabled in-home elbow rehabilitation for brachial plexus injuries using deep-reinforcement-learning-assisted telepresence robots”, *Sensors*, 24 (4) (2024), p. 1273

- [161] F. Naseer, M.N. Khan, A. Rasool, N. Ayub, “A novel approach to compensate delay in communication by predicting teleoperator behaviour using deep learning and reinforcement learning to control telepresence robot”, *Electron. Lett.*, 59 (9) (2023), Article e12806
- [162] C. G. Le Prell and O. H. Clavier, “Effects of noise on speech recognition: Challenges for communication by service members,” *Hearing Research*, vol. 349, pp. 76–89, 2017.
- [163] J. A. Gonzalez-Lopez, A. Gomez-Alanis, J. M. Martin Donas, J. L. Perez-Cordoba, and A. M. Gomez, “Silent speech interfaces for speech restoration: A review,” *IEEE Access*, vol. 8, pp. 177995–178021, 2020.
- [164] P. Lopez-Meyer, J. A. del Hoyo Ontiveros, H. Lu, and G. Stemmer, “Efficient end-to-end audio embeddings generation for audio classification on target applications,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, 2021, pp. 601–605.
- [165] D.-X. Wang, M.-S. Jiang, F.-L. Niu, Y.-D. Cao, and C.-X. Zhou, “Speech enhancement control design algorithm for dual-microphone systems using β -NMF in a complex environment,” *Complexity*, vol. 2018, pp. 1–13, 2018.
- [166] W. Nabi, N. Aloui, and A. Cherif, “Speech enhancement in dual-microphone mobile phones using Kalman filter,” *Applied Acoustics*, vol. 109, pp. 1–4, 2016.
- [167] M. Wand, M. Janke, and T. Schultz, “Tackling speaking mode varieties in EMG-based speech recognition,” *IEEE Transactions on Biomedical Engineering*, vol. 61, pp. 2515–2526, 2014.
- [168] G. K. Anumanchipalli, J. Chartier, and E. F. Chang, “Speech synthesis from neural decoding of spoken sentences,” *Nature*, vol. 568, pp. 493–498, 2019.
- [169] S. Duraivel, et al., “High-resolution neural recordings improve the accuracy of speech decoding,” *Nature Communications*, vol. 14, p. 6938, 2023.
- [170] H. C. Liu, et al., “An epidermal sEMG tattoo-like patch as a new human-machine interface for patients with loss of voice,” *Microsystems & Nanoengineering*, vol. 6, p. 16, 2020.
- [171] Y. Wang, et al., “All-weather, natural silent speech recognition via machine-learning-assisted tattoo-like electronics,” *NPJ Flexible Electronics*, vol. 5, p. 20, 2021.

- [172] T.-S. Le, J. An, Y. Huang, Q. Vo, J. Boonruangkan, T. Tran, S.-W. Kim, G. Sun, and Y.-J. Kim, “Ultrasensitive anti-interference voice recognition by bio-inspired skin-attachable self-cleaning acoustic sensors,” *ACS Nano*, 2019, doi: 10.1021/acsnano.9b06354.
- [173] Analog Devices, “ADMP421 datasheet,” [Online]. Available: <https://www.analog.com/media/en/technical-documentation/obsolete-data-sheets/ADMP421.pdf>
- [174] K. Nakamura, T. Toda, H. Saruwatari, and K. Shikano, “Speaking-aid systems using GMM-based voice conversion for electrolaryngeal speech,” *Speech Communication*, vol. 54, pp. 134–146, 2012.
- [175] S. Yuan, et al., “Non-acoustic speech sensing system based on flexible piezoelectric,” in *Proc. ACM Conf. Embedded Networked Sensor Systems*, 2022, pp. 1055–1060.
- [176] L. Han, et al., “Health monitoring via heart, breath, and Korotkoff sounds by wearable piezoelectret patches,” *Advanced Science*, vol. 10, p. 2301180, 2023.
- [177] J. Z. Zhao, et al., “Self-powered speech recognition system for deaf users,” *Cell Reports Physical Science*, vol. 3, p. 101168, 2022.
- [178] S. Lee, et al., “An electret-powered skin-attachable auditory sensor that functions in harsh acoustic environments,” *Advanced Materials*, vol. 34, p. 2205537, 2022.
- [179] Q. S. Yang, et al., “Mixed-modality speech recognition and interaction using a wearable artificial throat,” *Nature Machine Intelligence*, vol. 5, pp. 169–180, 2023.
- [180] C. Wang, et al., “Detection of non-joint areas tiny strain and anti-interference voice recognition by micro-cracked metal thin film,” *Nano Energy*, vol. 34, pp. 578–585, 2017.
- [181] R. Sridharan, “Epidemiology of epilepsy,” *Current Science*, vol. 82, no. 6, pp. 664–670, 2002.
- [182] E. Faught, “Epilepsy case studies,” *Neurologic Clinics*, vol. 24, no. 2, pp. 291–307, May 2006, doi: 10.1016/j.ncl.2006.01.001.

