Human-Computer Interaction using Neuromuscular Signals

Team Members:

Rabin Nepal [073/BEX/331]

Rhimesh Lwagun [073/BEX/333]

Sanjay Rijal [073/BEX/342]

Upendra Subedi [073/BEX/347]

Project Supervisor:

Mr. Dinesh Baniya Kshatri

Lecturer

Department of Electronics and Computer Engineering

Institute of Engineering, Thapathali Campus

March 5, 2021

Presentation Outline

- Motivation
- Project Objectives
- Scope of Project
- Project Applications
- Instrumentation
- Self-Recorded EMG Dataset
- Methodology
- Results
- Analysis and Discussion
- Future Enhancements
- Conclusion
- Appendix (Project Budget)
- Appendix (Project Timeline)
- References

Motivation

Project Objectives

- To extract and transfer EMG signals from articulatory muscles to a computer

- To process and convert the articulated speech signals to text

Scope of Project

- **Project Capabilities:**
    - Detects words that a person is articulating internally
    - Displays the articulated words on monitor as text

- **Project Limitations:**
    - Detection of sentences and symbols
    - Two way interaction between user and computer

Project Applications

- **Silent Means of Communication**
    - Covert systems where privacy is crucial
- **Novel Human Computer Interface**
    - New interaction method for speech impaired or general people
- **Improvement of Speech Recognition Models**
    - Additional information about muscles involved during speech

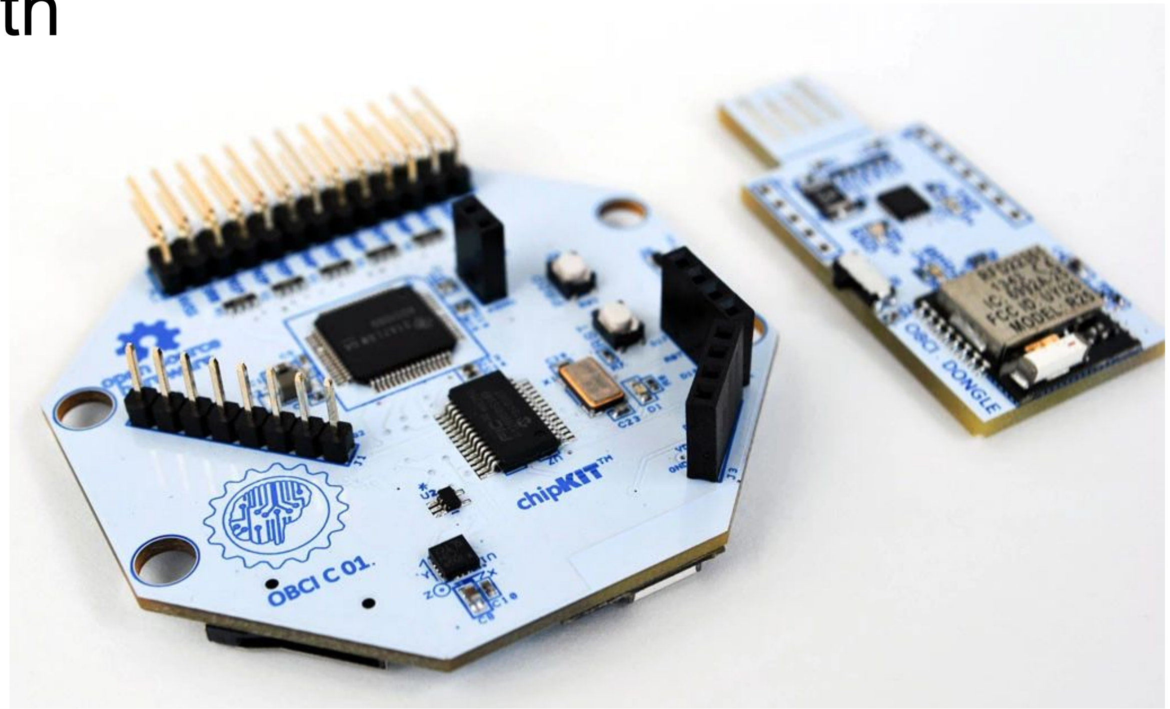
Instrumentation-1
[Hardware Components]

- **EMG Electrodes**
  - Gold-plated, cup shaped
  - Reusable, high selectivity and non-invasive

- **Electrolyte**
  - Ten20 Conductive Paste
  - Reduces skin-electrode interface impedance
  - Specially manufactured for EEG, EMG signals

Instrumentation-2
[Hardware Components]

- **Cyton Board**
  - Has 8 channel, 24 bit ADC with built-in PGA
  - 32-bit μC as processor
  - Wirelessly transmits data through bluetooth module



Instrumentation - 3
[Software Platforms]

- **Python**
  - Mainly used Scipy, Librosa, Scikit Learn, Tensorflow
  - Used for feature extraction and ML implementation

- **OpenBCI GUI**
  - Used for visualizing and recording EMG signals
  - Allows to configure gain and digital inputs of Cyton board

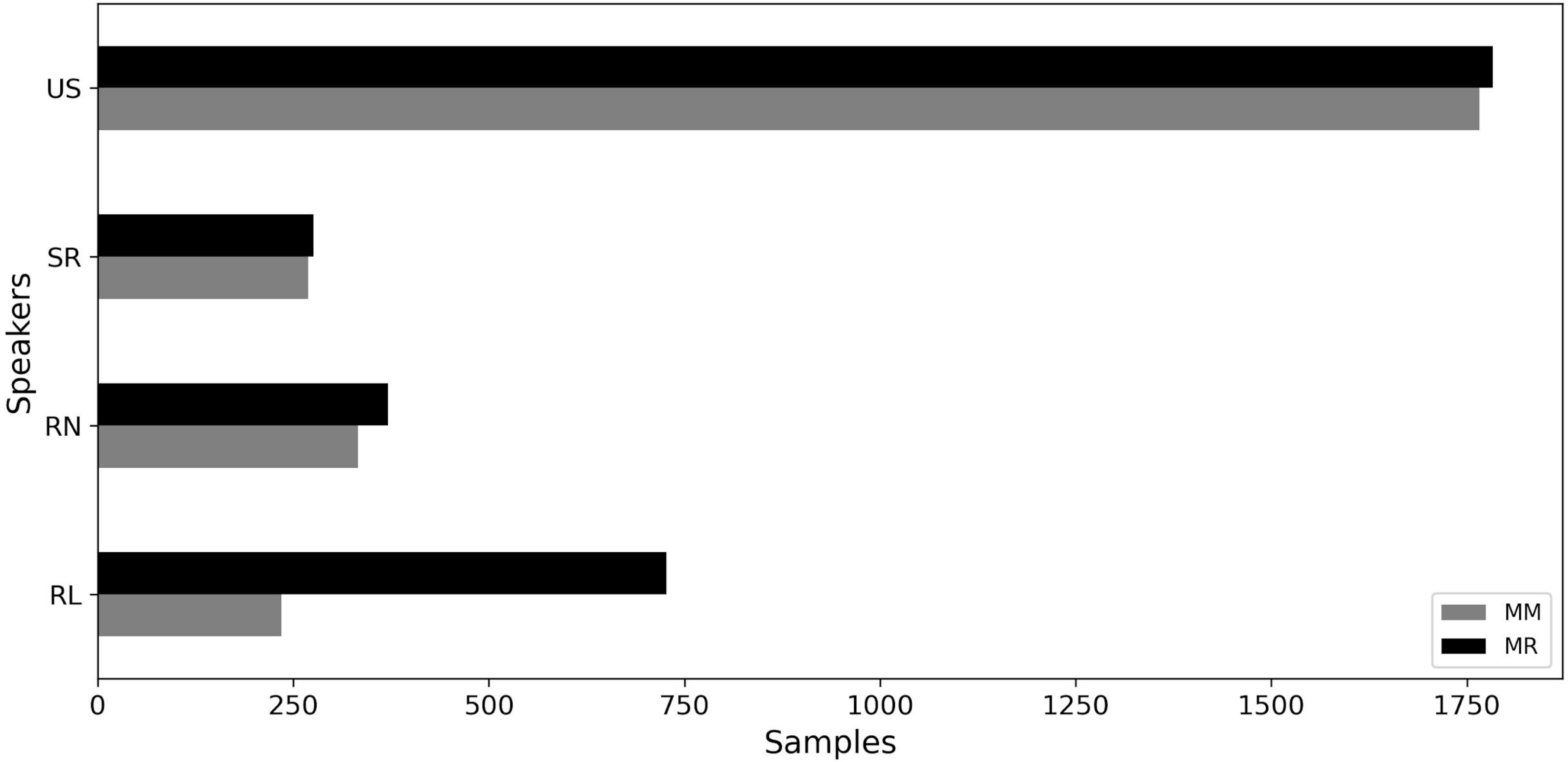
Self-Recorded EMG Dataset - 1
[Dataset Specifications]

Recording Mode	Sampling Rate (Hz)	Channels	Duration (hh:mm:ss)	Speaker Count	Session Count
MM	250	8	01:00:01	4	6
MR	250	8	01:31:59	4	6

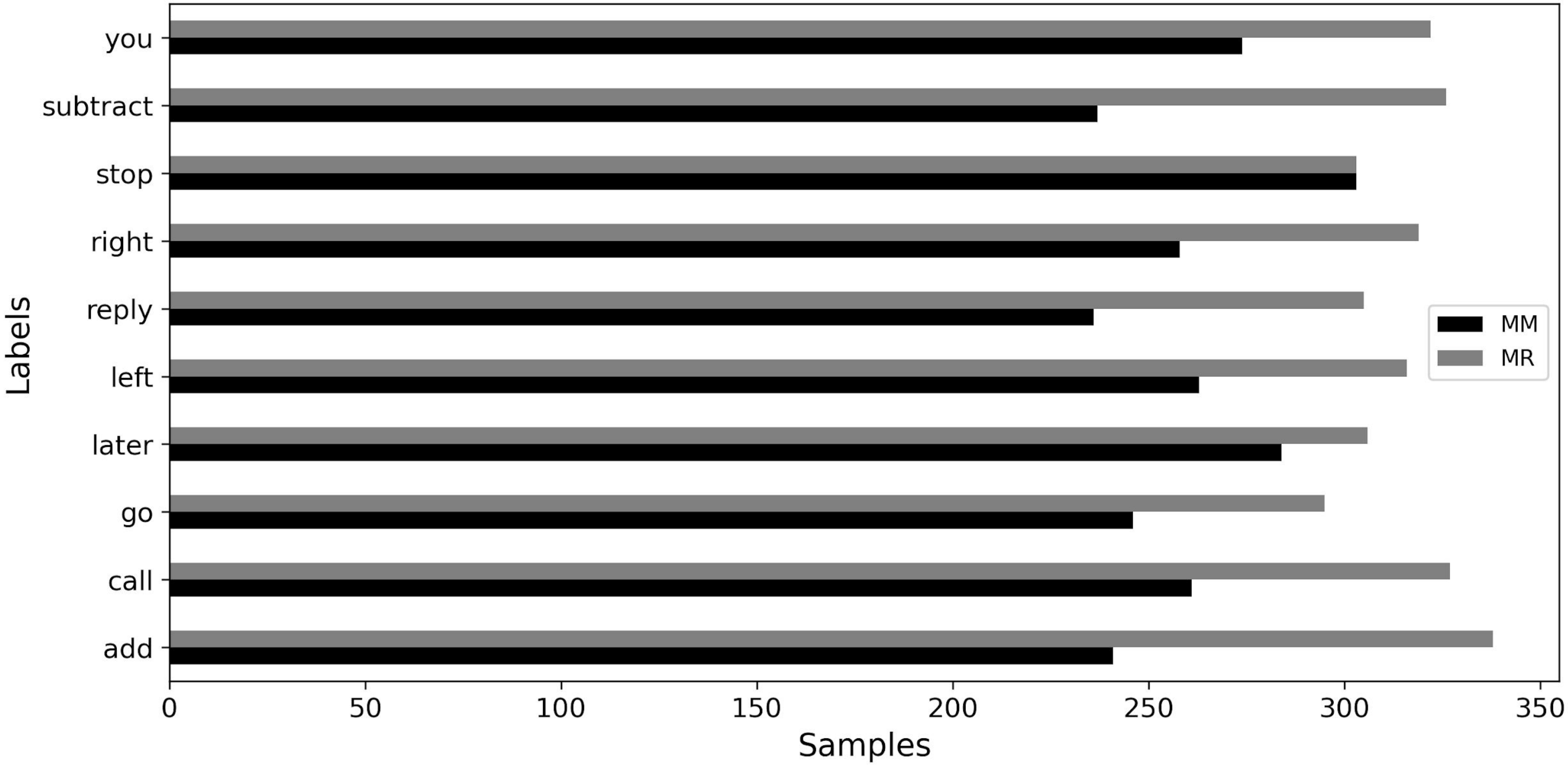
MM: Muscle Movement
MR: Mentally Rehearsed

Self-Recorded EMG Dataset - 2

[Samples per Speaker]

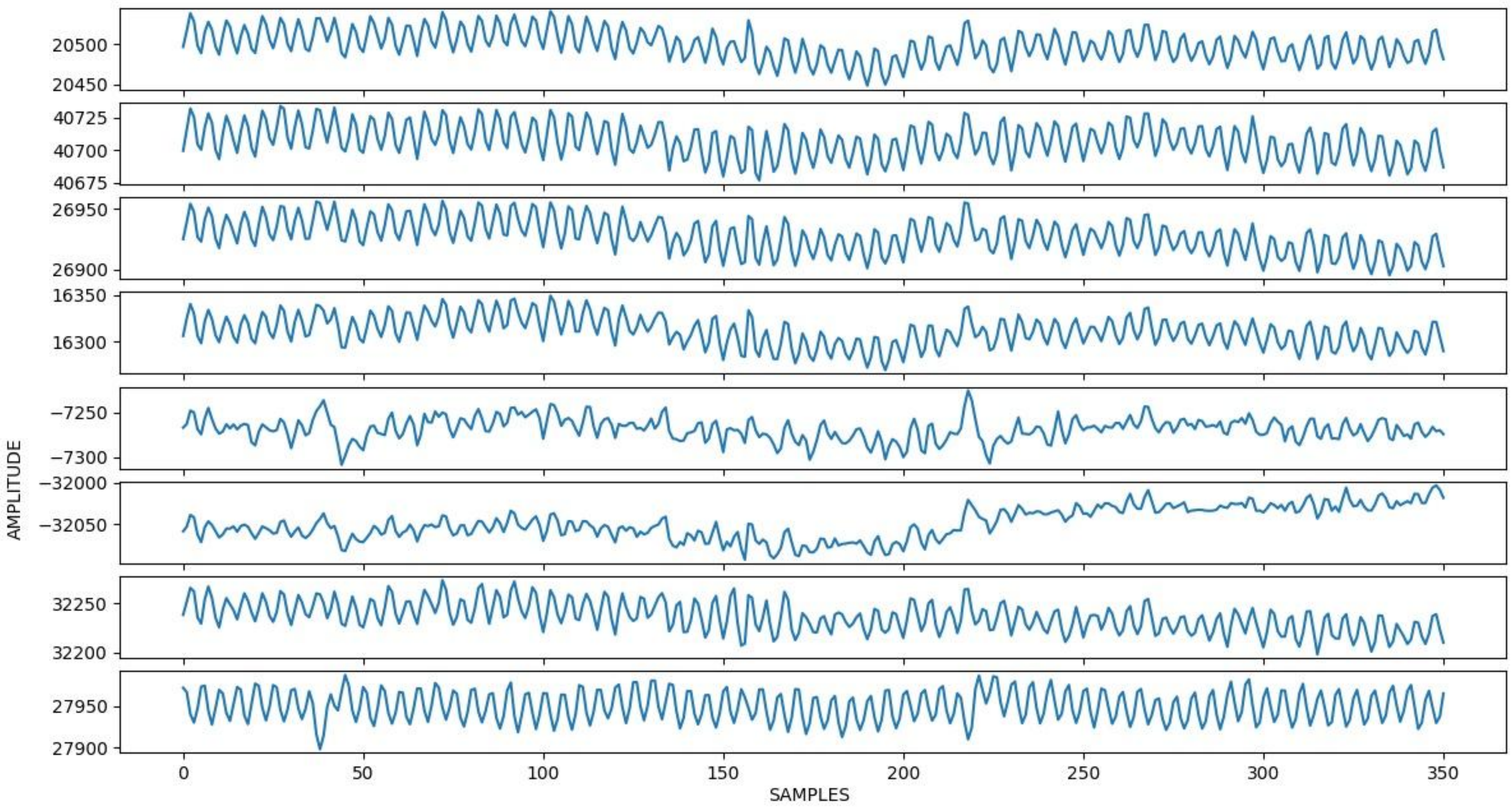


Self-Recorded EMG Dataset - 3
[Word Distribution]

