# Silent Speech Recognition in Nepali

Rabin Nepal, Rhimesh Lwagun, Sanjay Rijal, Upendra Subedi, Dinesh Baniya Kshatri
Department of Electronics and Computer Engineering
Thapathali Campus, Institute of Engineering, Tribhuvan University, Nepal

## INTRODUCTION

Speech is a convenient way to interact with smart electronic gadgets. However, normal audible speech is predisposed to noisy external environments and subjected to privacy issues. This research work breaks down the human speech process and exploits the bio-signals generated during internal articulation, which is a part of the speech generation process. Sentences uttered silently by a user are imperceptibly sent to a remote device for seamless human computer interaction.

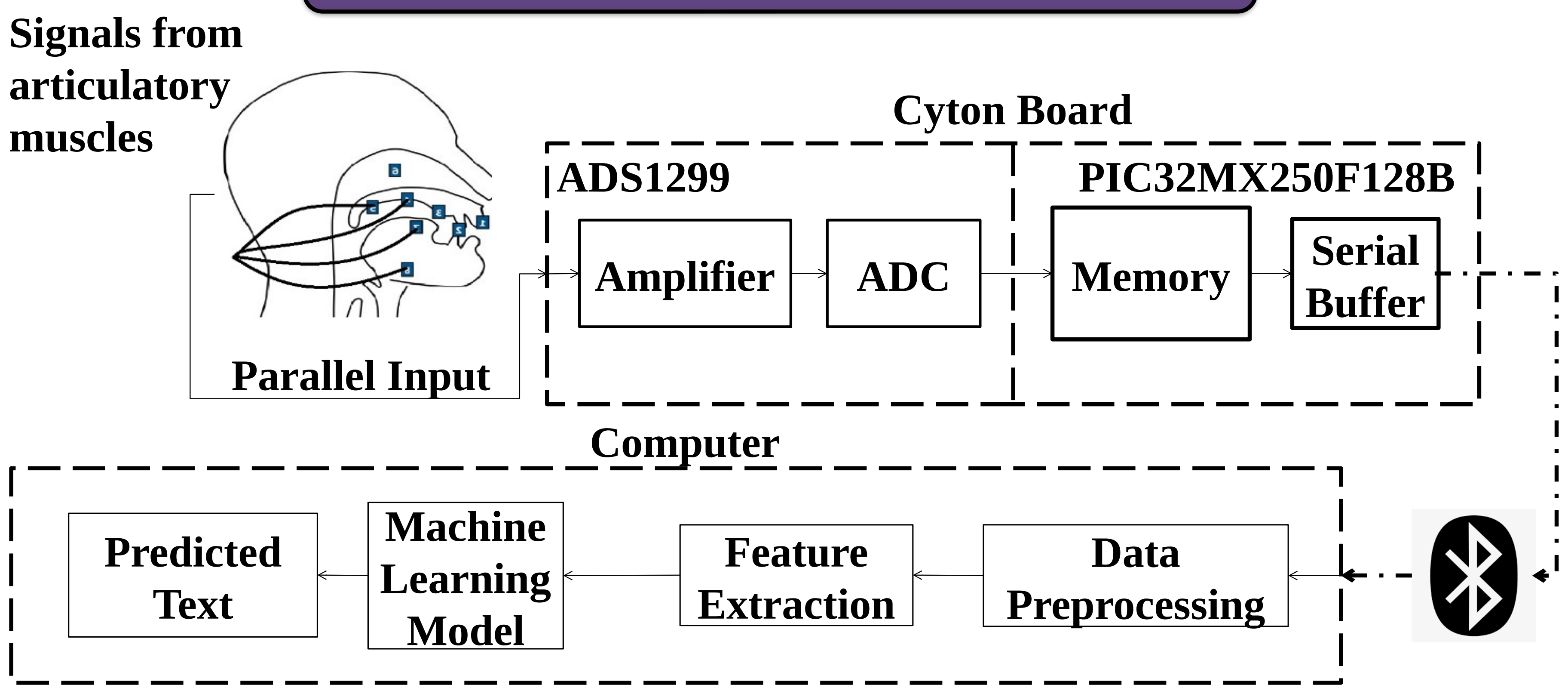
## OBJECTIVE

To process and analyze bio-signals from speech articulator muscles for recognizing silently uttered Nepali sentences

## METHODOLOGY

- Selected muscles around the face and the neck region formed recording sites for bio-signals
- The sites were gel coated to reduce impedance between skin and electrode
- Signals recorded from articulator network were normalized and filtered to remove line noise and ECG artifacts
- Short-time Fourier transform provided the optimum features required to train a convolutional neural network