- [183] E. C. Galizia and H. J. Faulkner, "Seizures and epilepsy in the acute medical setting: Presentation and management," *Clinical Medicine*, vol. 18, no. 5, pp. 409–413, 2018.
- [184] World Health Organization, "Epilepsy," [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/epilepsy>
- [185] Mayo Clinic, "Epilepsy," [Online]. Available: <https://www.mayoclinic.org/diseases-conditions/epilepsy/symptoms-causes/syc-20350093>
- [186] Centers for Disease Control and Prevention, "SUDEP," [Online]. Available: <https://www.cdc.gov/epilepsy/about/sudep/index.htm>
- [187] M. Strein, J. P. Holton-Burke, L. R. Smith, and G. M. Brophy, "Prevention, treatment, and monitoring of seizures in the intensive care unit," *Journal of Clinical Medicine*, vol. 8, no. 8, p. 1177, 2019.
- [188] J. Desai, "Electroencephalography (EEG) data collection and processing through machine learning," M.S. thesis, Univ. Arkansas, 2014. [Online]. Available: <https://scholarworks.uark.edu/etd/2160>
- [189] S. J. M. Smith, "EEG in the diagnosis, classification, and management of patients with epilepsy," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, pp. ii2–ii7, 2005.
- [190] C. Bigan, "A recursive time-frequency processing method for neural networks recognition of EEG seizures," in *Neural Networks and Expert Systems in Medicine and Healthcare*, Singapore: World Scientific, 1998, pp. 67–73.
- [191] M. Sukriti and D. Mitra, "Epilepsy seizure detection using kurtosis-based VMD's parameters selection and bandwidth features," *Biomedical Signal Processing and Control*, vol. 64, p. 102255, 2021.
- [192] V. A. Kumari and R. Chitra, "Classification of diabetes disease using support vector machine," *International Journal of Engineering Research and Applications*, vol. 3, pp. 1797–1801, 2013.
- [193] E. Venkatesan, "Performance analysis of decision tree algorithms for breast cancer classification," *Indian Journal of Science and Technology*, vol. 8, no. 1, pp. 1–8, 2015.

- [194] A. Sarwar, “Intelligent naïve Bayes approach to diagnose diabetes type-2,” *International Journal of Computer Applications*, 2012.
- [195] M. R. Kousarrizi, F. Seiti, and M. Teshnehlab, “An experimental comparative study on thyroid disease diagnosis based on feature subset selection and classification,” 2012.
- [196] K. Vembandasamy, R. Sasipriya, and E. Deepa, “Heart diseases detection using naive Bayes algorithm,” *IJISSET*, vol. 2, pp. 441–444, 2015.
- [197] M. B. T. Noor, et al., “Application of deep learning in detecting neurological disorders from magnetic resonance images: A survey,” *Brain Informatics*, vol. 7, no. 1, 2020.
- [198] A. Nayyar, L. Gadhavi, and N. Zaman, “Machine learning in healthcare: Review, opportunities and challenges,” in *Machine Learning and the Internet of Medical Things in Healthcare*, Academic Press, 2021, pp. 23–45.
- [199] K. Shailaja, B. Seetharamulu, and M. A. Jabbar, “Machine learning in healthcare: A review,” in *Proc. Int. Conf. Electronics, Communication and Aerospace Technology (ICECA)*, 2018.
- [200] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, “Automated EEG analysis of epilepsy: A review,” *Knowledge-Based Systems*, vol. 45, pp. 147–165, 2013.
- [201] B. Abbaszadeh, C. A. D. Teixeira, and M. C. E. Yagoub, “Feature selection techniques for the analysis of discriminative features in temporal and frontal lobe epilepsy: A comparative study,” *The Open Biomedical Engineering Journal*, vol. 15, no. 1, pp. 1–15, 2021.
- [202] M. S. Siddiqui, R. Morales-Menendez, X. Huang, and N. Hussain, “A review of epileptic seizure detection using machine learning classifiers,” *Brain Informatics*, vol. 7, no. 1, 2020, doi: 10.1186/s40708-020-00105-1.
- [203] A. S. M. Murugavel and S. Ramakrishnan, “Hierarchical multi-class SVM with ELM kernel for epileptic EEG signal classification,” *Medical & Biological Engineering & Computing*, vol. 54, no. 1, pp. 149–161, 2015, doi: 10.1007/s11517-015-1351-2.
- [204] I. Wijayanto, A. Humairani, S. Hadiyoso, A. Rizal, D. L. Prasanna, and S. L. Tripathi, “Epileptic seizure detection on a compressed EEG signal using energy