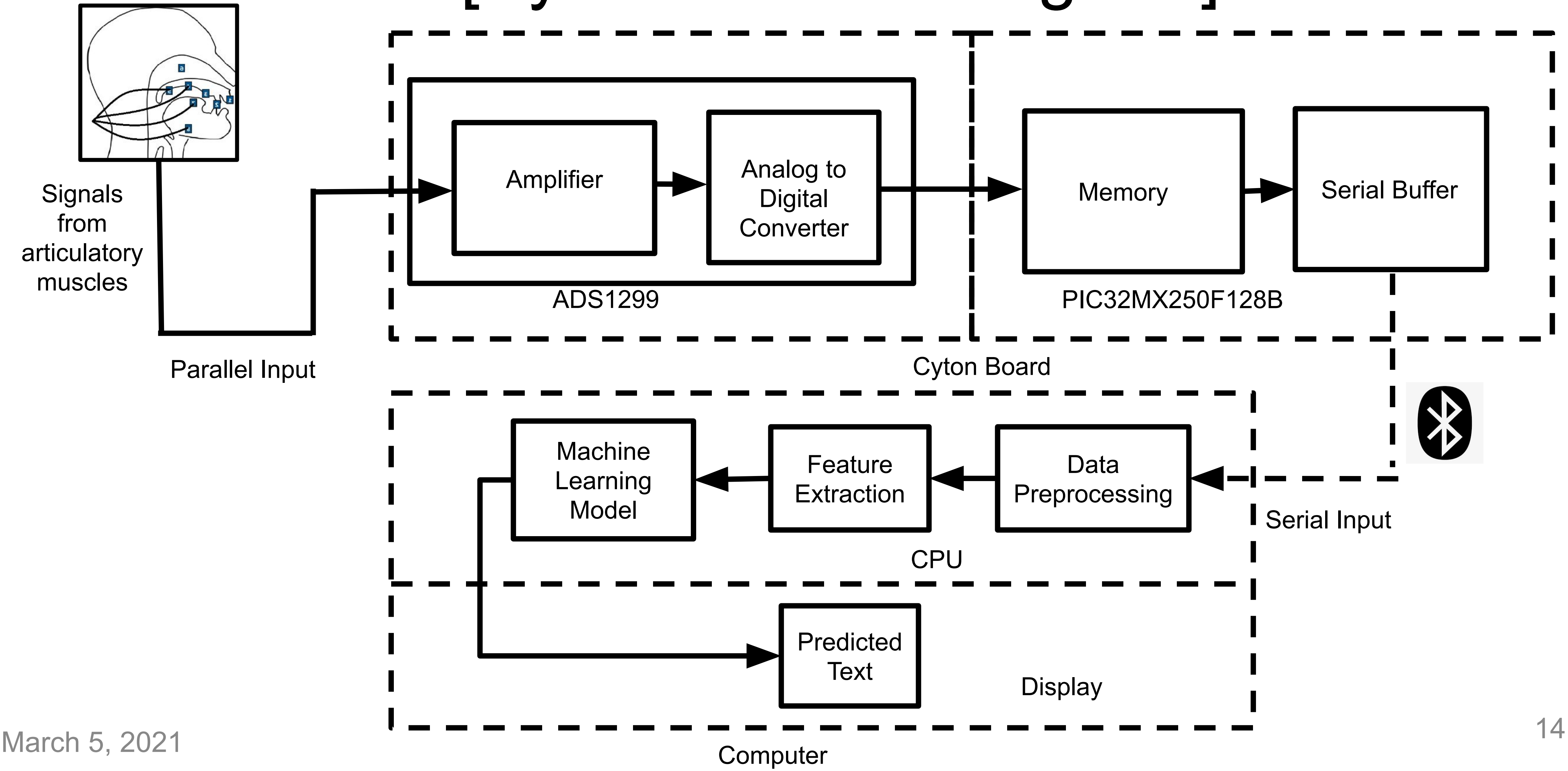


Self-Recorded EMG Dataset - 4

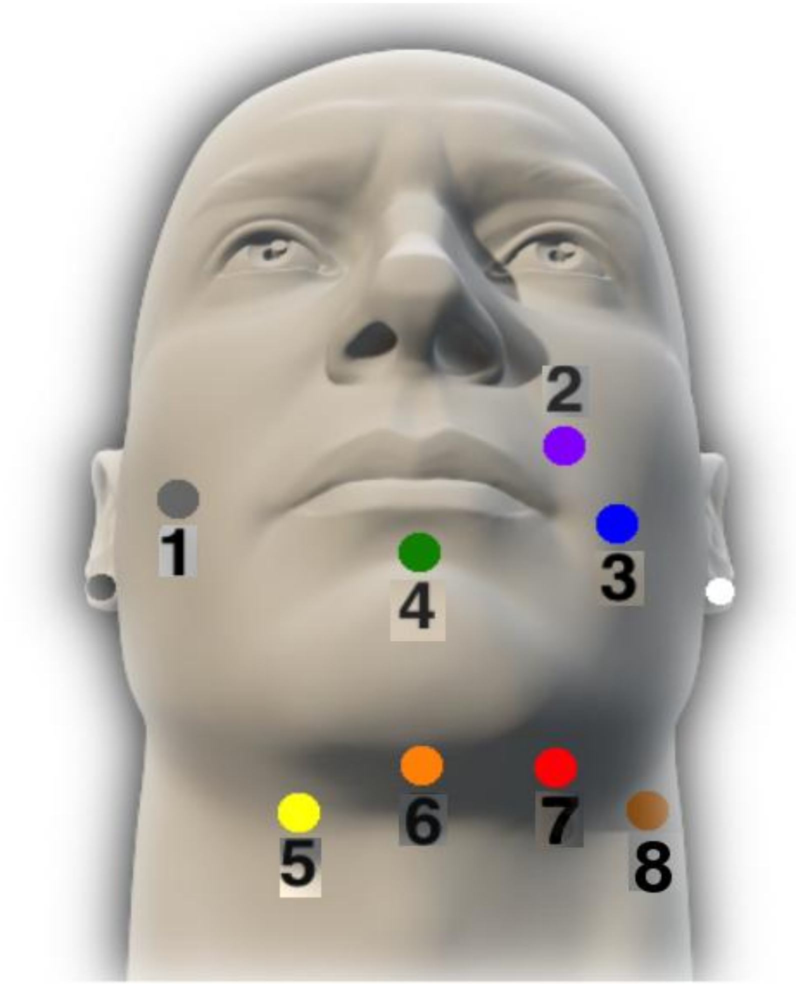
[Raw Signals for Word “CALL” in MR Mode]



Methodology - 1
[System Block Diagram]



Methodology - 2
[Muscle Selection]



EMG Channel	Muscle Name
1	Levator Angulis Oris
2	Zygomaticus Minor
3	Zygomaticus Major
4	Orbicularis Oris
5	Omohyoid
6	Anterior Belly of Digastric
7	Mylohyoid
8	Platysma

Methodology - 3
[Signal Acquisition and Transmission]

- Micro-volt EMG signals get amplified by ADS1299 with a gain of 24
- ADS1299 samples signal at frequency of 250Hz

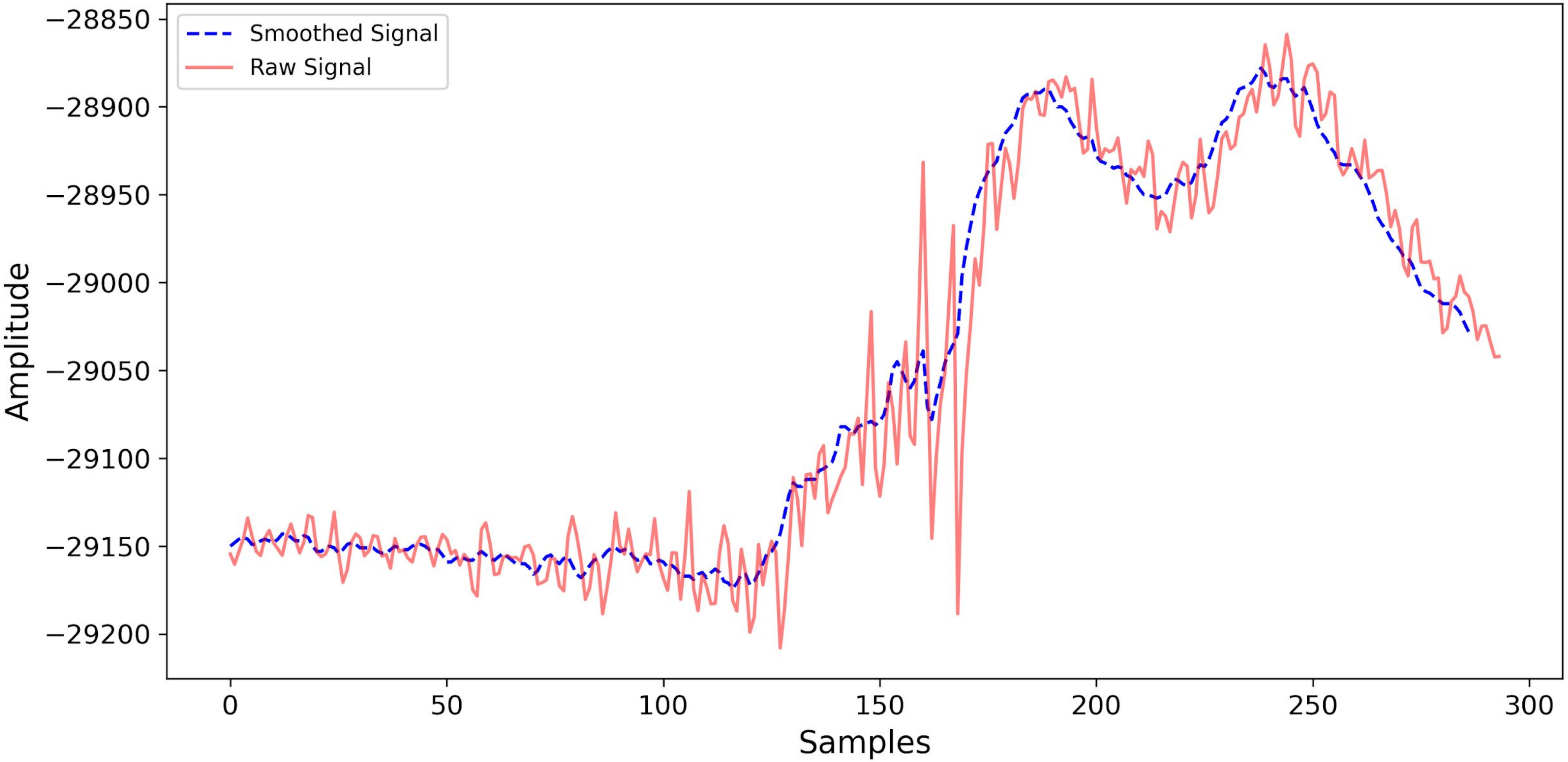
  Nyquist criteria ($f_s \geq 2 \times f_m$, $f_m = 50$ Hz)

- PIC32 receives data from ADS1299 through SPI
- Bluetooth device (RFD22301) sends data from PIC32 to computer

Methodology - 4
[Signal Smoothing and Normalization]

- **Signal Smoothing**
  - Smoothed by applying 8 point moving average
  - Provides resistance to sudden change in amplitude
- **Mean Normalization**
  - Correction of DC offset drift
  - Reduction of low frequency noise from the hardware