## SYSTEM BLOCK DIAGRAM



Signals from articulatory muscles

Parallel Input

Cyton Board

ADS1299

Amplifier

ADC

PIC32MX250F128B

Memory

Serial Buffer

Computer

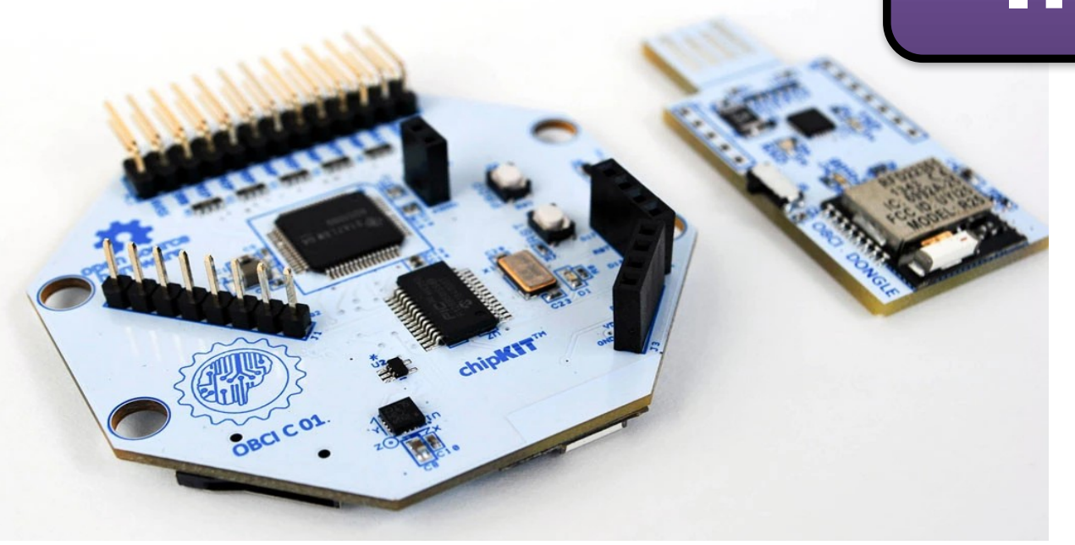
Data Preprocessing

Feature Extraction

Machine Learning Model

Predicted Text

## INSTRUMENTATION

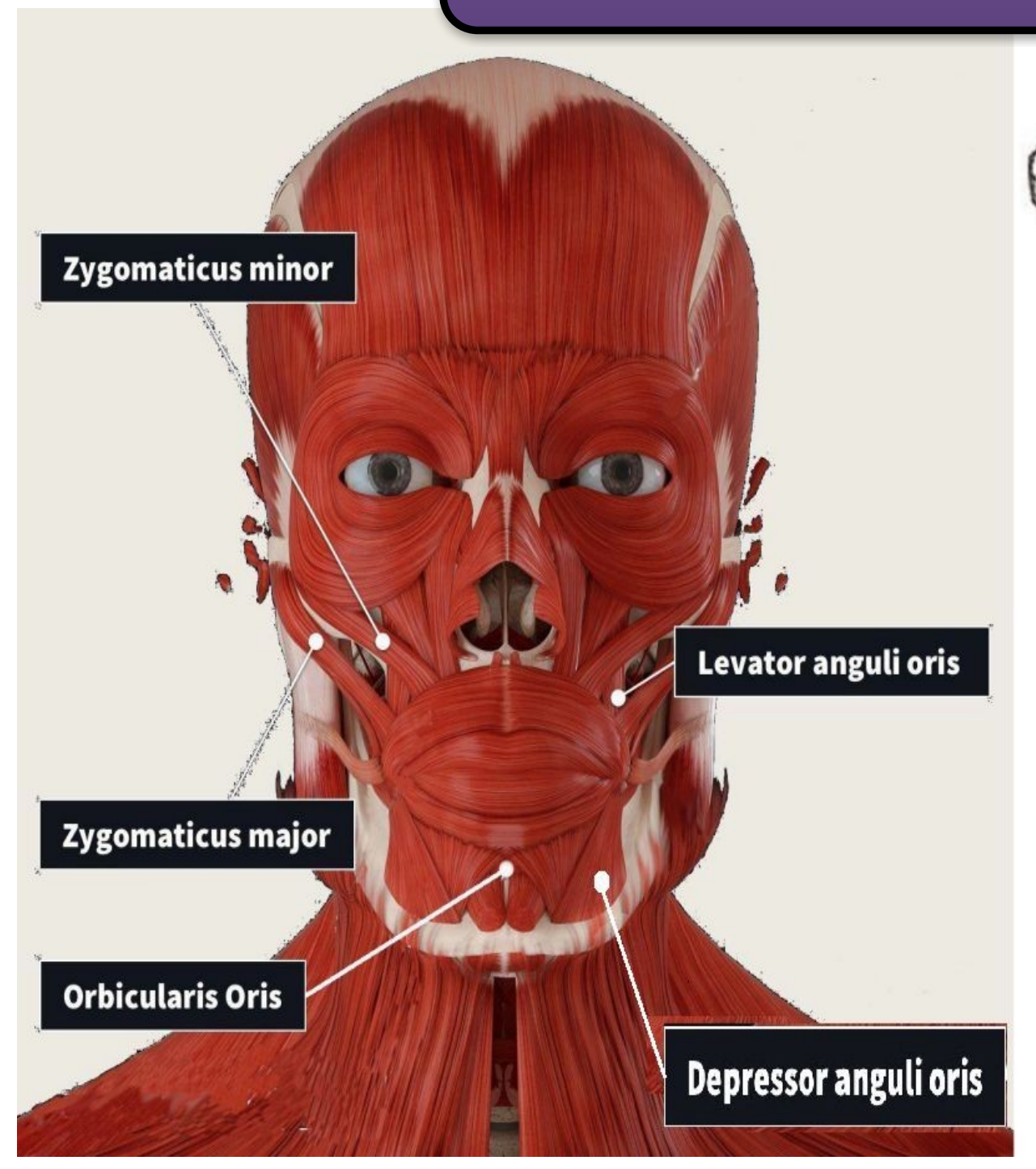