- measurement,” *Biomedical Signal Processing and Control*, vol. 85, p. 104872, 2023, doi: 10.1016/j.bspc.2023.104872.
- [205] R. Sarić, D. Jokić, N. Beganovic, L. Gurbeta, and A. Badnjevic, “FPGA-based real-time epileptic seizure classification using artificial neural network,” *Biomedical Signal Processing and Control*, vol. 62, p. 102106, 2020, doi: 10.1016/j.bspc.2020.102106.
- [206] A. Saidi, S. Ben Othman, W. Kacem and S. Ben Saoud, "FPGA implementation of EEG signal analysis system for the detection of epileptic seizure," *2018 International Conference on Advanced Systems and Electric Technologies (ICASET)*, Hammamet, Tunisia, 2018, pp. 415-420, doi: 10.1109/ASET.2018.8379892.
- [207] M. Altaf and J. Yoo, “A 1.83 μ J/classification, 8-channel, patient-specific epileptic seizure classification SoC using a non-linear support vector machine,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 1, pp. 49–60, 2016, doi: 10.1109/TBCAS.2014.2386891.
- [208] L. Feng, Z. Li, and Y. Wang, “VLSI design of SVM-based seizure detection system with on-chip learning capability,” *IEEE Transactions on Biomedical Circuits and Systems*, vol. 12, no. 1, pp. 171–181, 2018, doi: 10.1109/TBCAS.2017.2762721.
- [209] V. Harpale and V. Bairagi, “FPGA-based architecture implementation for epileptic seizure detection using one-way ANOVA and genetic algorithm,” *Biomedical & Pharmacology Journal*, vol. 12, pp. 1543–1553, 2019, doi: 10.13005/bpj/1785.
- [210] P. Indira and R. J. Krishna, “Optimized adaptive neuro fuzzy inference system (OANFIS) based EEG signal analysis for seizure recognition on FPGA,” *Biomedical Signal Processing and Control*, vol. 66, p. 102484, 2021, doi: 10.1016/j.bspc.2021.102484.
- [211] M. Sahani, S. K. Rout, and P. K. Dash, “FPGA implementation of epileptic seizure detection using semisupervised reduced deep convolutional neural network,” *Applied Soft Computing*, vol. 110, p. 107639, 2021, doi: 10.1016/j.asoc.2021.107639.

- [212] Y. S. Akgul, C. Kambhamettu and M. Stone, "Automatic extraction and tracking of the tongue contours," in *IEEE Transactions on Medical Imaging*, vol. 18, no. 10, pp. 1035-1045, Oct. 1999, doi: 10.1109/42.811315.
- [213] Álvaro J. Jiménez Palenzuela, Flavius Frasincar, Maria Mihaela Truşcă, "Modeling Second Language Acquisition with pre-trained neural language models," *Expert Systems with Applications*, Volume 207, 2022, 117871, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2022.117871>.
- [214] I. Fasel and J. Berry, "Deep Belief Networks for Real-Time Extraction of Tongue Contours from Ultrasound During Speech," *2010 20th International Conference on Pattern Recognition*, Istanbul, Turkey, 2010, pp. 1493-1496, doi: 10.1109/ICPR.2010.369
- [215] K. Al-Hammuri, F. Gebali, I. Thirumarai Chelvan, and A. Kanan, "Tongue contour tracking and segmentation in lingual ultrasound for speech recognition: A review," *Diagnostics*, vol. 12, p. 2811, 2022, doi: 10.3390/diagnostics12112811.
- [216] A. Johnson-Long, X. Huo, M. Ghovanloo, and M. Shinohara, "Dual-task motor performance with a tongue-operated assistive technology compared with hand operations," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, p. 1, 2012, doi: 10.1186/1743-0003-9-1.
- [217] A. Verma, N. Singh, V. Khanna, et al., "Automated tongue contour extraction from ultrasound sequences using signal enhancing neural network and energy minimized spline," *Multimedia Tools and Applications*, vol. 83, pp. 57511–57530, 2024, doi: 10.1007/s11042-023-17813-5.
- [218] M. Feng, Y. Wang, K. Xu, H. Wang and B. Ding, "Improving Ultrasound Tongue Contour Extraction Using U-Net and Shape Consistency-Based Regularizer," *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, 2021, pp. 6443-6447, doi: 10.1109/ICASSP39728.2021.9414420.
- [219] "Ultrasuite TAL corpus," [Online]. Available: https://ultrasuite.github.io/data/tal_corpus/