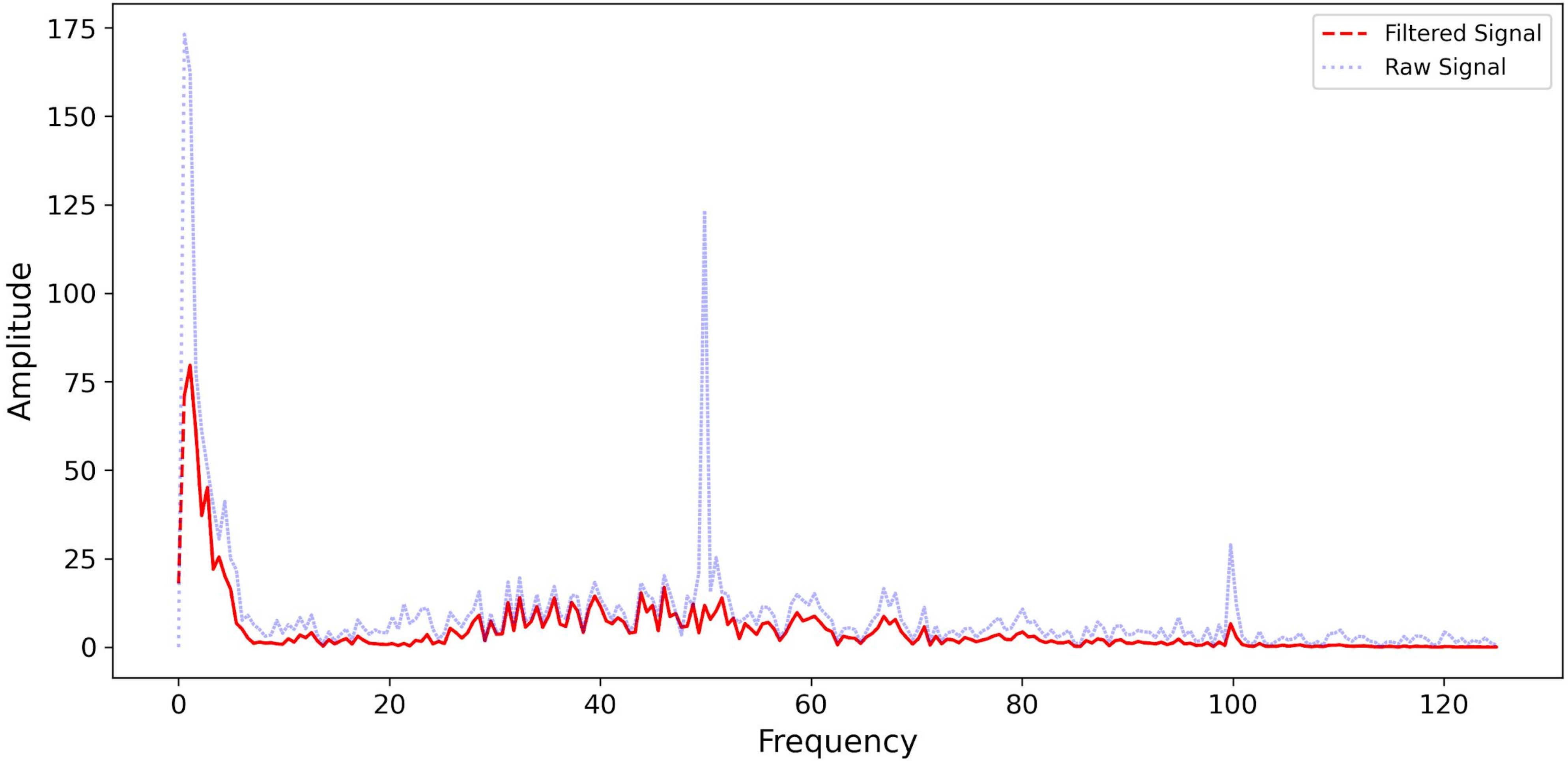
Methodology - 5
[Signal Smoothing and Normalization]



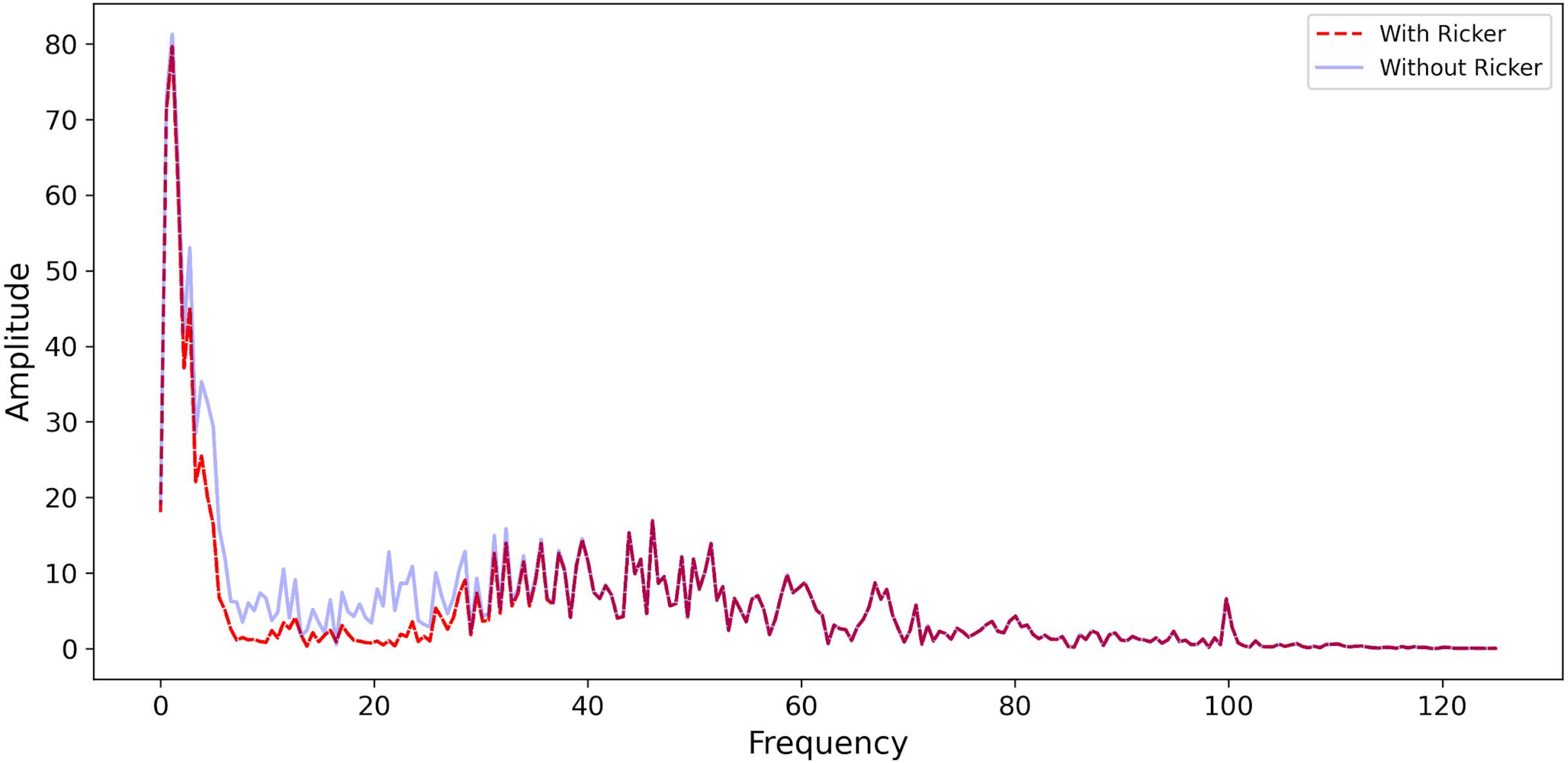
Methodology - 6
[Signal Filtration]

- First order digital Butterworth HPF of cut-off frequency at 1.5 Hz
- Notch filter to remove power line noise at 50 Hz and its harmonics
- First order digital Butterworth LPF of cutoff frequency at 50 Hz
- ECG artifacts suppressed by implementing Ricker wavelet

Methodology - 7
[Signal Filtration]



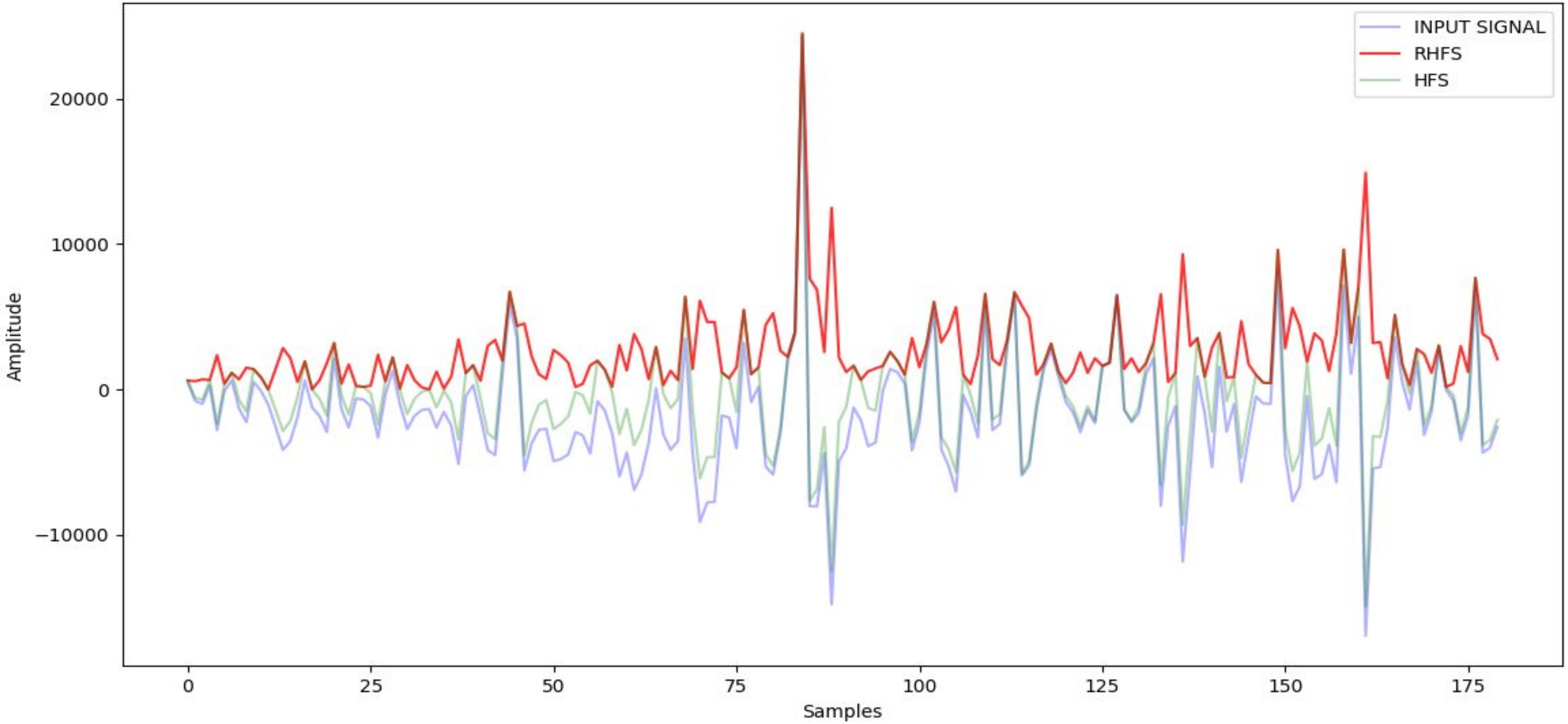
Methodology - 8
[Signal Filtration]



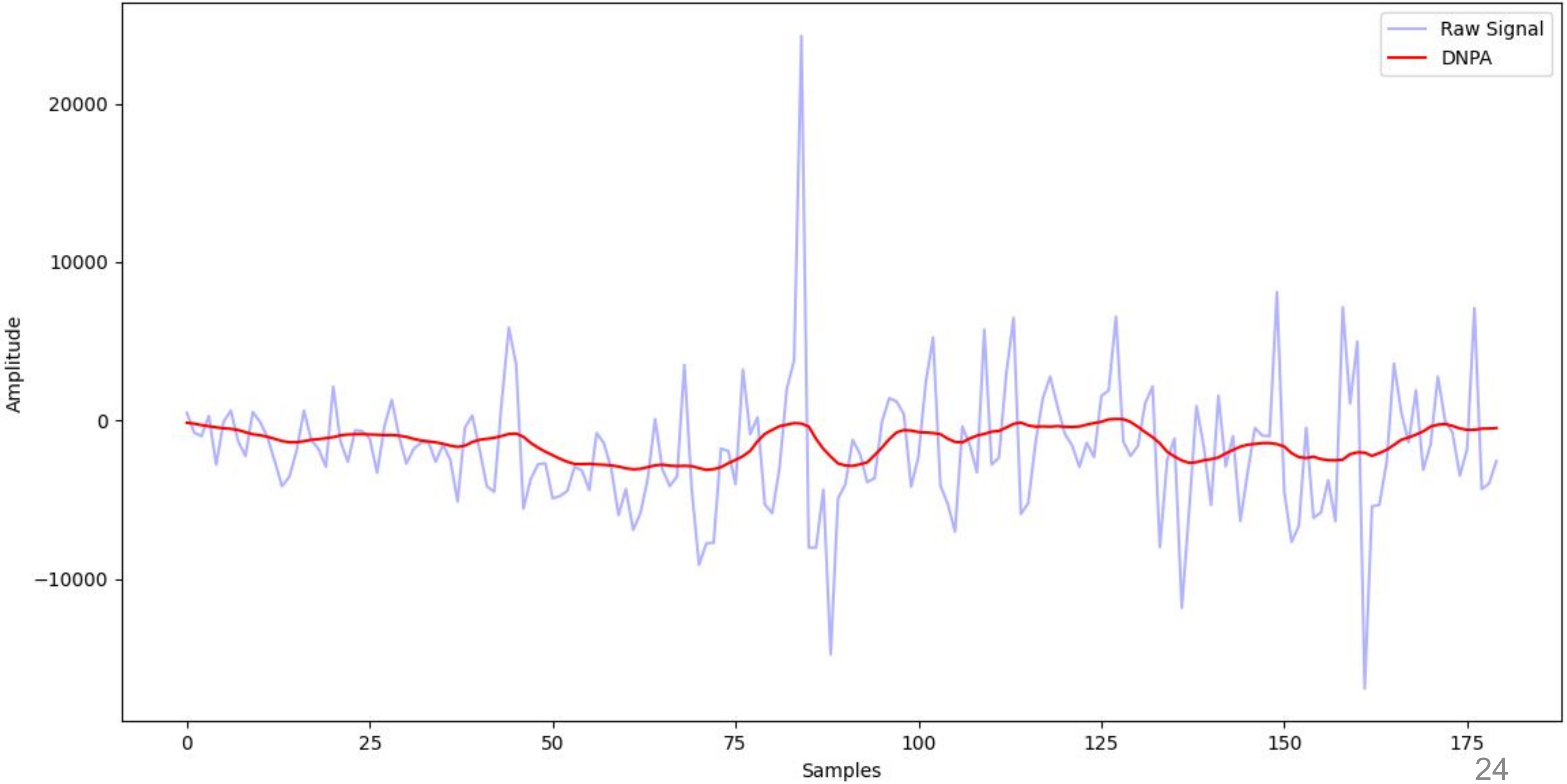
Methodology - 9
[Length Normalization and Feature Extraction]

- **Length Normalization**
  - Make data of equal length within 95th percentile
- **Feature Extraction**
  - **Temporal Features:** Zero Crossing Rate, Nine Point Double Average, High Frequency Signal and Frame Based Power
  - **Spectral Features:** STFT