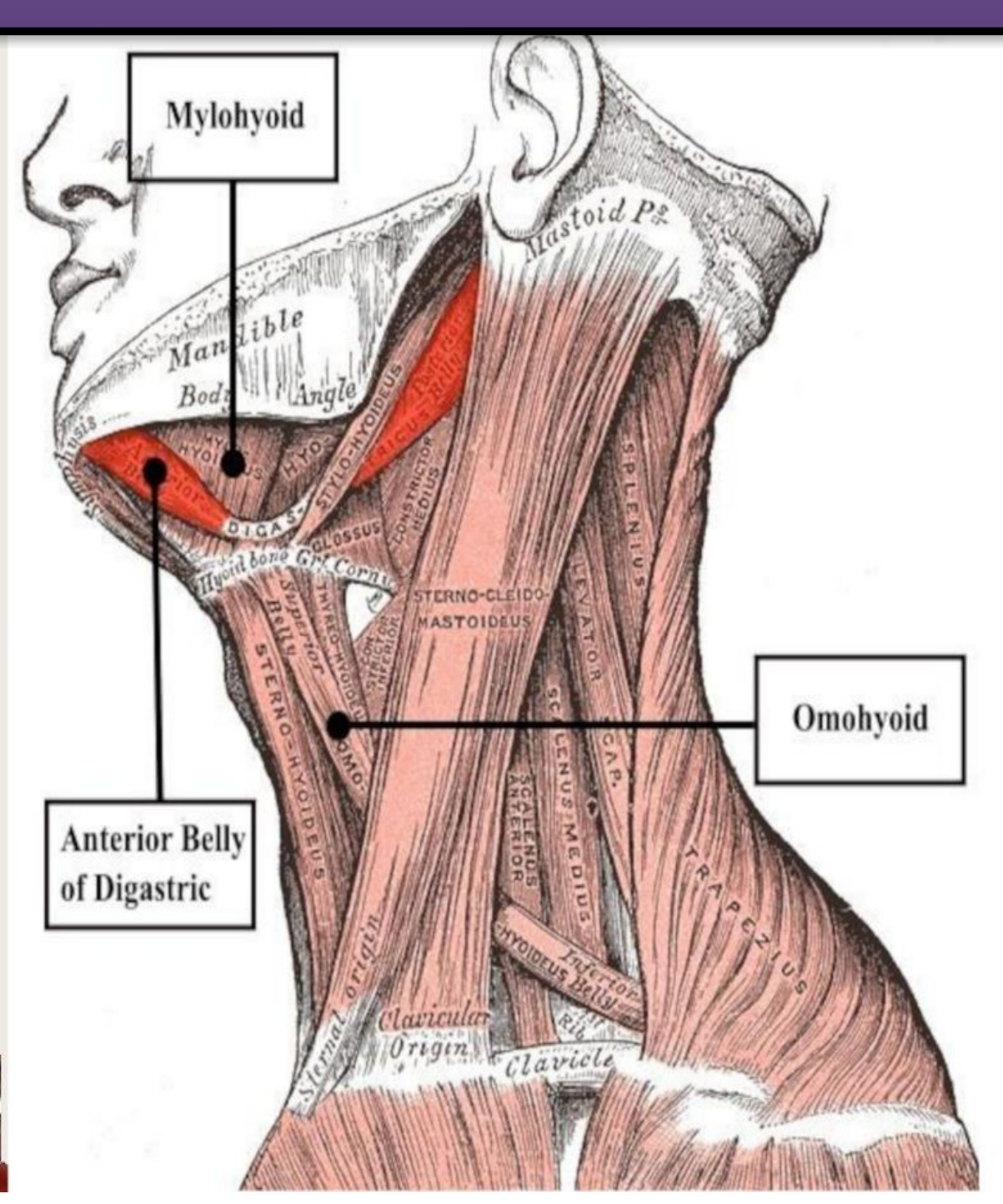
OpenBCI Cyton Board

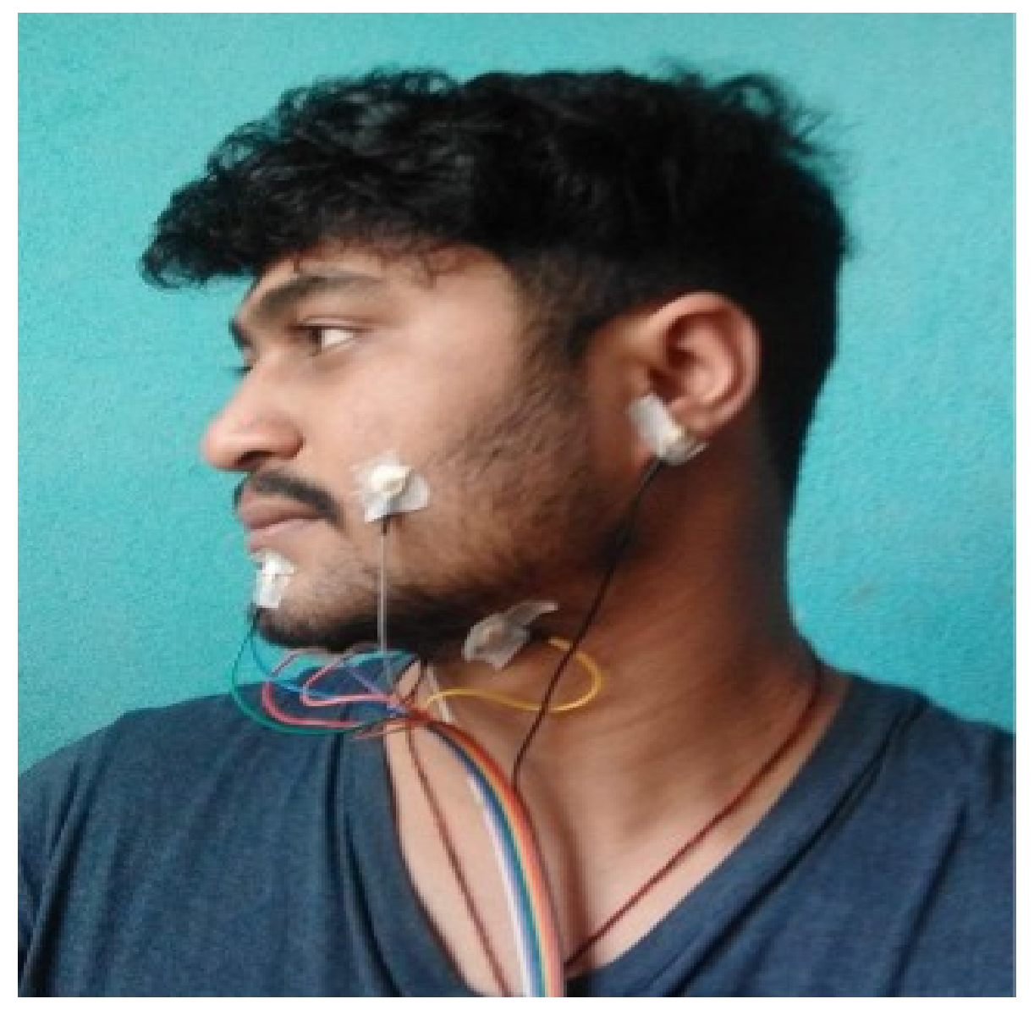
Goldcup EMG Electrodes

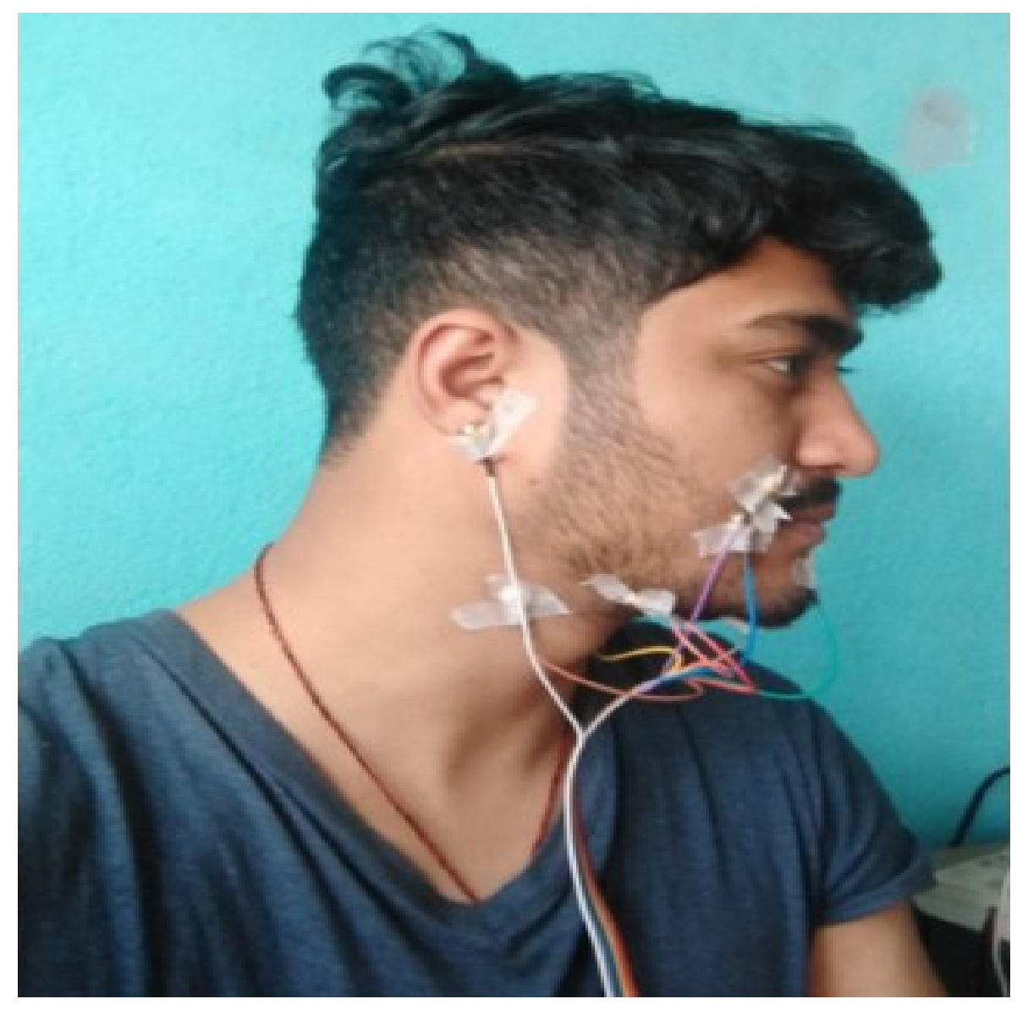
Electrolyte Ten20 Paste