- [220] Y. Yu and S. T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1260–1270, 2002, doi: 10.1109/TIP.2002.804276.
- [221] S. H. Kang and J. Y. Kim, "Application of fast non-local means algorithm for noise reduction using separable color channels in light microscopy images," *International Journal of Environmental Research and Public Health*, vol. 18, no. 6, p. 2903, Mar. 2021, doi: 10.3390/ijerph18062903.
- [222] C. Laporte and L. Ménard, "Multi-hypothesis tracking of the tongue surface in ultrasound video recordings of normal and impaired speech," *Medical Image Analysis*, vol. 44, pp. 98–114, 2018, doi: 10.1016/j.media.2017.12.003.
- [223] E. Karimi, L. Ménard, and C. Laporte, "Fully-automated tongue detection in ultrasound images," *Comput. Biol. Med.*, vol. 111, p. 103335, 2019, doi: 10.1016/j.compbimed.2019.103335.
- [224] D. L. Prasanna, S. L. Tripathi and I. Wijayanto, "Classifying the Tongue Contours for Assistive Devices," *2024 IEEE International Conference of Electron Devices Society Kolkata Chapter (EDKCON)*, Kolkata, India, 2024, pp. 1-5, doi: 10.1109/EDKCON62339.2024.10870771.
- [225] Reka Trencsenyi and Laszlo Czap, "Possible Methods for Combining Tongue Contours of Dynamic MRI and Ultrasound Records," *Acta Polytechnica Hungarica*, 2021, vol 18, no. 14.
- [226] Y. Akgul, C. Kambhamettu, and M. Stone, "Automatic extraction and tracking of the tongue contours," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 1035–1045, 1999, doi: 10.1109/42.811315.
- [227] B. Liu and J. Liu, "Overview of image noise reduction based on non-local mean algorithm," *MATEC Web of Conferences*, vol. 232, p. 03029, 2018, doi: 10.1051/mateconf/201823203.
- [228] J. Zhu, W. Styler, and I. Calloway, "A CNN-based tool for automatic tongue contour tracking in ultrasound images," *arXiv preprint arXiv:1907.10210*, 2019, doi: 10.48550/arXiv.1907.10210.
- [229] R. Kaur and P. S. Maan, "M-band ridgelet transform to remove speckle noise from medical images," *International Journal of Computer Applications*, vol. 85, no. 9, pp. 1–5, Jan. 2014, doi: 10.5120/14866-2958.

Appendix A

List of Publications

As an author

- ❖ D. L. Prasanna and S. L. Tripathi, "Machine Learning Classifiers for Speech Detection," *2022 IEEE VLSI Device Circuit and System (VLSI DCS)*, Kolkata, India, 2022, pp. 143-147, doi: 10.1109/VLSIDCS53788.2022.9811452.
- ❖ D. L. Prasanna and S. L. Tripathi, "Machine and Deep-Learning techniques for text and speech processing," in *Machine Learning Algorithms for Signal and Image Processing*, 2022, pp. 115–128. doi: 10.1002/9781119861850.ch7.
- ❖ D. L. Prasanna and S. L. Tripathi, "A review on tongue based assistive technology, devices and FPGA processors using Machine Learning module," *Wireless Personal Communications*, vol. 134, no. 1, pp. 151–170, Jan. 2024, doi: 10.1007/s11277-024-10897-8.
- ❖ D. L. Prasanna, S. L. Tripathi, M. Mahmud, and I. Wijayanto, "High attenuation electromagnetic interface filter for effective processing of audio signals," *Engineering Research Express*, vol. 6, no. 3, p. 035309, Jul. 2024, doi: 10.1088/2631-8695/ad5fdf.
- ❖ D. L. Prasanna, S. L. Tripathi and I. Wijayanto, "Classifying the Tongue Contours for Assistive Devices," *2024 IEEE International Conference of Electron Devices Society Kolkata Chapter (EDKCON)*, Kolkata, India, 2024, pp. 1-5, doi: 10.1109/EDKCON62339.2024.10870771.