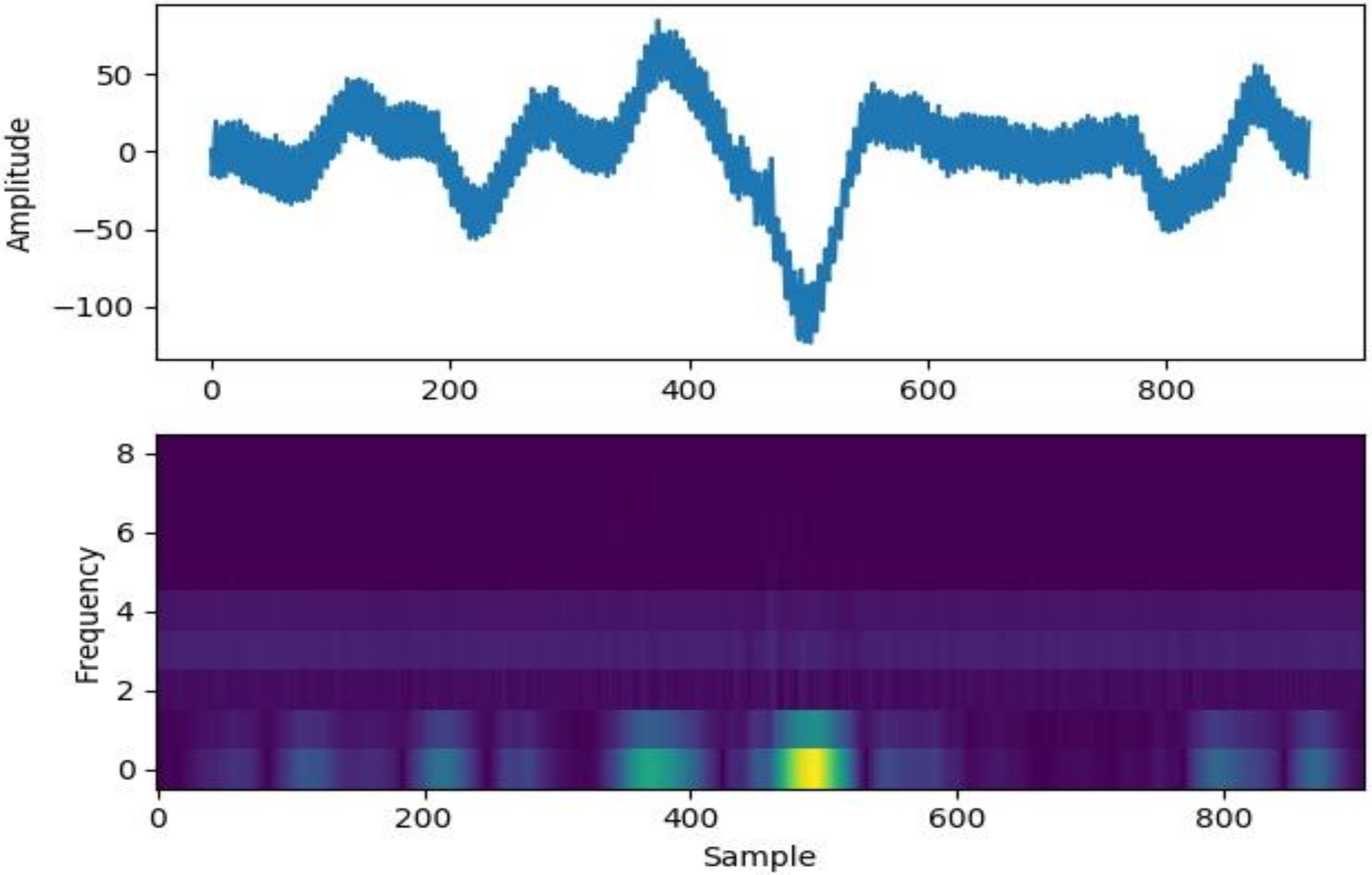
Methodology - 10
[Time Domain Features: HFS and RHFS]



Methodology - 11
[Time Domain Feature: DNPA]



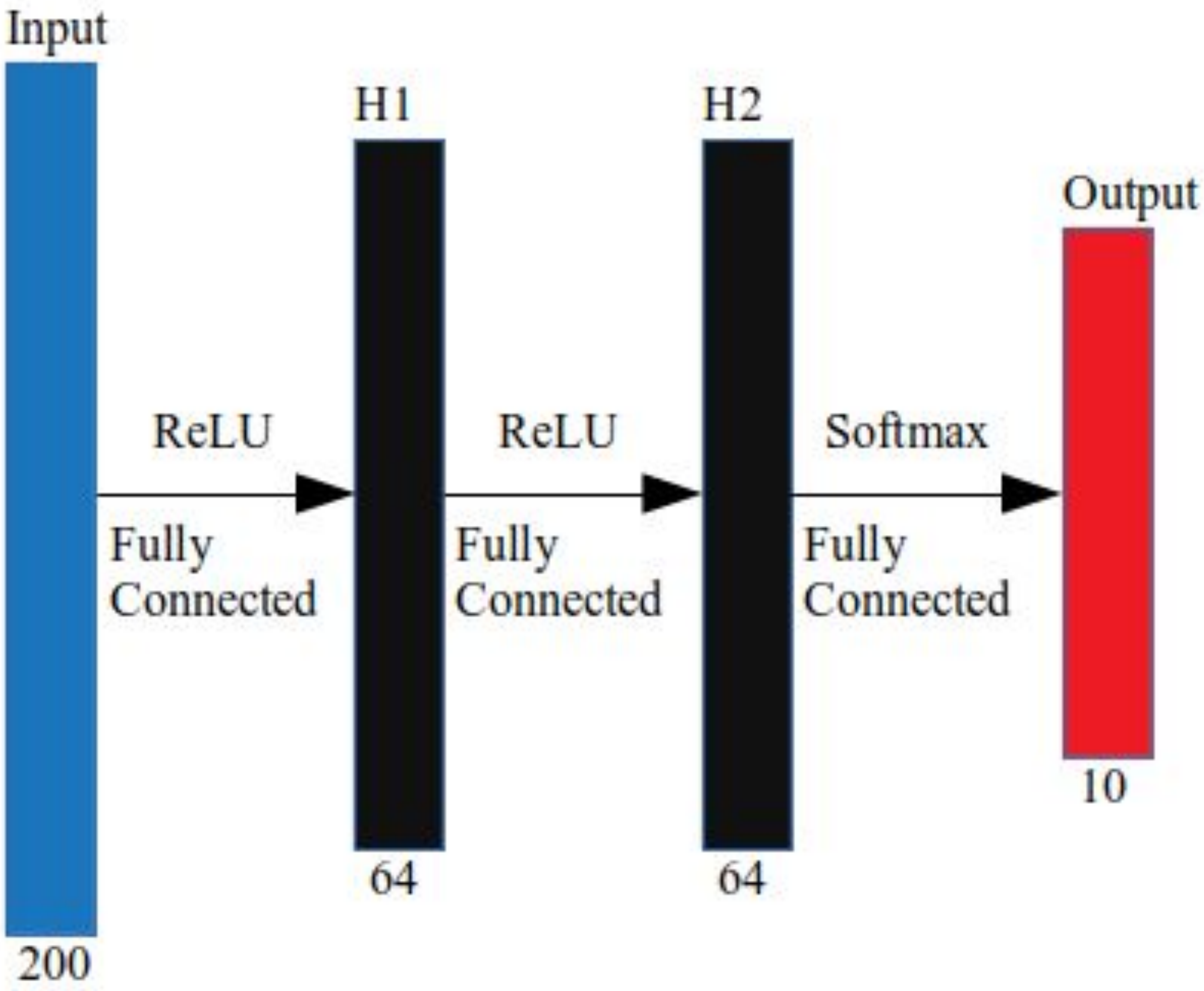
Methodology - 12
[Frequency Domain Feature: STFT]



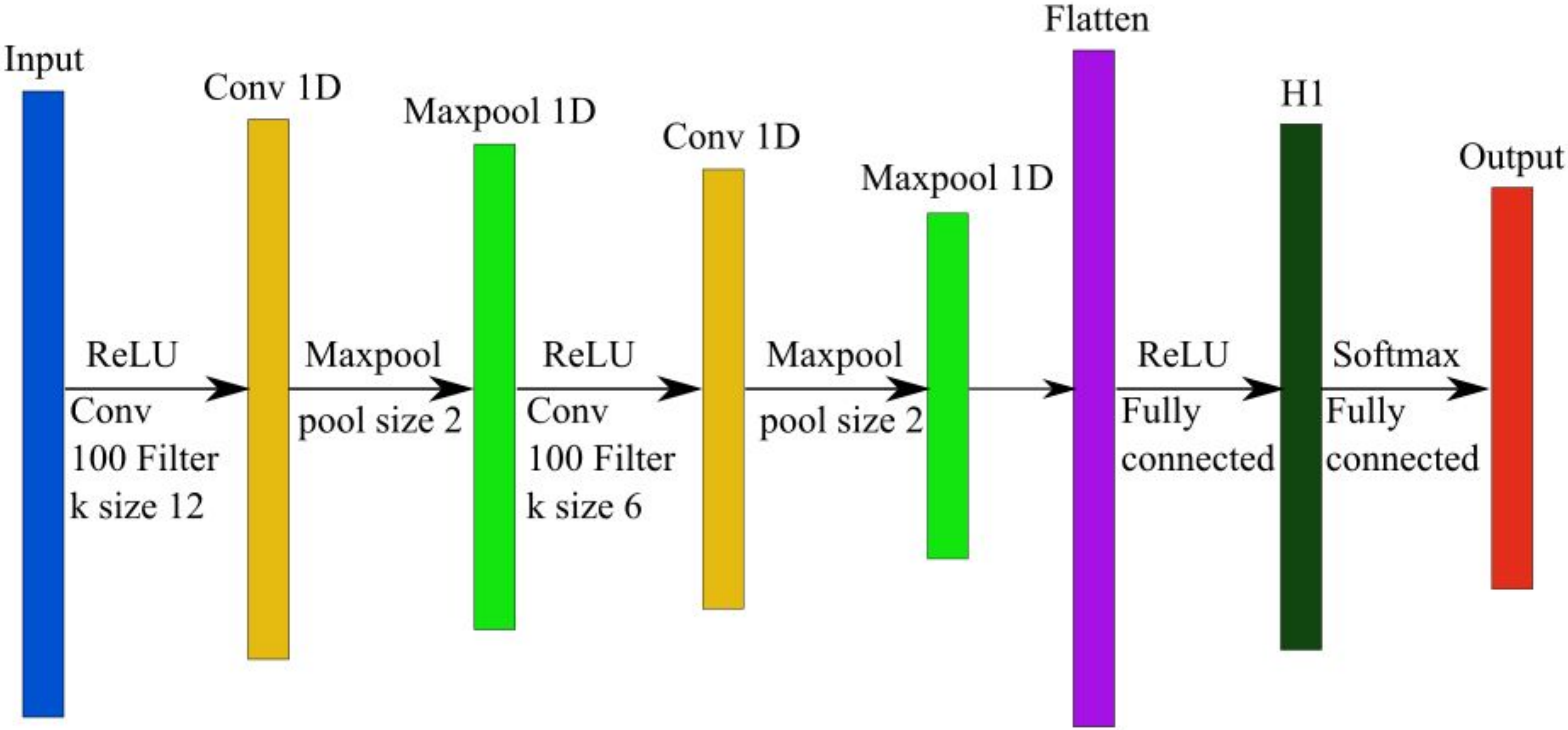
Methodology - 13
[Data Preparation]

- **Label Encoding**
  - Labels in string format changed to numeric value
- **Data Scaling**
  - Standardization (Z-score)
- **Dimension Reduction**
  - Independent Component Analysis (100 Components)

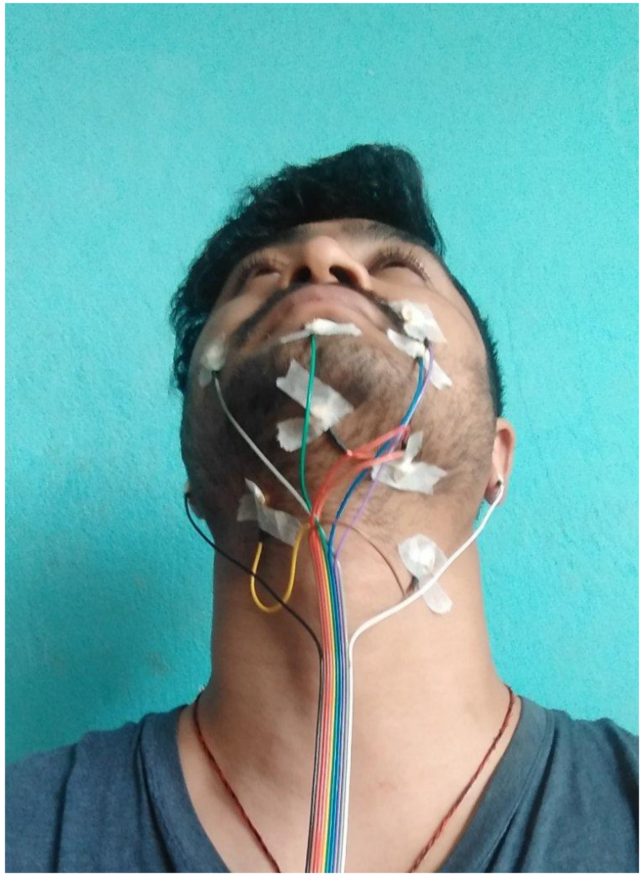
Methodology - 14
[MLP Architecture]

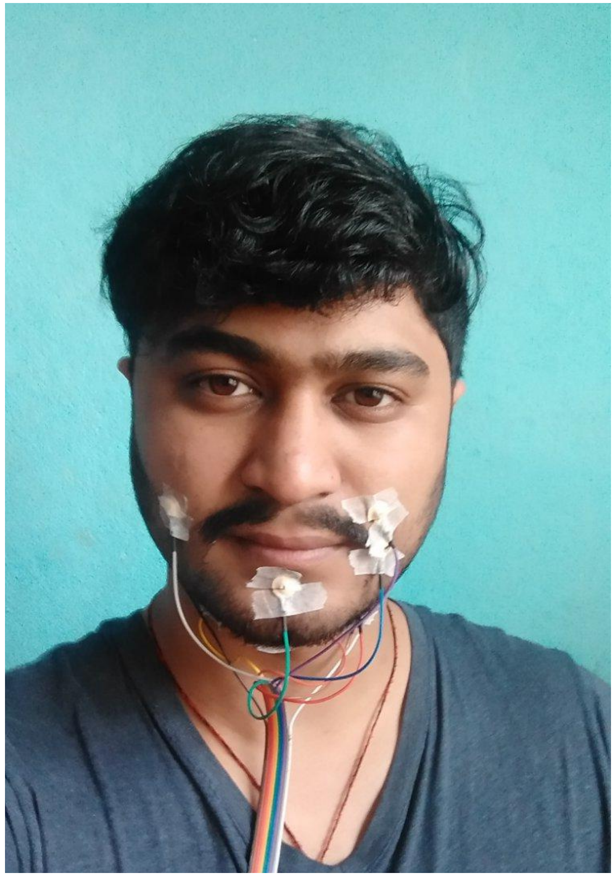


Methodology - 15
[CNN Architecture]