## SPEECH ARTICULATOR MUSCLES



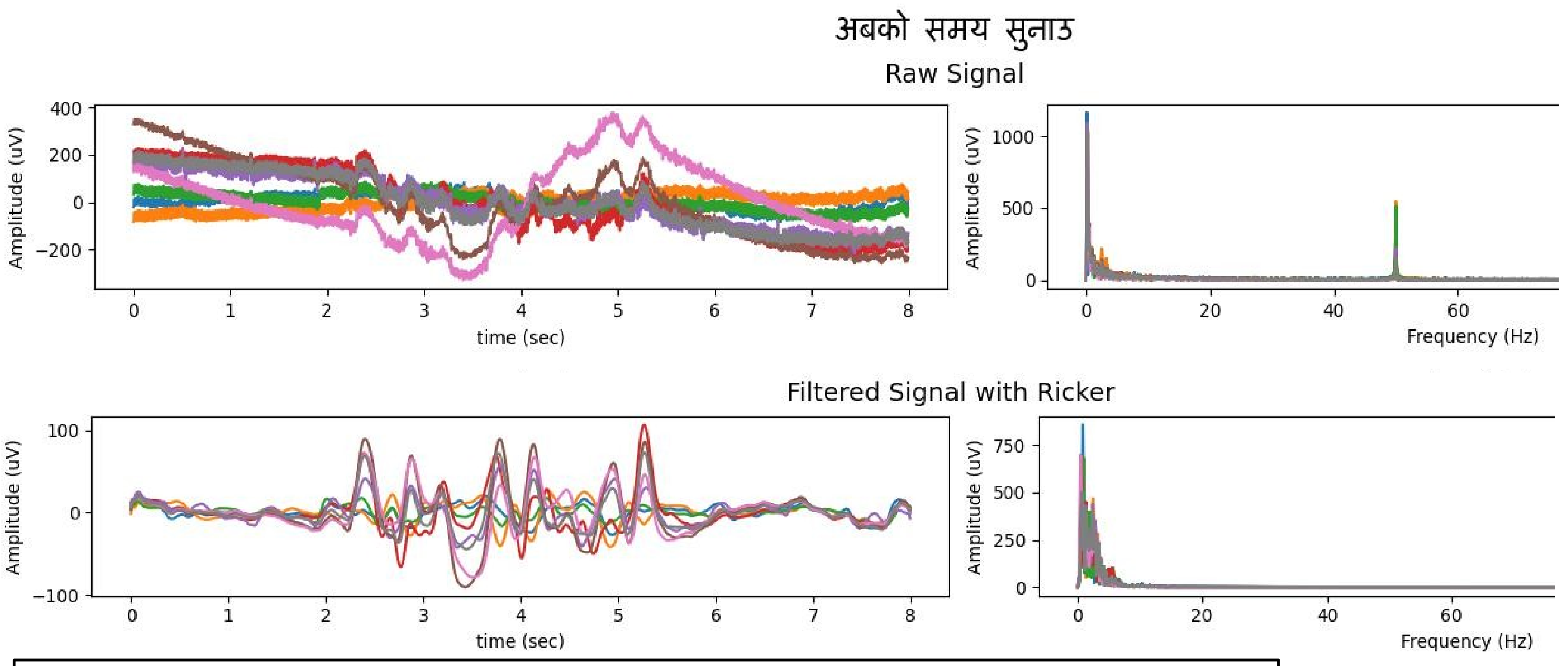


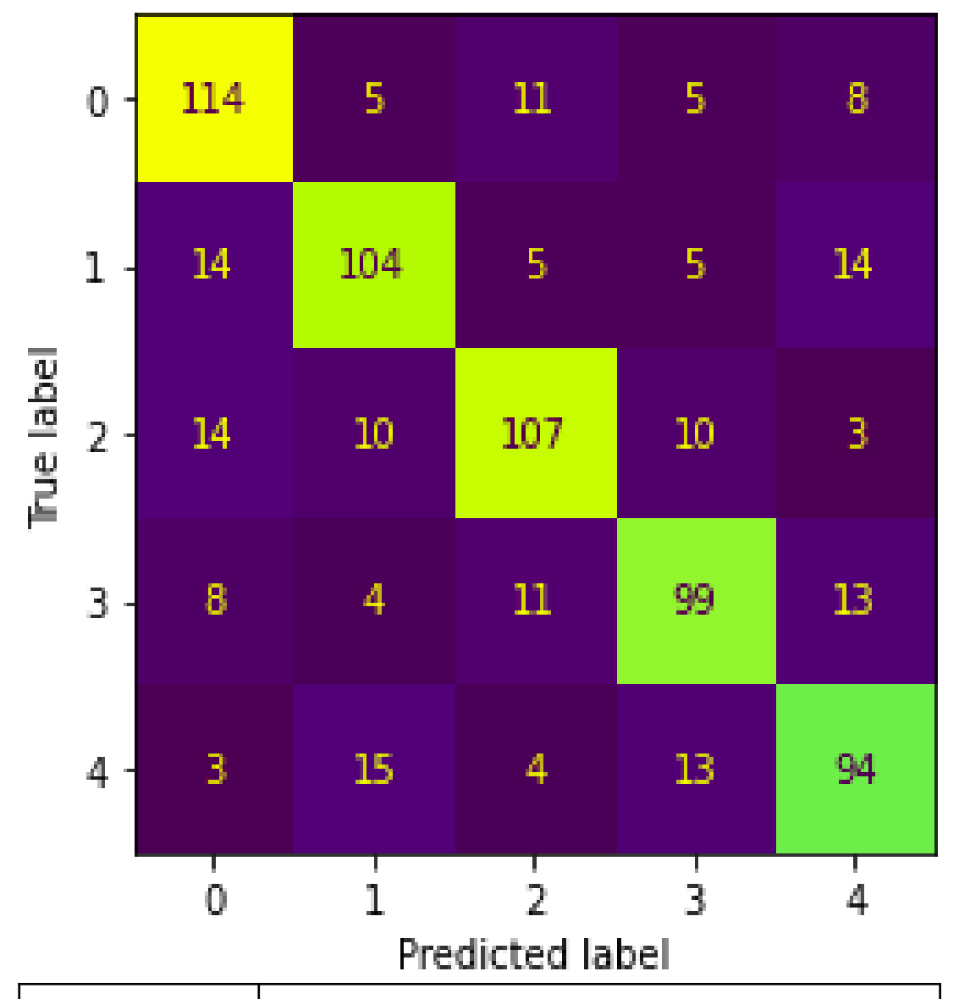




| EMG Channel | Muscle Name |
|---|---|
| 1. | Levator Angulis Oris |
| 2. | Zygomaticus Minor |
| 3. | Zygomaticus Major |
| 4. | Orbicularis Oris |
| 5. | Omohyoid |
| 6. | Anterior Belly of Digastric |
| 7. | Mylohyoid |
| 8. | Depressor Anguli Oris |