As a co-author

- ❖ I. Wijayanto, A. Humairani, S. Hadiyoso, A. Rizal, D. L. Prasanna, and S. L. Tripathi, "Epileptic seizure detection on a compressed EEG signal using energy measurement," *Biomedical Signal Processing and Control*, vol. 85, p. 104872, Mar. 2023, doi: 10.1016/j.bspc.2023.104872.

- ❖ S. L. Tripathi, D. L. Prasanna, and M. Mahmud, “A Low-Power Audio Processing Using Machine Learning Module on FPGA and Applications,” in *Explainable Machine Learning Models and Architectures*, Wiley, 2023, pp. 49–63. doi: 10.1002/9781394186570.ch4.
- ❖ S. L. Tripathi, L. P. Dasari, I. Wijayanto, D. Ghai, and M. Mahmud, “Assistive technology for neuro-rehabilitation applications using machine learning techniques,” in *Machine Learning Models and Architectures for Biomedical Signal Processing*, Elsevier eBooks, 2024, pp. 179–191. doi: 10.1016/b978-0-443-22158-3.00008-9.

Appendix B

List of Courses and Certifications

- ❖ **RISC-V VLSI Implementation Flow: RTL2GDS**, Faculty Development Programme (FDP) conducted by Electronics & ICT Academy supported by Ministry of Electronics and Information Technology (MeitY), Government of India, April 2021.
- ❖ **System Design Methodologies for Embedded, IoT, AI, & HPC using Intel FPGA**, Faculty Development Programme (FDP) conducted by Electronics & ICT Academy supported by Ministry of Electronics and Information Technology (MeitY), Government of India, April 2021.
- ❖ **GaN on Silicon front-end electronics in advanced microwave and millimeter-wave systems for space and defence**, conducted by Prof. Ernesto Limiti, University of Rome, Tor -Vergata, Italy organized by Electronics and Communication Engineering held on August 26th, 2021.
- ❖ Short Term Course on **VLSI Design Tools** organized by Lovely Professional University w.e.f. June 27, 2022, to July 02, 2022
- ❖ **Learn the Fundamentals of VHDL and FPGA Development**, Udemy.
- ❖ **Digital Systems: From Logic Gates to Processors**, Universitat Autònoma de Barcelona, Coursera.
- ❖ **VHDL Circuit Design and FPGAs with VIVADO and MODELSIM**, Udemy.
- ❖ **Machine Learning**, Stanford University, Coursera.
- ❖ **FPGA Design with MATLAB & Simulink**, Udemy.
- ❖ **I2C, SPI, UART (RS232), VGA IN VHDL for FPGA interfacing**, Udemy.
- ❖ **Fundamentals of Digital Image and Video Processing**, Northwestern University, Coursera.
- ❖ **Research Exchange Program**, Telkom University, Bandung, Indonesia 2023.
- ❖ Workshop on **Explainable Artificial Intelligence Powered Tools for the Early Detection and Management of Dementia**, Co-Organized by Cognitive Computing and Brain Informatics Research Group of Nottingham Trent

University (UK) and School of Electronics and Electrical Engineering, Lovely Professional University (India) with support from Applied Intelligence and Informatics Lab, UK, June 2024.