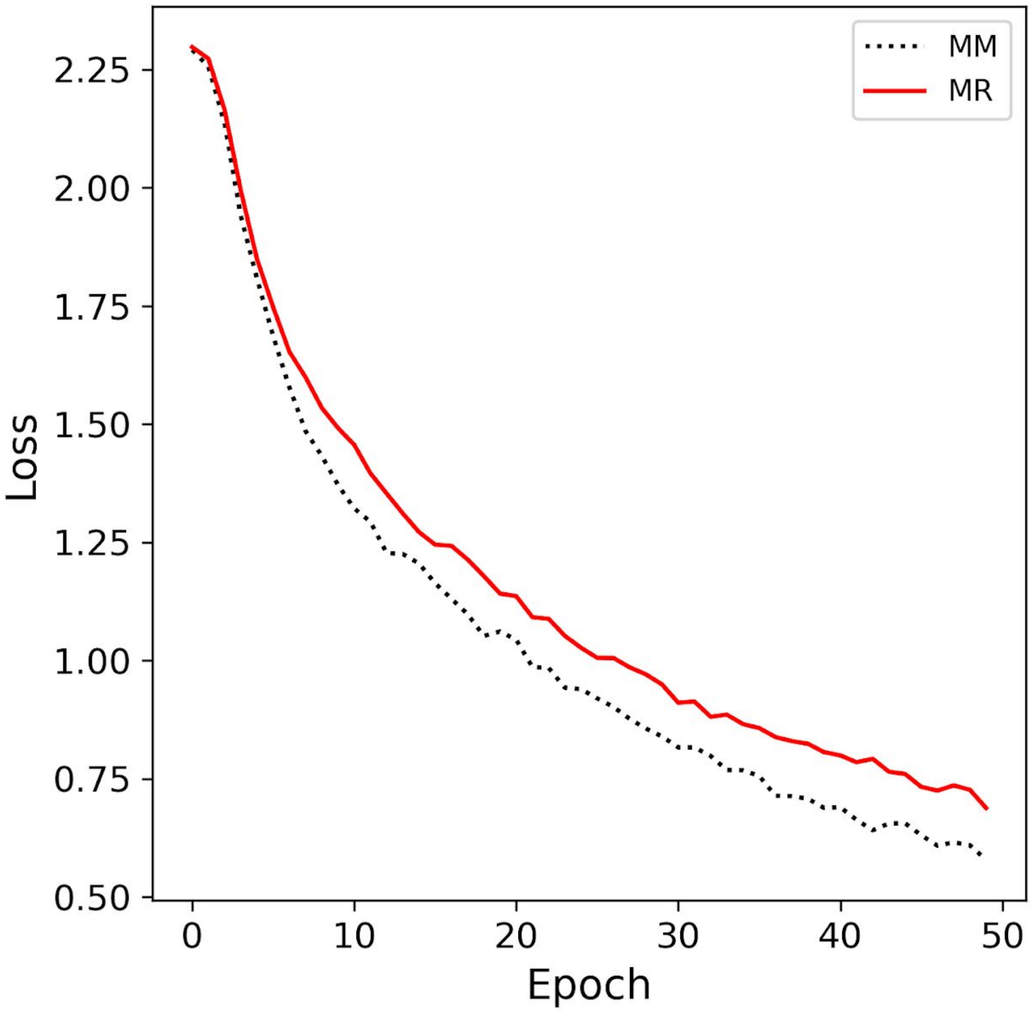
Results - 1
[OpenBCI Hardware Setup]

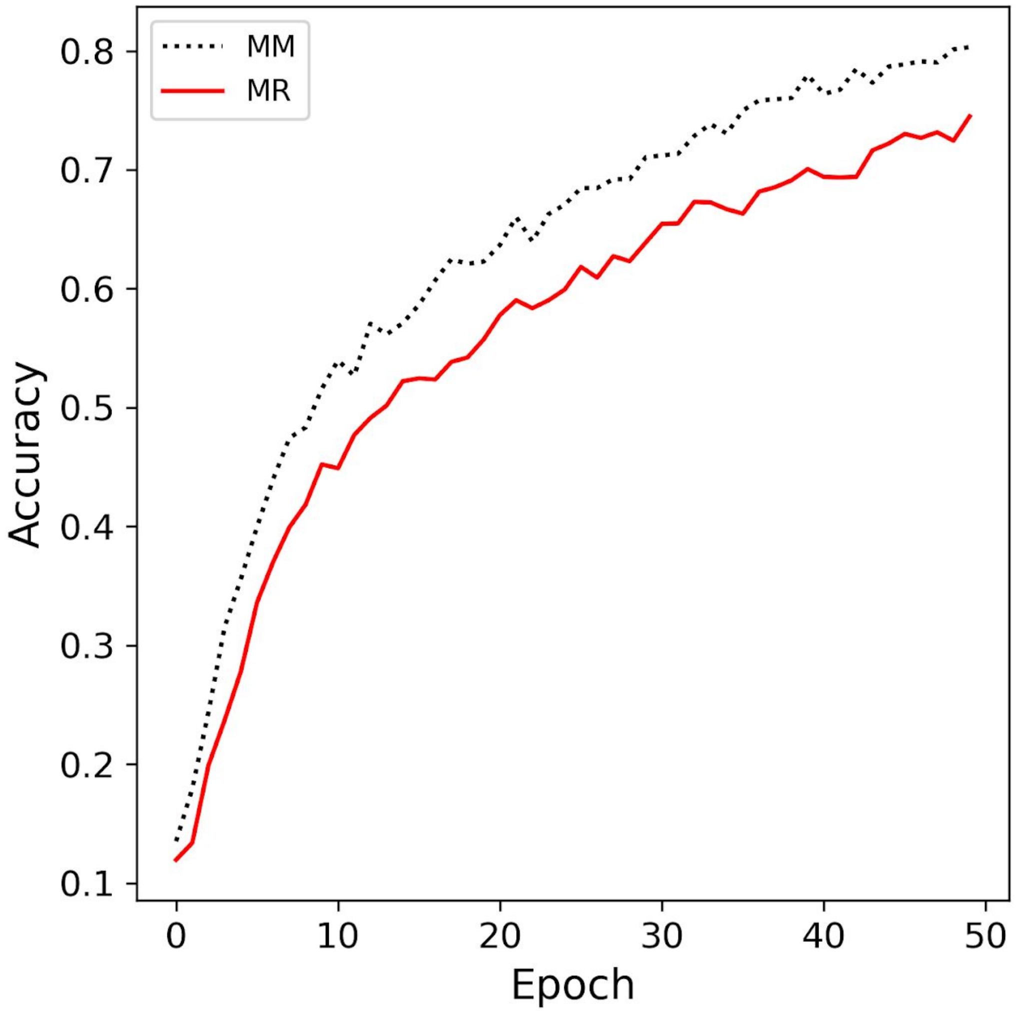




Results - 2
[OpenBCI GUI]

Results - 3
[MLP Accuracy and Loss]





Modes	Accuracy		Loss	
	Train	Validation	Train	Validation
Muscle Movement (MM)	0.8032	0.7265	0.5788	1.5239
Mentally Rehearsed (MR)	0.7449	0.6880	0.6883	1.2620

Results - 4
[Confusion Matrix from MLP]

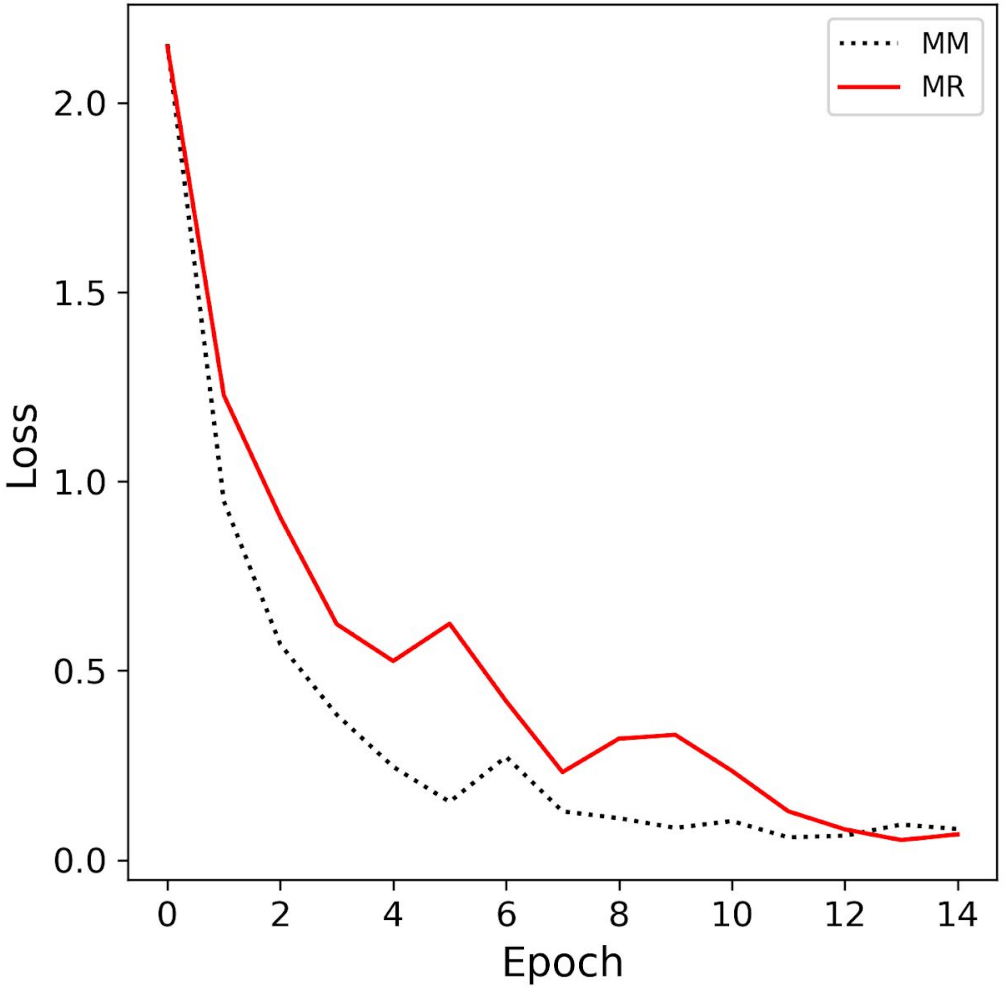
Actual Label \ Predicted Label	add	call	go	later	left	reply	right	stop	subtract	you
add	22	1	0	1	0	0	0	0	0	1
call	3	17	0	1	0	0	2	0	0	0
go	1	0	11	5	3	0	1	1	0	0
later	2	0	3	16	2	0	0	0	0	0
left	0	1	0	4	13	2	3	0	0	0
reply	0	1	0	1	6	13	2	0	0	0
right	2	0	1	0	3	4	13	0	0	0
stop	1	0	0	0	1	0	1	18	0	2
subtract	0	0	0	0	0	0	0	4	19	1
you	0	1	1	0	0	0	0	1	3	19

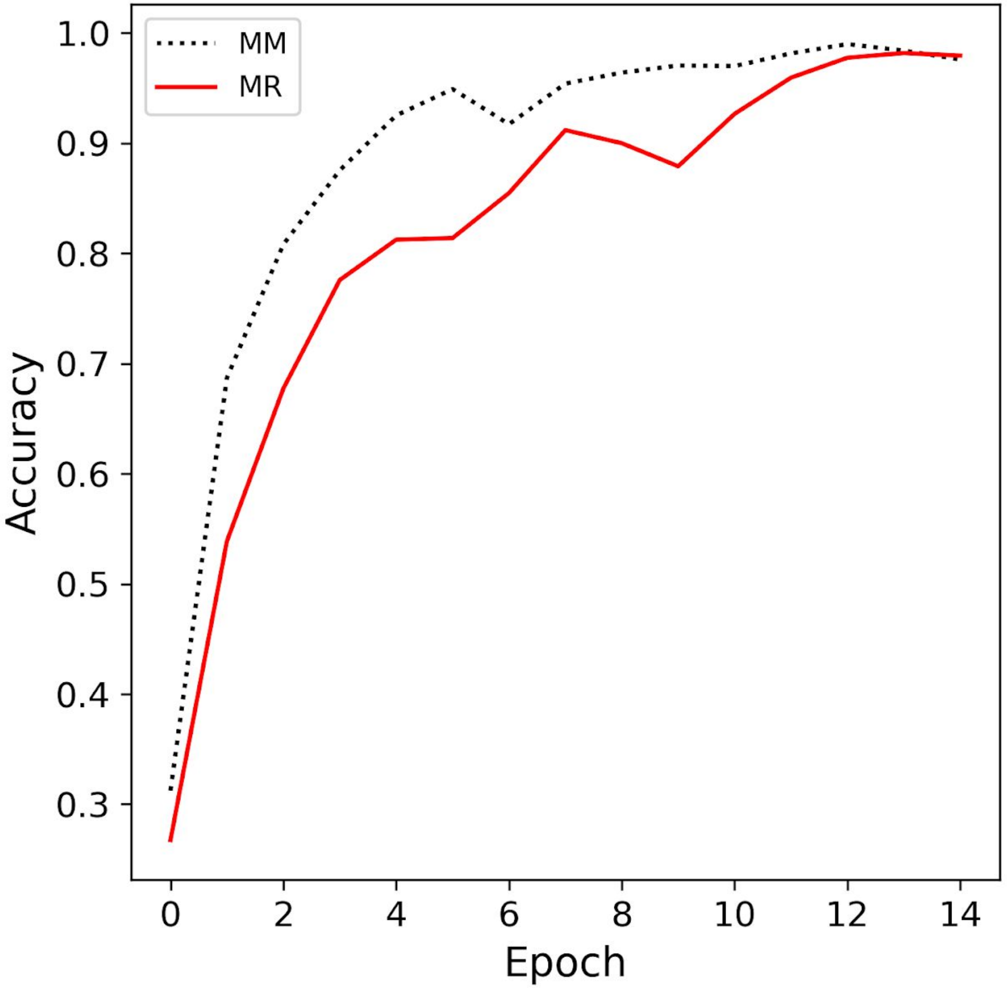
Mentally Rehearsed (MR)

Actual Label \ Predicted Label	add	call	go	later	left	reply	right	stop	subtract	you
add	20	1	0	0	0	0	0	0	0	1
call	3	15	1	1	0	0	0	0	0	1
go	1	1	11	4	0	1	1	1	0	0
later	1	1	0	17	2	2	1	1	1	0
left	1	0	0	0	17	0	2	1	0	0
reply	0	0	1	1	1	16	0	0	0	1
right	0	0	0	2	5	1	11	2	0	0
stop	0	0	4	0	0	0	1	18	1	1
subtract	0	0	0	2	0	0	2	2	15	1
you	0	0	0	0	0	0	0	0	3	22

Muscle Movement (MM)