## FINDINGS AND ANALYSIS

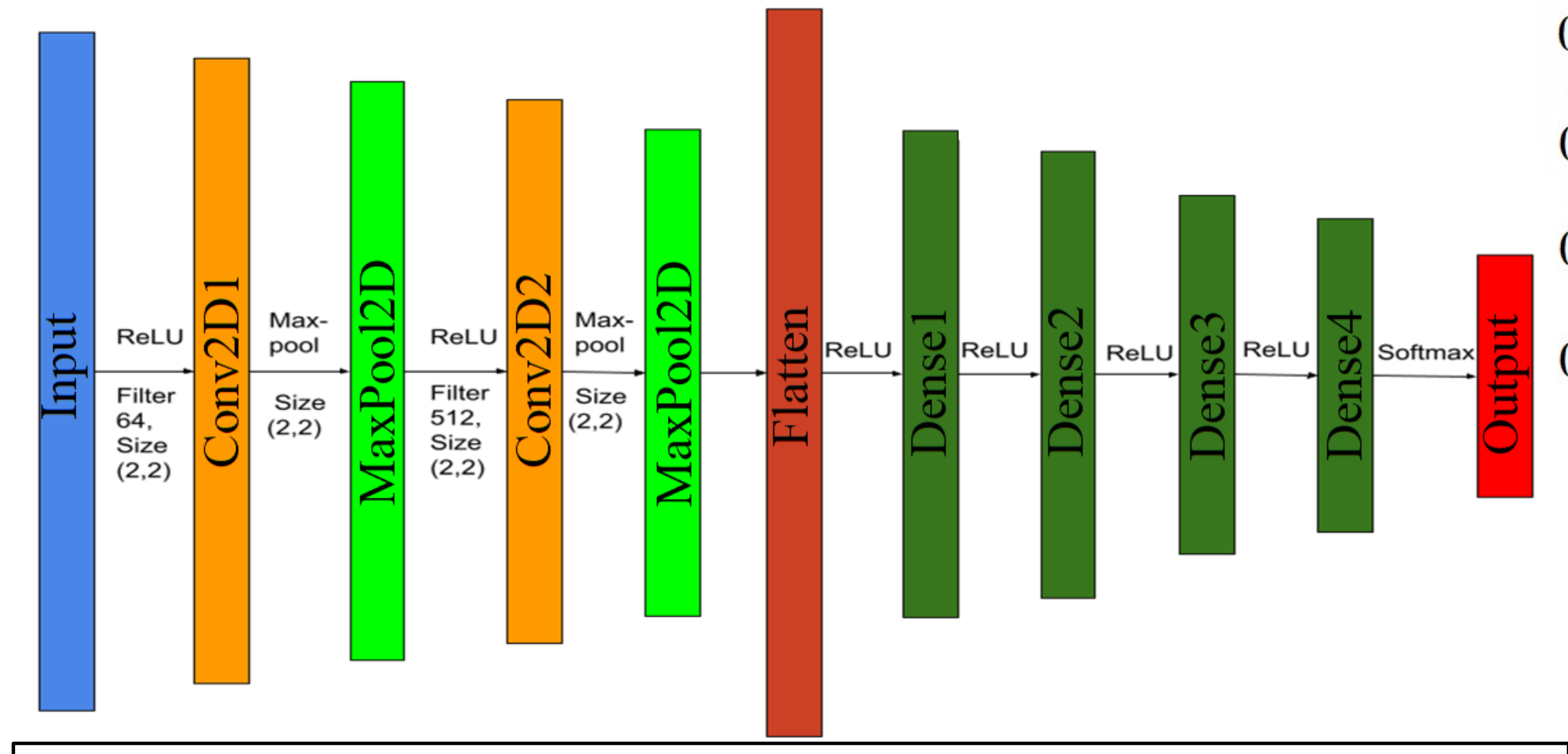


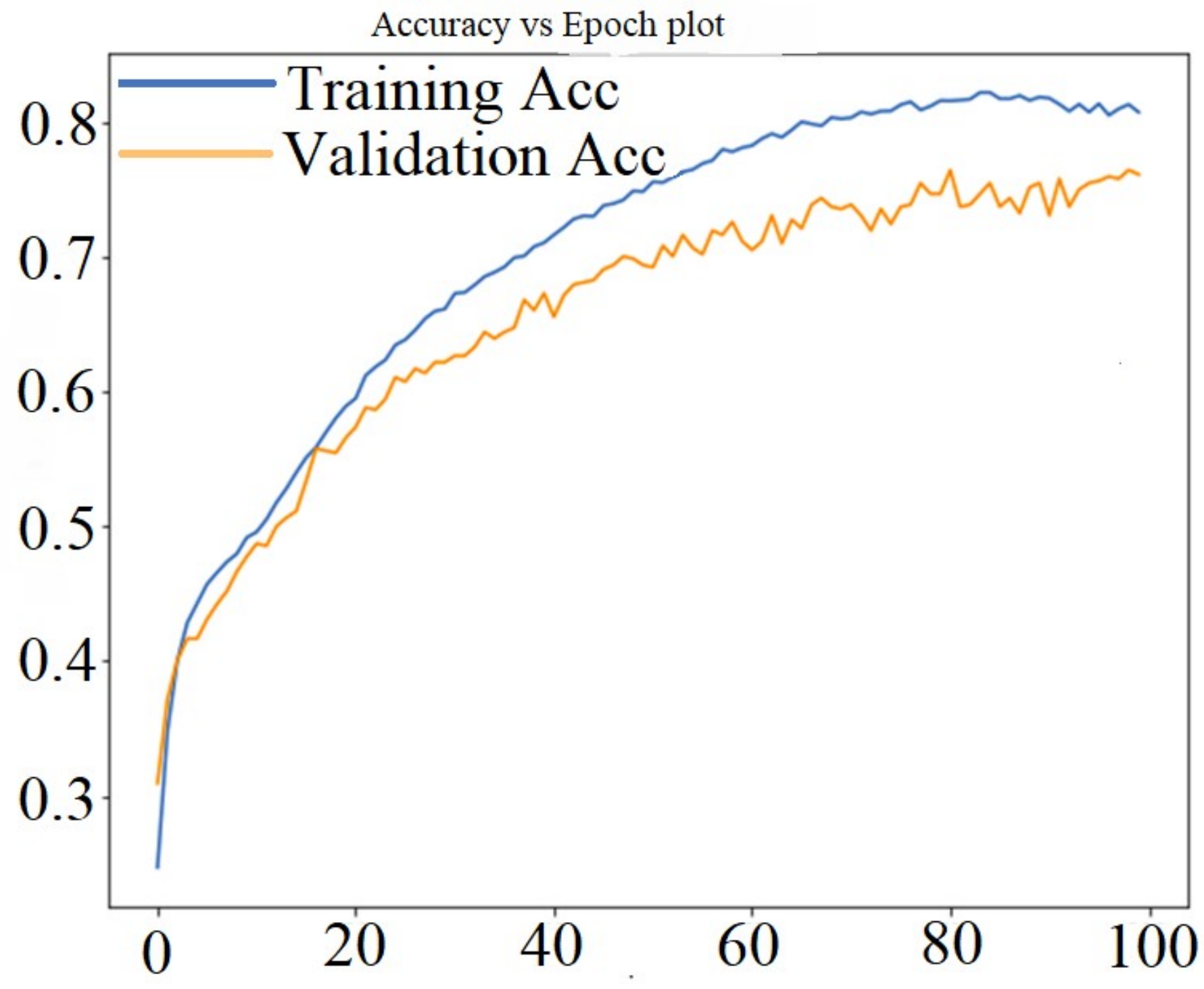


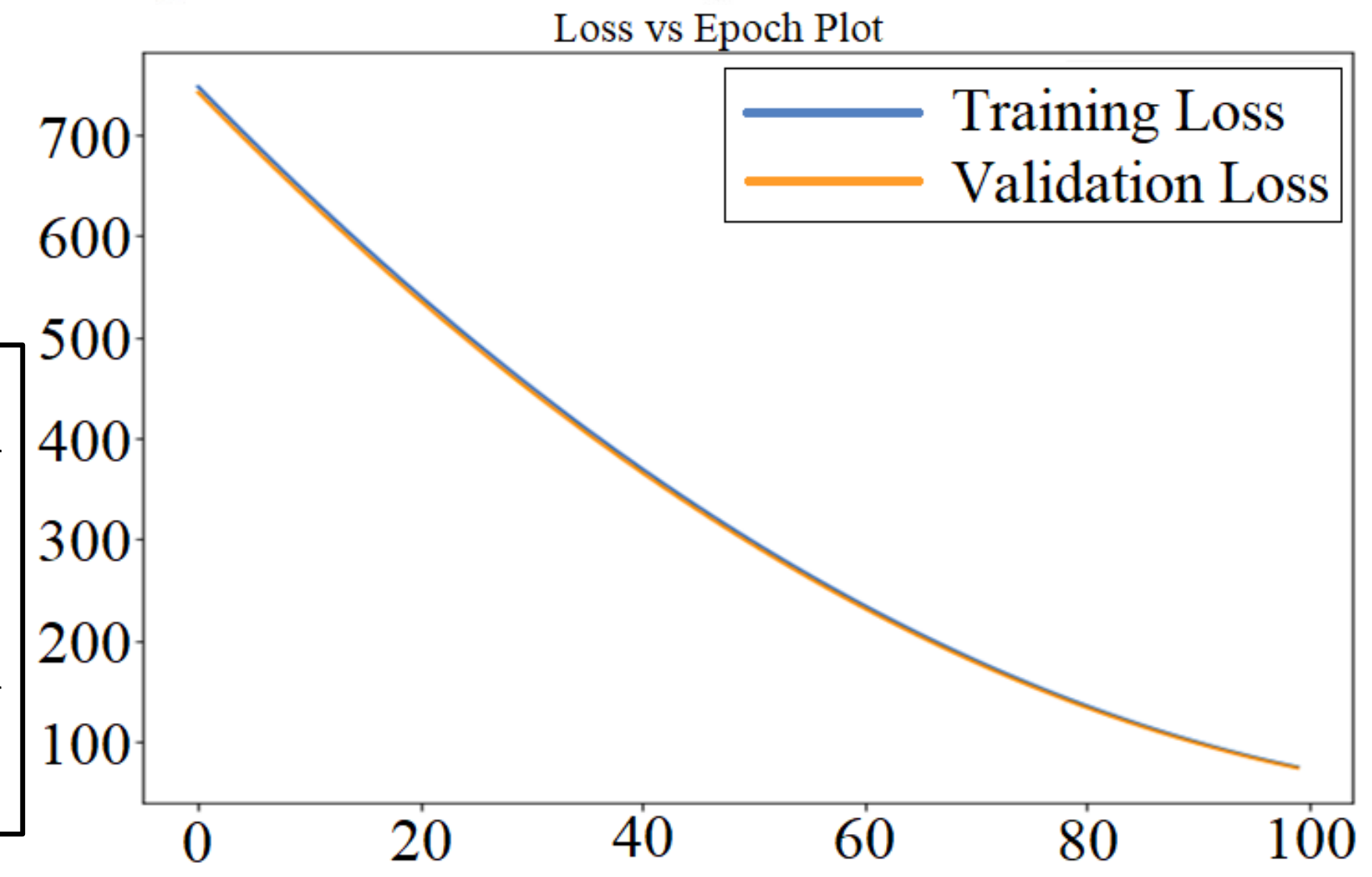
| | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| 0 | 114 | 5 | 11 | 5 | 8 |
| 1 | 14 | 104 | 5 | 5 | 14 |
| 2 | 14 | 10 | 107 | 10 | 3 |
| 3 | 8 | 4 | 11 | 99 | 13 |
| 4 | 3 | 15 | 4 | 13 | 94 |

| Label | Sentences |
|---|---|
| 0 | बत्तिको अवस्था बदल |
| 1 | आजको मौसम बताउ |
| 2 | 'एउटा सङ्गीत बजाउ |
| 3 | 'पङ्खाको स्थिती बदल |
| 4 | अबको समय सुनाउ |

- Raw signal included DC offsets, line noise and ECG artifacts
- Normalization and notch filtration removed the DC offsets and line noise
- ECG artifacts were removed via Ricker wavelet







- The CNN model consisted of 2 convolution, 2 max pooling layers and 4 dense layers
- Accuracy of about 80% in the training dataset and 74.75% in the testing dataset was achieved

## RECOMMENDATIONS

- Explore other machine learning models that perform better with time series inputs
- Consider the accuracy and consistency in electrode placement during dataset creation
- Take into account muscle fatigue during long sessions of silent speech recordings

## CONCLUSION

- Silent speech recognition by decoding neuromuscular signals is feasible with a permissible error rate
- System performance greatly depends upon electrode arrangement, utterance rate and the quality of the training dataset