Certificates



Ref. no: EICT/2020-21/JFDP/RISV/087

Electronics & ICT Academy
Supported by Ministry of Electronics and Information Technology (MeitY), Govt. of India
Indian Institute of Technology Guwahati

Participation Certificate

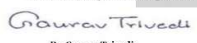


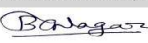
Mr./Ms./Dr. _____

This is to certify that **Ms. Dasari Lakshmii Prasanna**

of **Lovely Professional University**

has participated in a two-week Online Faculty Development Programme on "RISC-V VLSI Implementation Flow: RTL2GDS" jointly organised by Electronics and ICT Academies held from **27 March - 10 April, 2021** under the "Scheme of financial assistance for setting up of Electronics and ICT Academies" of the Ministry of Electronics and Information Technology (MeitY), Government of India. He/she has also taken part in the hands on sessions during the program. We wish him/her a very good luck for his/her future endeavour. This Faculty Development Programme is at par with other Quality Improvement Programmes*.


Dr. Gaurav Trivedi
Principal Coordinator & Co-PL E&ICT Academy
IIT Guwahati
Assam


Dr. Bal Chand Nagar
Joint Principal Coordinator
NIT Patna
Bihar


Dr. Meena Panchore
Joint Principal Coordinator
NIT Patna
Bihar


Prof. D. Boudchandani
Joint Principal Coordinator
MNIT Jaipur
Rajasthan


Dr. Deepak Bharti
Joint Principal Coordinator
MNIT Jaipur
Rajasthan


Prof. P. N. Kondekar
Joint Principal Coordinator
IITDM Jabalpur
Madhya Pradesh

* As per minutes of the 3rd PRSG.



Supported by Ministry of Electronics and Information Technology (MeitY), Govt. of India
Indian Institute of Technology Guwahati

Ref. no: EICT/2021-22/JFDP/SD&IoT/551

Participation Certificate



Mr./Ms./Dr. _____

This is to certify that
Dasari Lakshmi Prasanna

of

Lovely Professional University

has participated in a two-week Online Faculty Development Programme on "System Design Methodologies for Embedded, IoT, AI, & HPC using Intel FPGA" jointly organised by Electronics and ICT Academies held from 19 - 30 April, 2021 under the "Scheme of financial assistance for setting up of Electronics and ICT Academies" of the Ministry of Electronics and Information Technology (MeitY), Government of India. He/she has also taken part in the hands on sessions during the program. We wish him/her a very good luck for his/her future endeavour. This Faculty Development Programme is at par with other Quality Improvement Programmes*.

Gaurav Trivedi

Dr. Gaurav Trivedi
Principal Coordinator and Co-PI, E&ICT Academy
IIT Guwahati
Assam

Sangeeta

Dr. Sangeeta Singh
Joint Principal Coordinator
NIT Patna
Bihar

Amit

Dr. Amit M. Joshi
Joint Principal Coordinator
MNIT Jaipur
Rajasthan

Pankaj Sharma

Dr. Pankaj Sharma
Joint Principal Coordinator
IIITDM Jabalpur
Madhya Pradesh

Satish

Dr. Satish Maheshwaram
Joint Principal Coordinator
NIT Warangal
Telangana

* As per minutes of the 3rd PRSG.

NATIONAL INSTITUTE OF TECHNICAL TEACHERS TRAINING AND RESEARCH



(Ministry of Education, Government of India)
Sector-26, Chandigarh-160019

Cert. No.: PDA-35/47/21-22

CERTIFICATE OF PARTICIPATION

This is to certify that Dr./Mr./Ms. **Dasari Lakshmi Prasanna** . Has Attended online webinar on "GaN on Silicon front-end electronics in advanced microwave and millimeter-wave systems for space and defence", conducted by Prof. Ernesto Limiti, University of Rome, Tor-Vergata, Italy organized by Electronics and Communication Engineering held on August 26th, 2021.

S.S. Gill

PROF.(DR.) S.S. GILL
Coordinator

B. Raj

DR. BALWINDER RAJ
Coordinator



HUMAN RESOURCE DEVELOPMENT CENTER

[Under the Aegis of Lovely Professional University, Jalandhar-Delhi G.T Road, Phagwara (Punjab)]



Certificate No. 251707

Certificate of Participation

This is to certify that **Ms. Dasari Lakshmi Prasanna** D/o Sh. Dasari Obulesh participated in **Short Term Course on VLSI Design Tools** organized by Lovely Professional University w.e.f. **June 27, 2022 to July 02, 2022** and obtained "O" Grade.