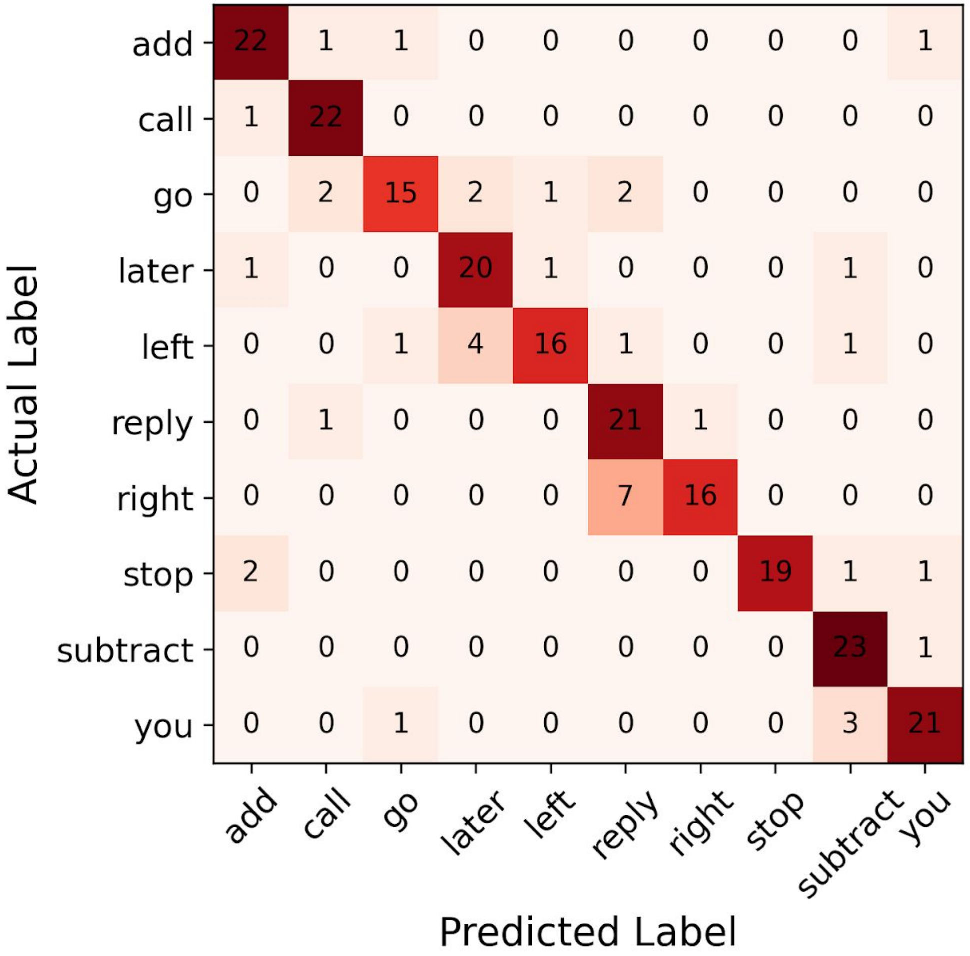
Results - 5
[CNN Accuracy and Loss]



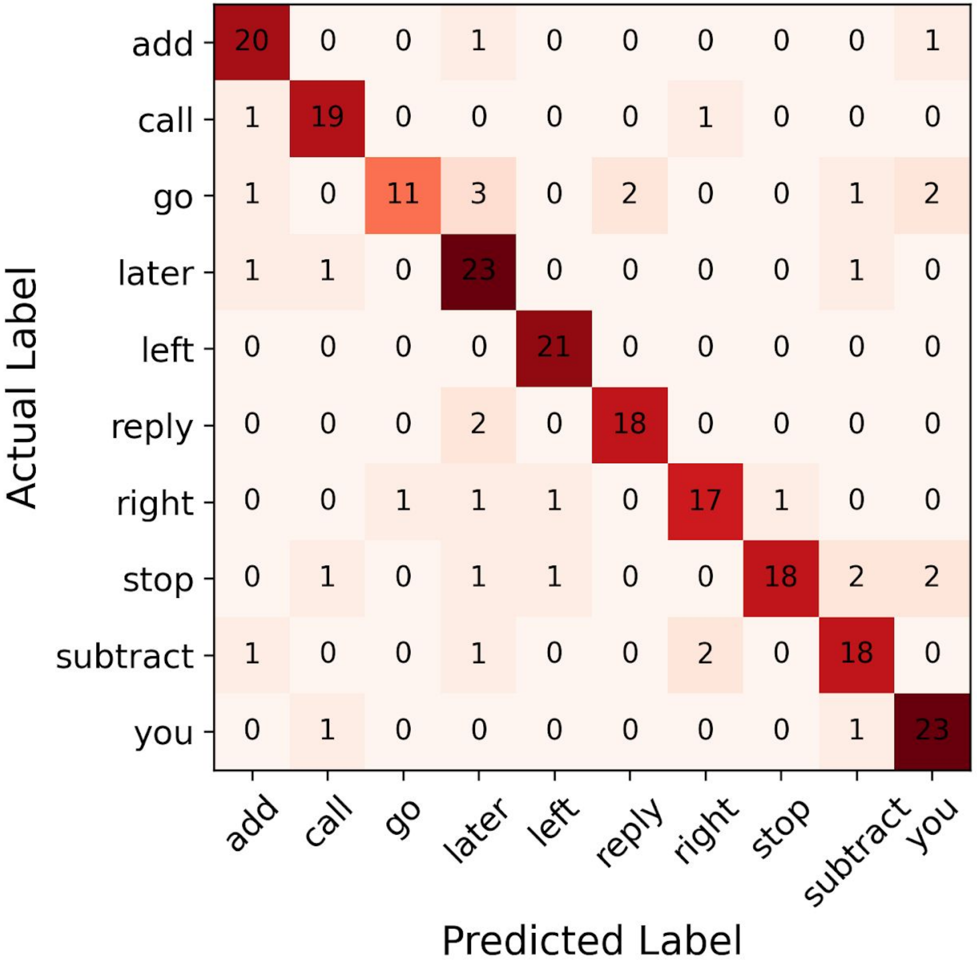


Modes	**Accuracy**		**Loss**	
	Train	**Validation**	**Train**	**Validation**
Muscle Movement (MM)	0.9761	0.8430	0.0819	1.1073
Mentally Rehearsed (MR)	0.9795	0.8333	0.0677	0.8627

Results - 6
[Confusion Matrix from CNN]



Mentally Rehearsed (MR)



Muscle Movement (MM)

Results - 7
[CNN Prediction on Terminal]

```
ACCURACY:  0.8470255136489868
LOSS:  0.6574686169624329


Actual Labels    Predicted Lables
   call          left
   go          go
   later          go
   stop          stop
   stop          stop
   later          later
   go          go
   add          add
   left          left
   you          you
   right          right
   stop          stop
   stop          right
   reply          left
   reply          reply
   right          right
   you          you
   later          later
   later          stop
```

Tested on speaker "US" data

```
ACCURACY:  0.09078404307365417
LOSS:  8.198955535888672


Actual Labels    Predicted Lables
   right          right
   stop          right
   call          subtract
   left          call
   later          reply
   left          left
   stop          right
   call          subtract
   you          call
   left          right
   right          reply
   right          left
   right          subtract
   right          left
   stop          you
   left          call
   you          call
   reply          left
   call          add
```

Tested on speaker "RL" data

Results - 8
[Model Summary]

Model	Mode	Feature Type	Accuracy	Validation Accuracy	Precision	Recall	F1 Score
MLP	MM	Temporal	81.32	75.34	0.75	0.75	0.75
		Spectral	60.74	49.33	0.49	0.49	0.47
		Temporal and Spectral	80.32	72.65	0.73	0.72	0.72
	MR	Temporal	68.92	67.95	0.69	0.68	0.67
		Spectral	60.68	55.50	0.63	0.60	0.60
		Temporal and Spectral	74.49	68.80	0.70	0.68	0.69
CNN	MM	Temporal	98.60	84.30	0.85	0.84	0.84
		Spectral	85.30	70.40	0.80	0.78	0.78
		Temporal and Spectral	97.61	84.30	0.85	0.84	0.84
	MR	Temporal	98.10	84.62	0.85	0.85	0.85
		Spectral	77.30	68.38	0.71	0.68	0.68
		Temporal and Spectral	97.95	83.33	0.85	0.83	0.83

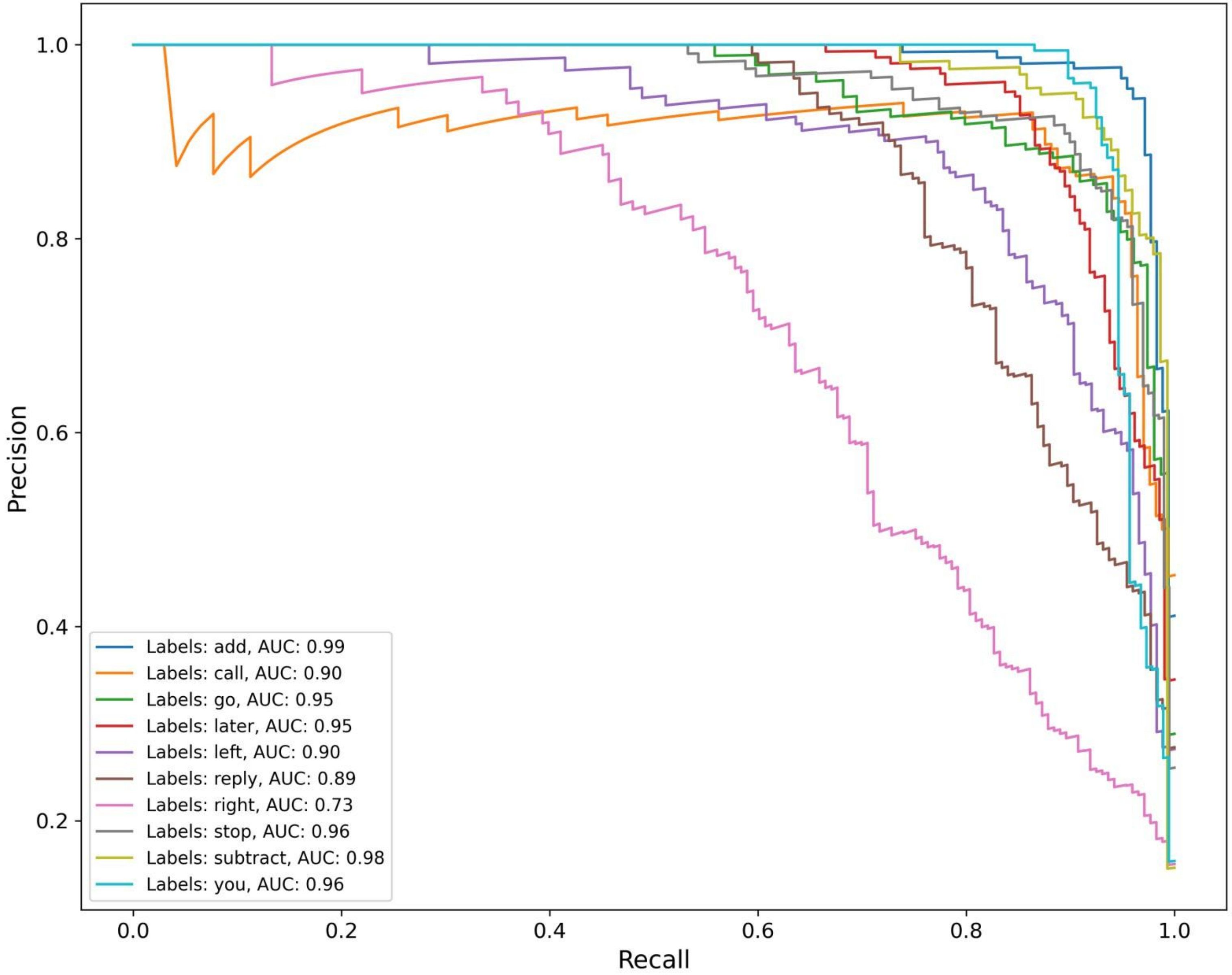
Analysis and Discussion - 1

- Line noise encountered at 50 Hz and its harmonics
- Before OpenBCI hardware was obtained, self-designed dual channel hardware was used
    - Difficulty in dataset recording due to less channels
    - Lower performance due to unoptimized design
- Bluetooth interference observed between nearby OpenBCI devices while recording dataset
- ECG artifacts encountered in the signal

Analysis and Discussion - 2

- Performance of Mentally Rehearsed (MR) mode is comparable to that of Mouth Movement (MM) mode due to more data instances
- Low accuracy in spectral features due to use of STFT magnitude instead of spectrogram as features
- Prediction accuracy in terminal is:
    - Higher for speaker “US” due to fewer but longer sessions while recording data
    - Lower for speaker “RL” due to multiple short sessions while recording data

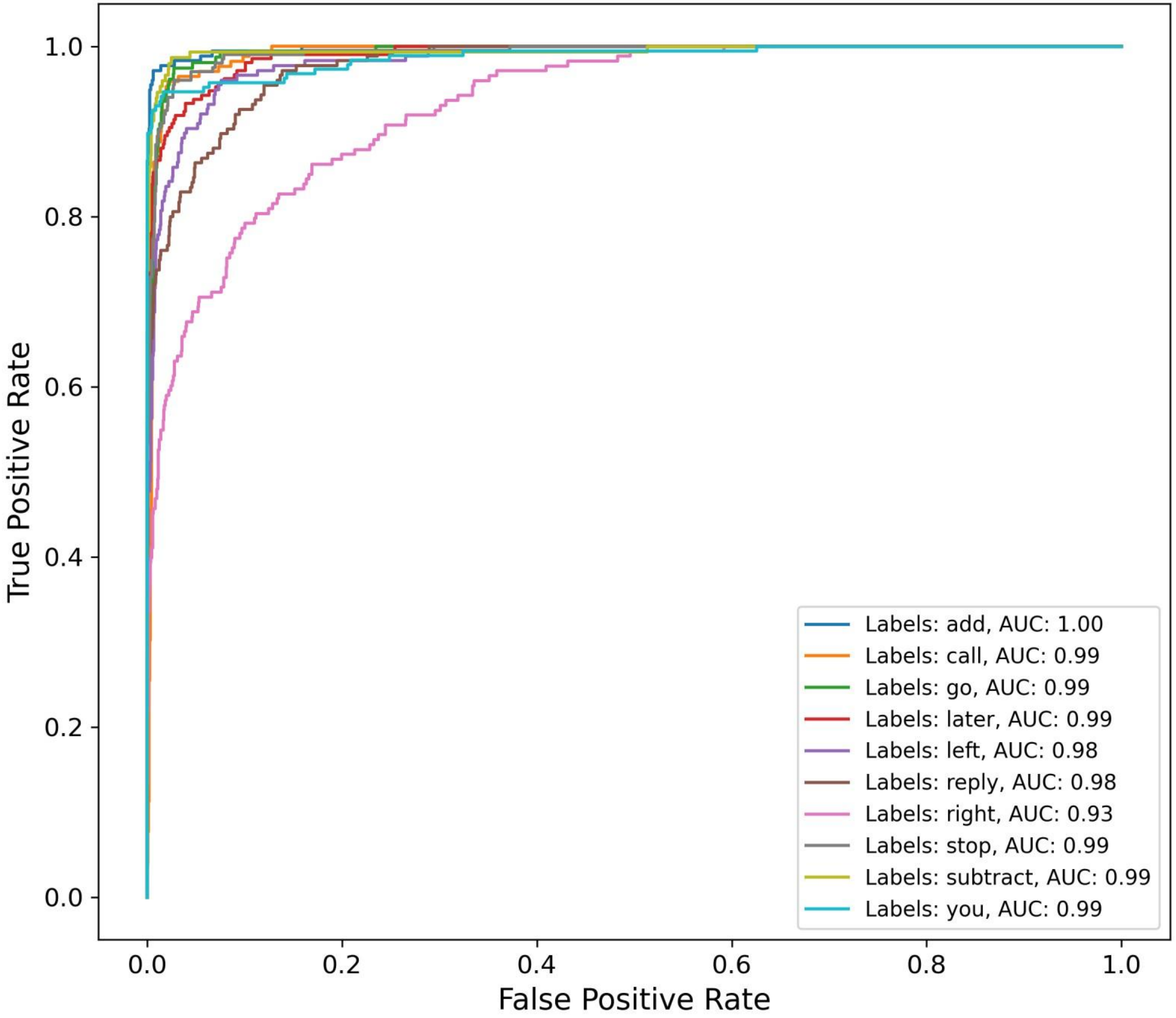
Analysis and Discussion - 3



Highest AUC Score: 0.99 (ADD)

Lowest AUC Score: 0.73 (RIGHT)

Analysis and Discussion - 4



Highest AUC Score: 1 (ADD)

Lowest AUC Score: 0.93 (RIGHT)

Future Enhancements

- Increase volume of the dataset by involving speakers of diverse age groups of both genders
- Optimize the number of EMG channels by avoiding muscles within close proximity of each other
- Incorporating additional features compatible with neuromuscular signal analysis
- Develop continuous and multilingual speech recognition system

Conclusion

- EMG signals were extracted successfully and transmitted to a remote computer
- Data was analyzed and processed for feature extraction and implemented on machine learning models
- The concept of human-computer interaction using neuromuscular signals was shown to be feasible within a permissible error rate
- Prediction of words was found to be dependent on speakers as well as sessions

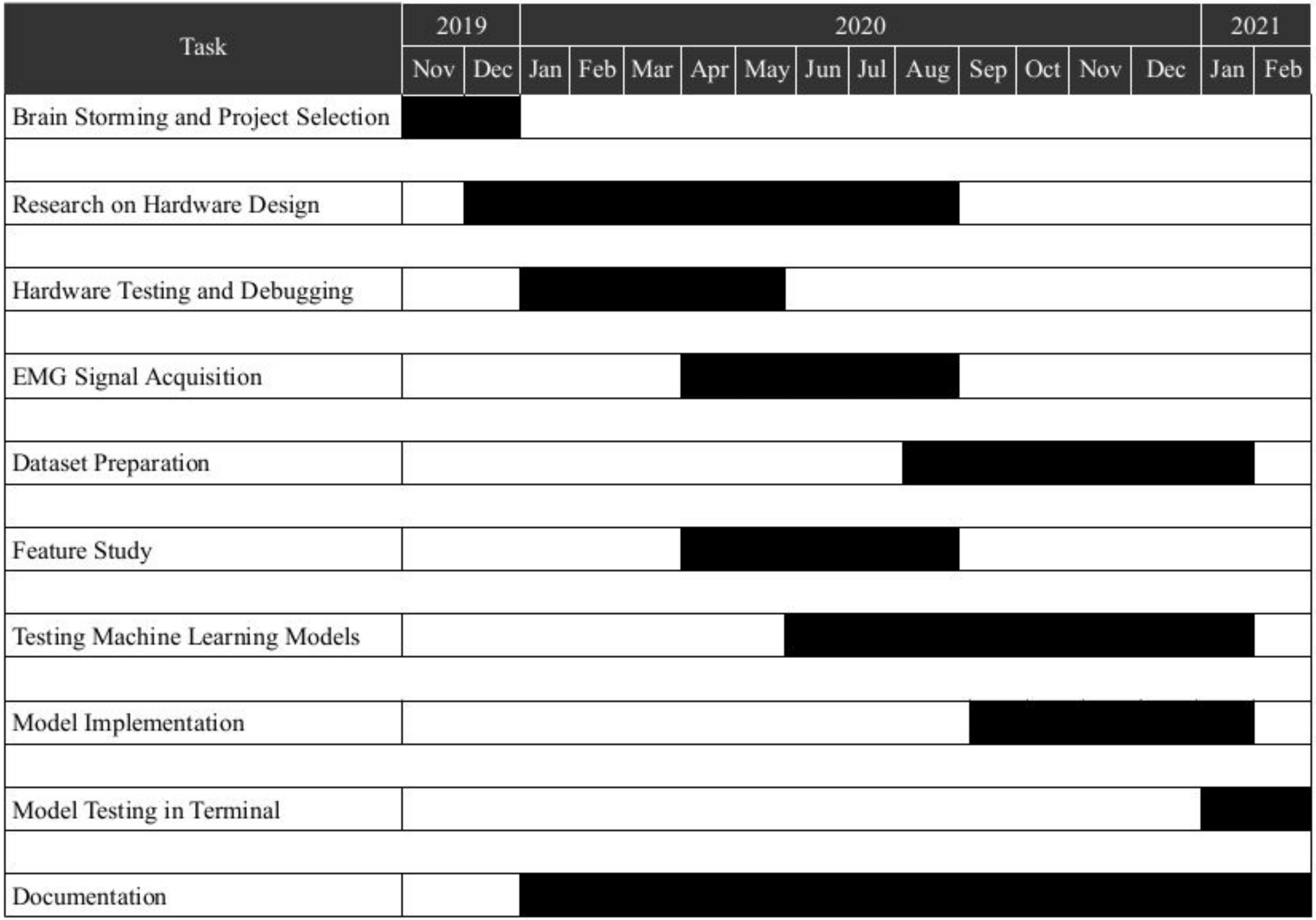
Appendix - 1
[Project Budget]

S.N.	Title	Qty(pcs)	Rate(NRs.)	Price (NRs.)	Remarks
1	Ag-AgCl Electrode (ECG EMG)	150	10/-	1,500/-	-
2	Instrumentation Amplifier (AD620)	5	114/-	570/-	-
3	Op-amp (OP37AJ)	10	114/-	1,140/-	-
4	Passive Electronic Components: • Resistors • Capacitors • Header pins • Diodes	-	-	2,000/-	-
5	AgCl Electrolyte	250ml	50/-	50/-	-
6	Arduino Uno	2	1,000/-	2,000/-	-
7	Single Sided PCB Board	2	250/-	500/-	-
8	Shielded RCA Cable	8 (1m)	215/-	1,720/-	-
9	Cyton Board	2	46,532/-	93,064/-	Donated
10	USB Dongle	2	11,633/-	23,266/-	Donated
11	Gold Cup Electrodes	2 set	3,490/-	6,980/-	Donated
12	Ten 20 Electrolyte	6	2,326/-	13,956	Donated
13	Miscellaneous	-	-	3,000/-	-
	Total			**1,60,212/-**	

Appendix - 2
[Project Timeline]

Project Start Date: 15 November 2019

Project End Date: 25 February 2021



References

[1] A. Kapur, "Human-Machine Cognitive Coalescence through an internal duplex interface," MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2018.

[2] Arnav Kapur,Shreys Kapur, Pattie Maes, "AlterEgo," Multimodel Interface, 2018.

[3] Michael Wand and Tanja Schultz, "SESSION-INDEPENDENT EMG-BASED SPEECH RECOGNITION," *Cognitive Systems Lab, Karlsruhe Institute of Technology, Adenauerring 4, 76131 Karlsruhe, Germany,* pp. 1-3.

References

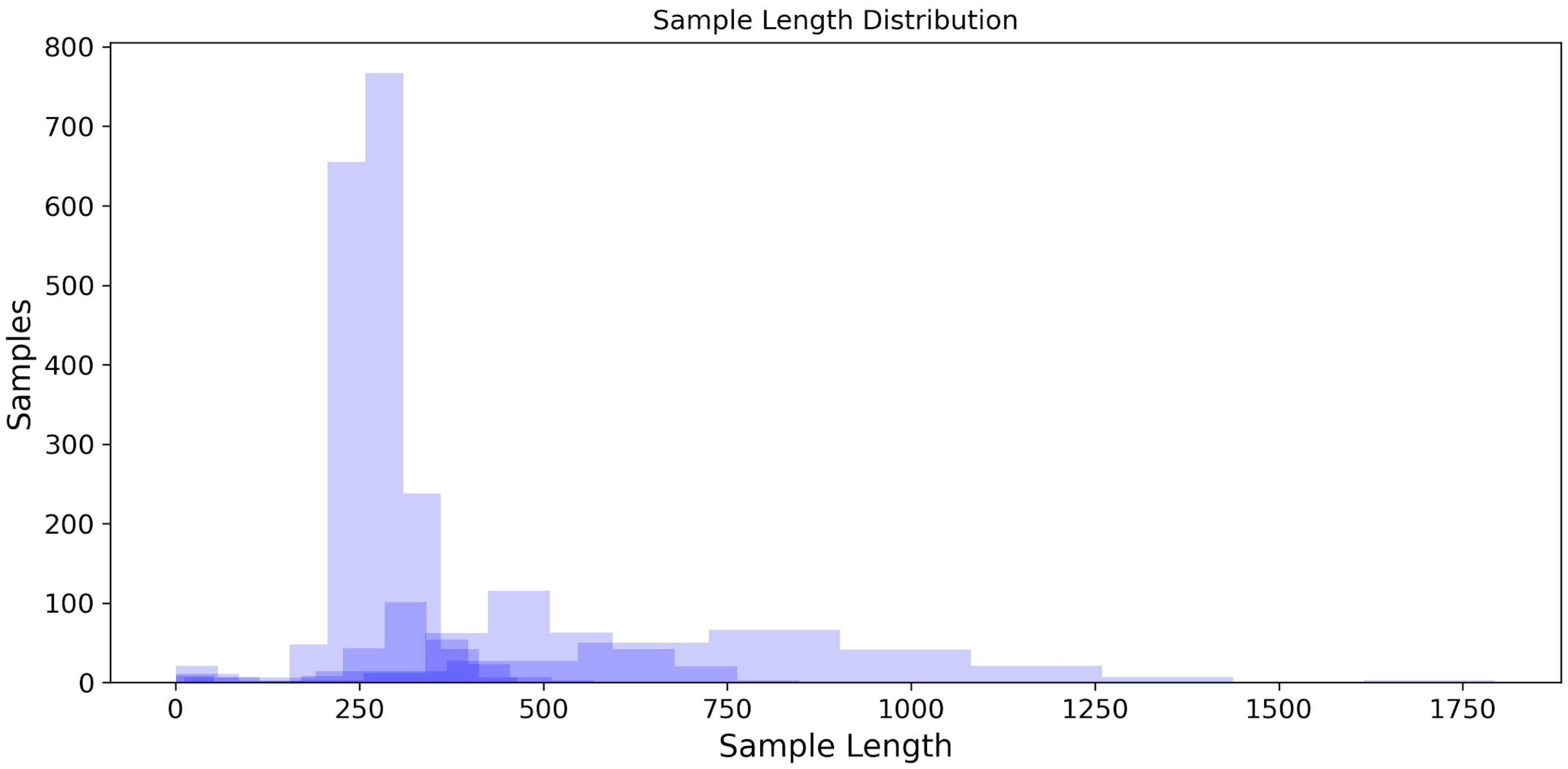
[4] G. Kamen, David A. Gabriel, "Essentials of Electromyography," in Human Kinetics, 2010, p. 57.

[5] Chuck Jorgensen and Kim Binsted, "Web Browser Control Using EMG Based Sub Vocal Speech Recognition," in 38th Hawaii International Conference on System Sciences, IEEE, Hawaii, 2005.

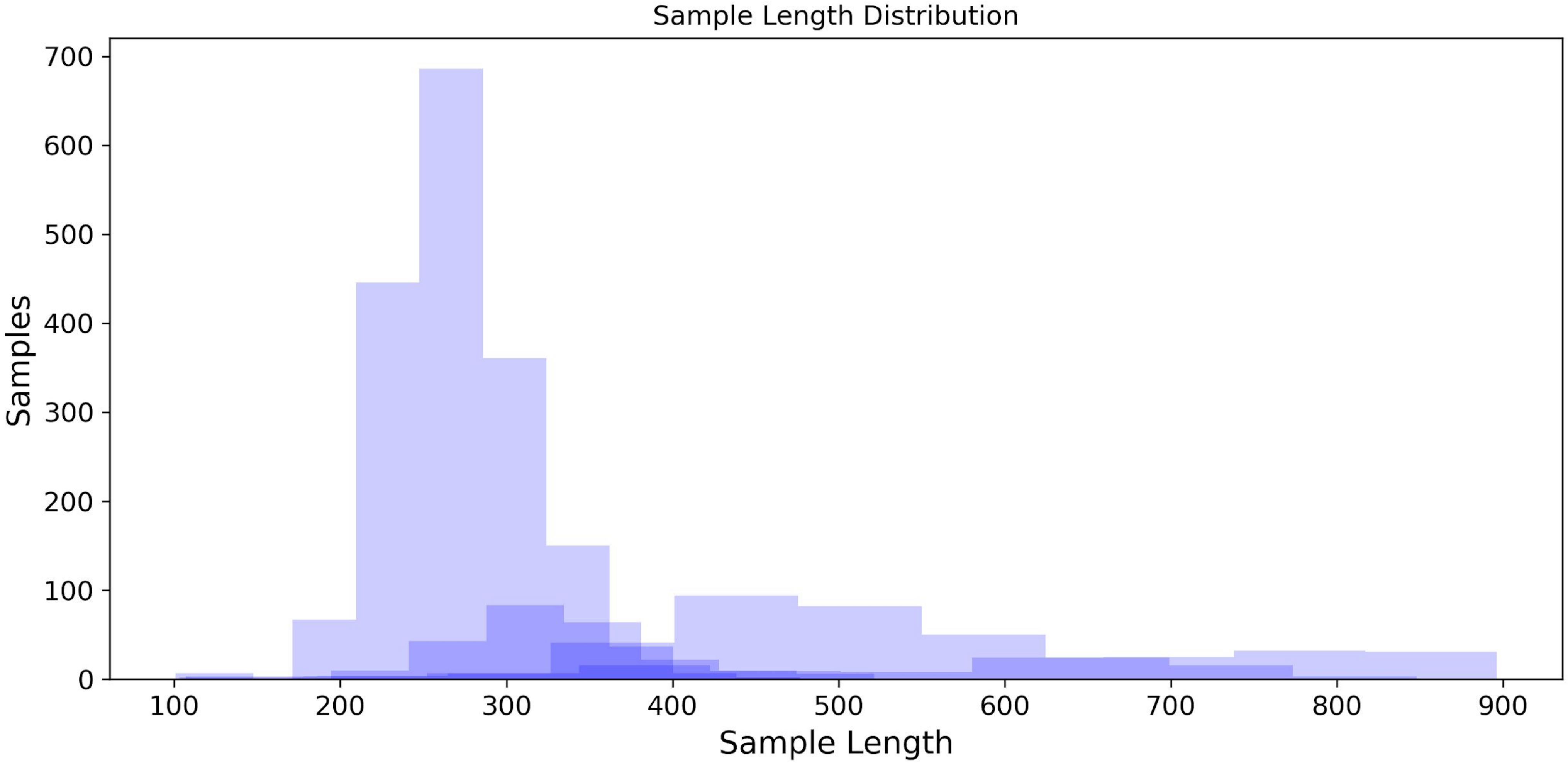
[6] Jacob MIllman and Christos C. Halkias, Integrated Electronics: Analog and Digital Circuits and systems, McGraw-Hill Kogakusha. Ltd., 1972, pp. 501-534

Thank You!!