Date of Issue: 02-07-2022
Place : Phagwara (Punjab), India

Prepared by
(Administrative Officer-Records)

Checked By
Program Coordinator

Head
Human Resource Development Center



SCHOOL OF
**ELECTRONICS AND
ELECTRICAL ENGINEERING**



Certificate No. 338606

Certificate of Participation

This is to certify that **Dr./Mr./Ms. Dasari Lakshmi Prasanna** of **SEEE, Lovely Professional University** has participated in the Co-Design Workshop on "Explainable Artificial Intelligence Powered Tools for the Early Detection and Management of Dementia" held on 19th June 2024 Co-Organized by Cognitive Computing and Brain Informatics Research Group of **Nottingham Trent University (UK)** and **School of Electronics and Electrical Engineering, Lovely Professional University (India)** with support from **Applied Intelligence and Informatics Lab, UK**.

Date of Issue : 06-07-2024

Place : Phagwara (Punjab), India

Prepared by
(Administrative Officer-Records)

Dr. Suman Lata Tripathi
Coordinator Professor
SEEE, LPU, India

Dr. Mufti Mahmud
Director, CCBI Research
Group, Nottingham Trent
University, UK

Dr. Gaurav Sethi
Professor & Head of School
LPU, India

2024 IEEE Electron Device Kolkata Conference



IEEE

EDKCON 2024



Certificate of Presentation

This is to certify that

Dr./Mr./Ms. Dasari Lakshmi Prasanna

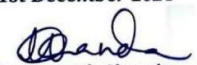
of Lovely Professional University

presented a paper

*in the International Conference on '2024 IEEE Electron Device Kolkata Conference organized by
IEEE EDS Kolkata Chapter held at Kolkata, India during 30 November - 1st December 2024*


Dr. Arpan Dasgupta
General Co-Chair, IEEE EDKCON 2024
Chair, IEEE EDS Kolkata Chapter


Prof. (Dr.) Angsuman Sarkar
General Co-Chair, IEEE EDKCON 2024
Ex-Chair, IEEE EDS Kolkata Chapter


Dr. Manash Chanda
General Co-Chair, IEEE EDKCON 2024
Immediate Past Chair, IEEE EDS
Kolkata Chapter

FAKULTAS TEKNIK ELEKTRO
PROGRAM STUDI S2 TEKNIK ELEKTRO
SCHOOL OF ELECTRICAL ENGINEERING
MAGISTER OF ELECTRICAL ENGINEERING

TERAKREDITASI / ACCREDITED : SK NO. 1601/SK/BAN-PT/AK-ISK/M/III/2022

TRANSKRIP AKADEMIK
ACADEMIC TRANSCRIPT

No 0000/AKD.1.14/SFTE.2.02/2023

NAMA / NAME : DASARI LAKSHMI PRASANNA
 NOMOR INDIK MAHASISWA / STUDENT'S ID : 2101221068
 TEMPAT & TANGGAL LAHIR / PLACE & DATE OF BIRTH : HINDHUPUR, ANDHRA PRADESH, 20 JULI 1995 / 20 JULY 1995
 TANGGAL LULUS / DATE OF COMPLETION : /
 STATUS MAHASISWA / STUDENT STATUS : STUDENT EXCHANGE

NO	KODE/ CODE	MATA KULIAH	SUBJECT	SEMESTER/ SEMESTER	SKS/ CREDIT	NILAI/ GRADE
1	TTI6Z2	PROYEK, SEMINAR DAN PUBLIKASI	PROJECT, SEMINAR AND PUBLICATION	1	2	A
2	TEI6K4	TOPIK PENELITIAN 2	RESEARCH TOPICS 2	2	4	A

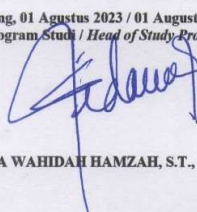
KETERANGAN NILAI / GRADE NOTICE

A = 4 AB = 3.5 B = 3 BC = 2.5 C = 2 D = 1 E = 0

PENCAPAIAN AKADEMIK / ACADEMIC ACHIEVEMENT

SKS (Satuan Kredit Semester) Total / Total Credits : 6 SKS / Credits
 IPK (Indeks Prestasi Kumulatif) / Grade Point Average : 4,00 (dari / out of 4,00)

Bandung, 01 Agustus 2023 / 01 August 2023
 Ka. Program Studi / Head of Study Program


 Dr. IDA WAHIDAH HAMZAH, S.T., M.T.

Halaman 1 dari 1 / Page 1 of 1