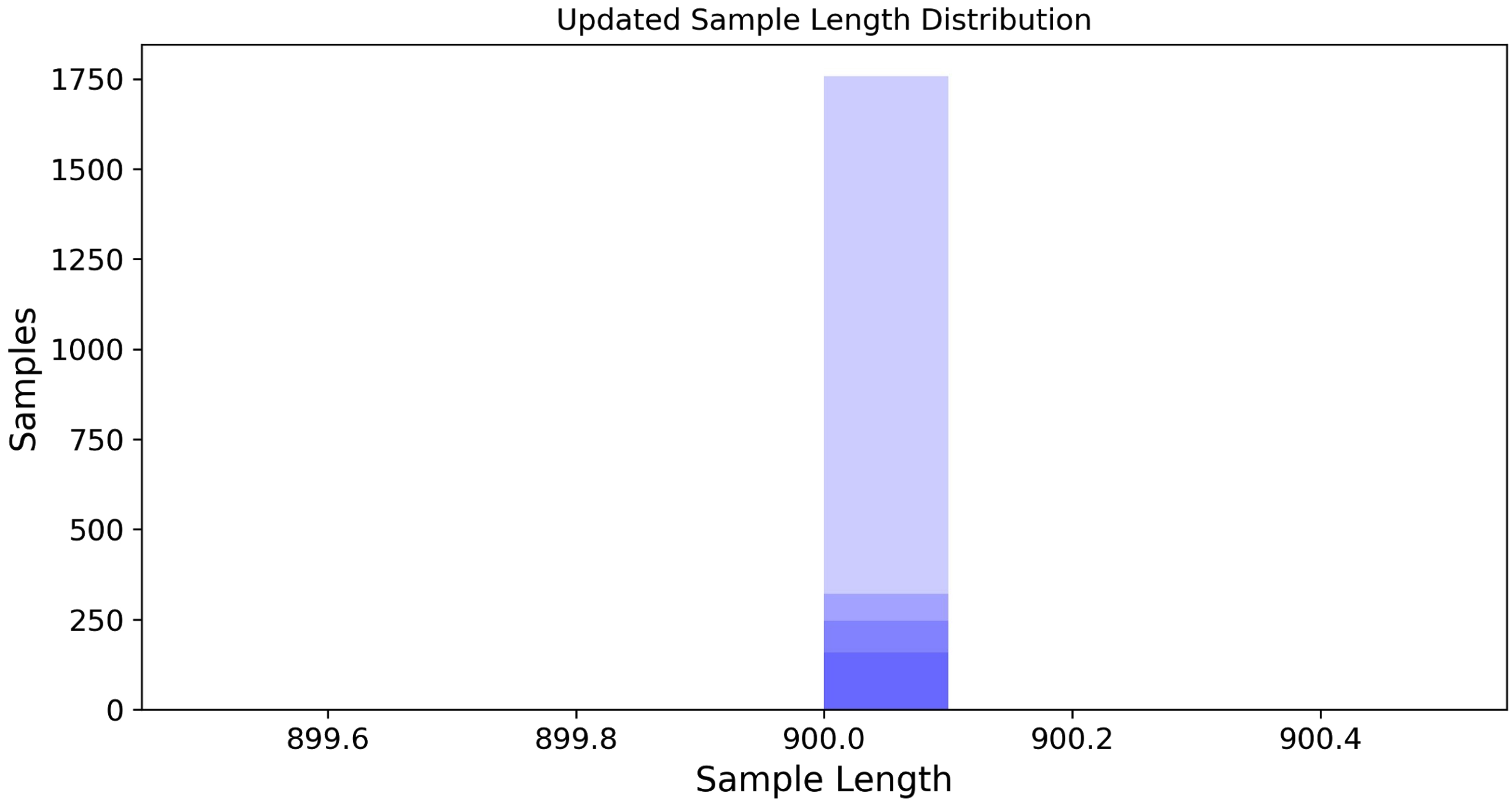
[Sample Length Distribution (Raw)]



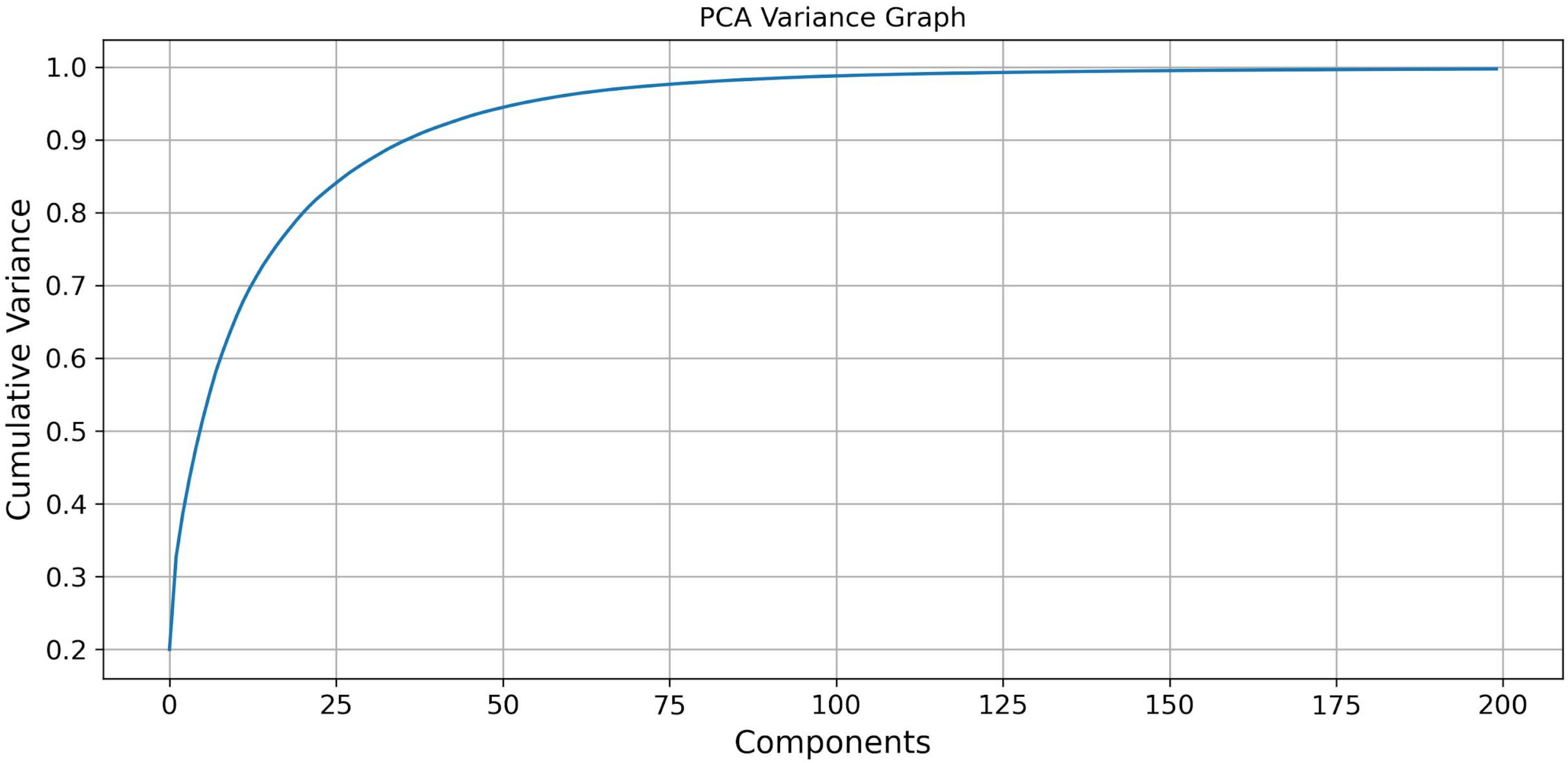
[Sample Length Distribution (Unnormalized)]



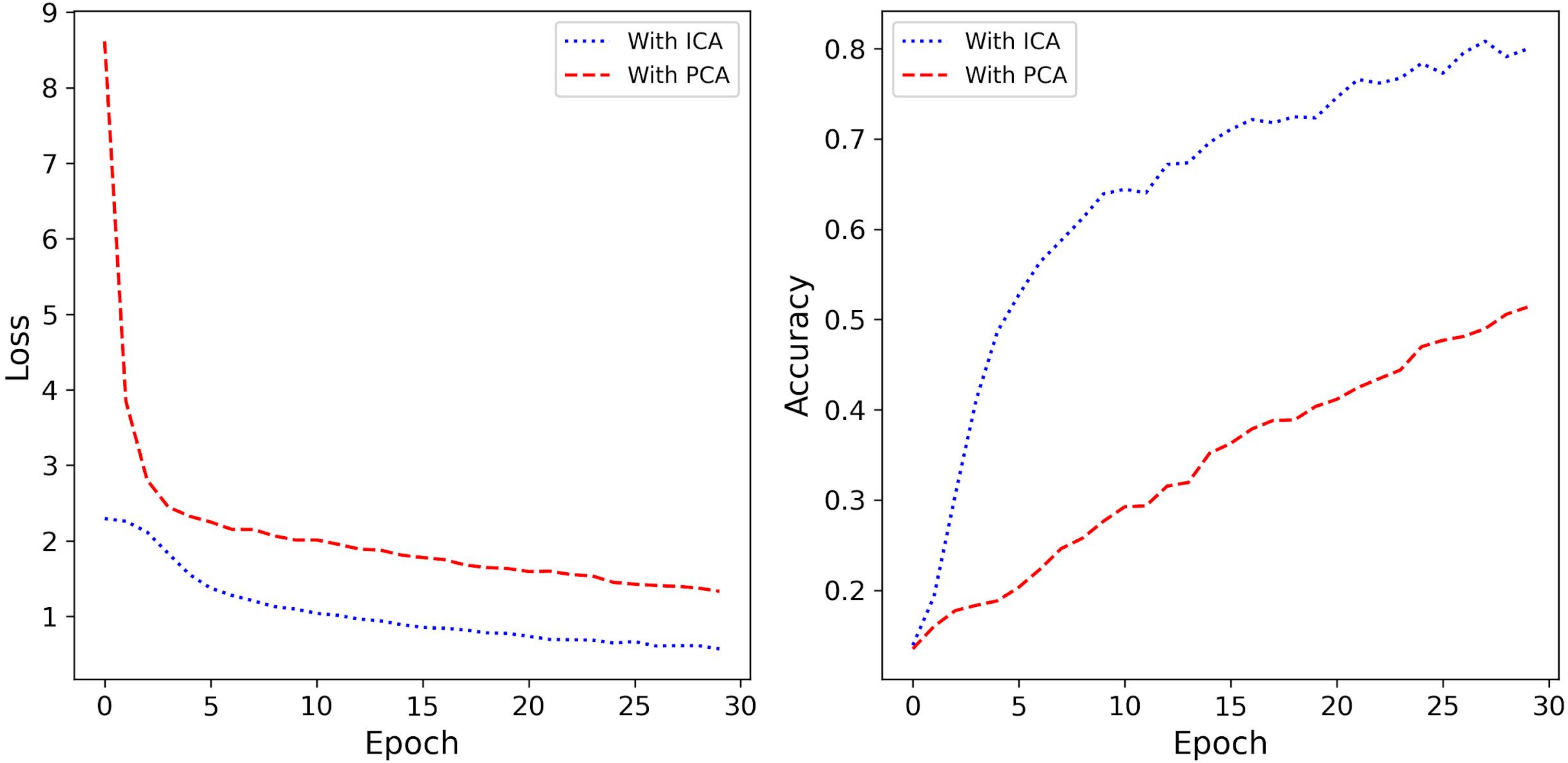
[Sample Length Distribution (Normalized)]



[PCA Variance Graph]

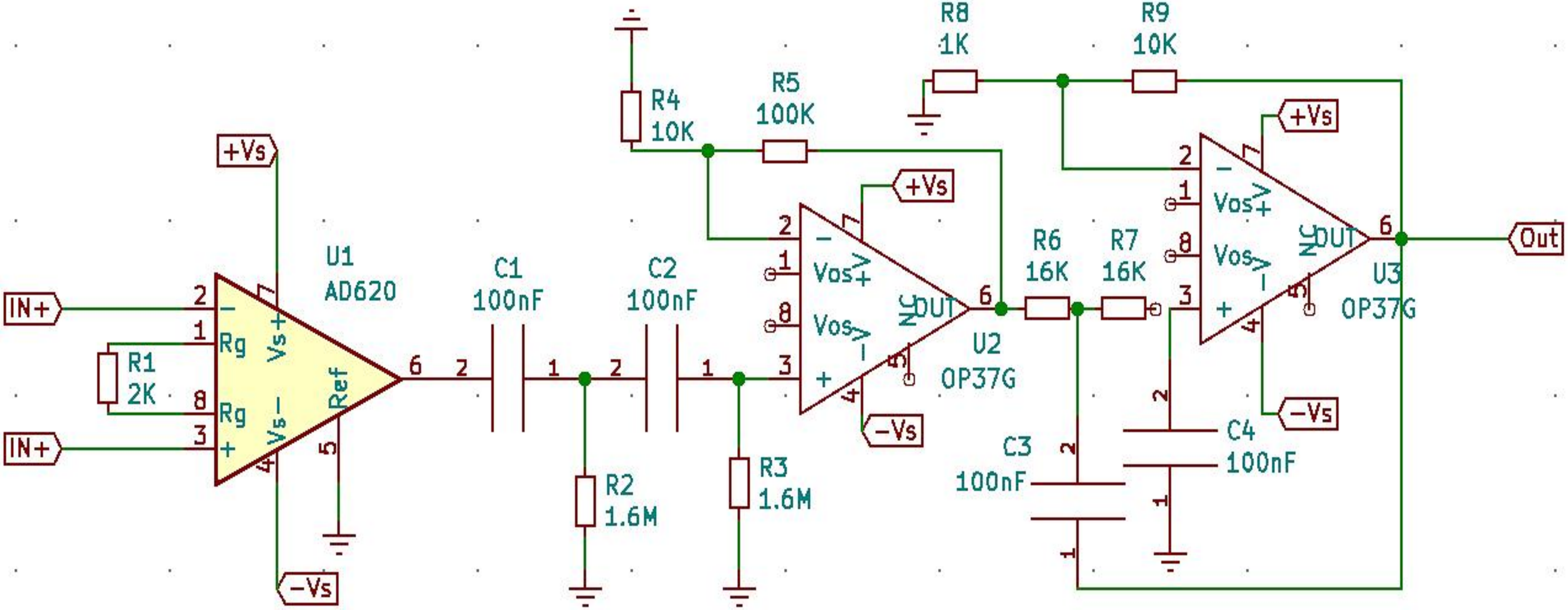


[PCA vs ICA]



With ICA
With PCA
Loss
Epoch
0
5
10
15
20
25
30
1
2
3
4
5
6
7
8
9
With ICA
With PCA
Accuracy
Epoch
0.2
0.3
0.4
0.5
0.6
0.7
0.8

[Single Channel Schematic]



[Multi-channel PCB Design]