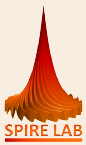
Low frequency characteristics are the key differentiators between dysarthric speech in ALS and healthy speech

Jhansi M[1], Aravind I[1], Yamini BK[2], Nalini A[3], Seena V[3], Saraswathi N[3], Ravi Y[3], Prasanta Kumar Ghosh[1]

[1]SPIRE Lab, Electrical Engineering, Indian Institute of Science (IISc), Bangalore, India
[2]Dept. of Speech Pathology and Audiology,
[3]Dept. of Neurology, NIMHANS, Bangalore, India



October 19, 2020

Introduction

Amyotrophic Lateral Sclerosis (ALS): A motor neuron disorder which occurs due to gradual degeneration of motor neurons

Introduction

Amyotrophic Lateral Sclerosis (ALS): A motor neuron disorder which occurs due to gradual degeneration of motor neurons

Objective of this work: To learn representations through neural network based approach for classification of dysarthric speech in ALS and healthy (HC) speech

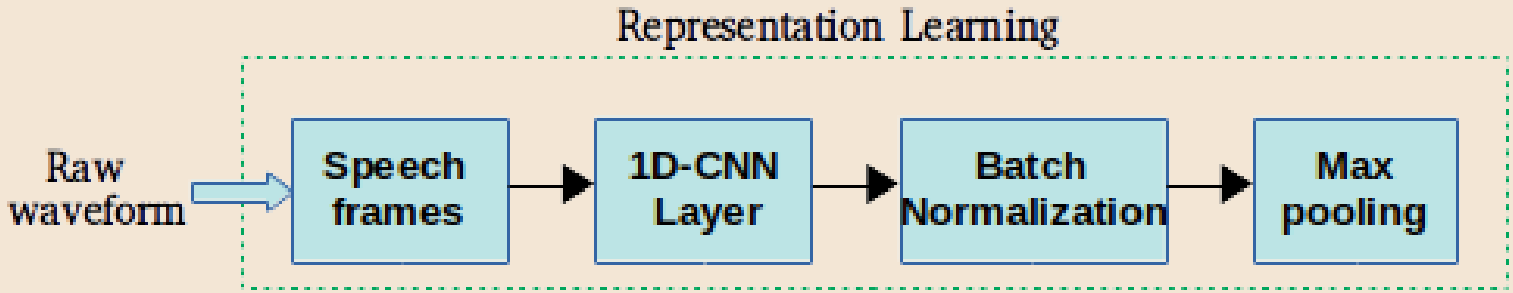
Introduction

Amyotrophic Lateral Sclerosis (ALS): A motor neuron disorder which occurs due to gradual degeneration of motor neurons

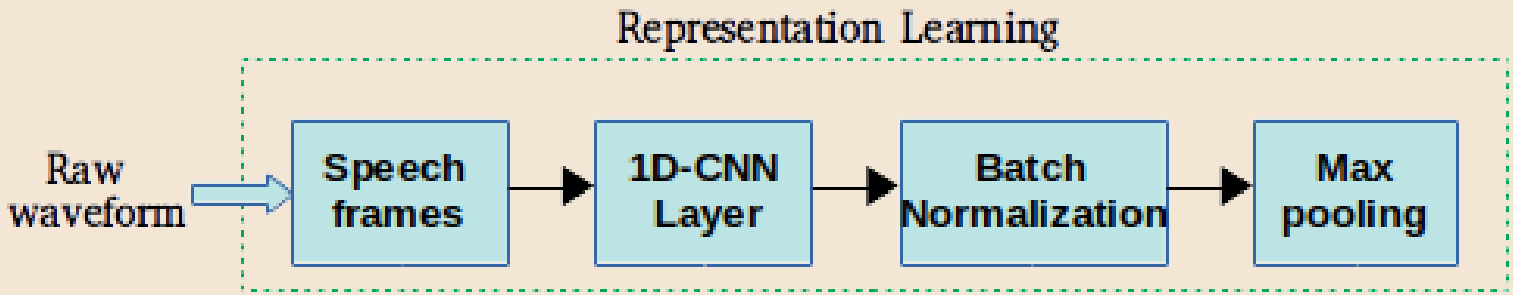
Objective of this work: To learn representations through neural network based approach for classification of dysarthric speech in ALS and healthy (HC) speech

Motivation: To develop a speech based application to detect and monitor the progression of ALS at an early stage

Proposed approach



Proposed approach



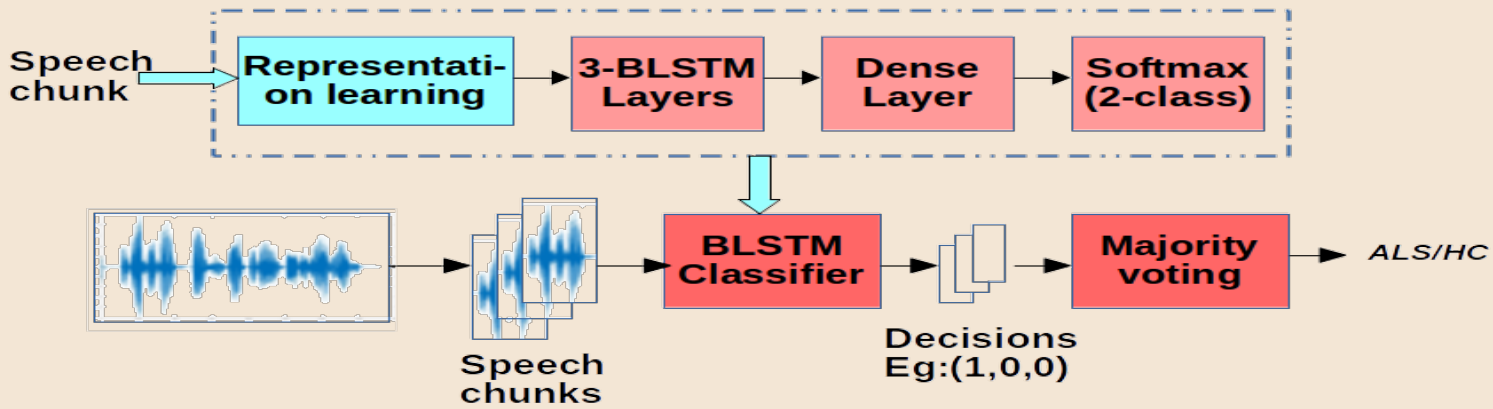


Figure: CNN-BLSTM architecture

Experimental setup:

- **Input features**: Raw speech waveform
- **Number of CNN filters**: 256 (size: 12X1)
- **Number of BLSTM layers**: 3 (each with 150 units)
- **Activation function**: ReLU log (softmax @ output)
- **Evaluation metric**: Classification accuracy
- **Baseline**: Mel Frequency Cepstral Coefficients (MFCCs)

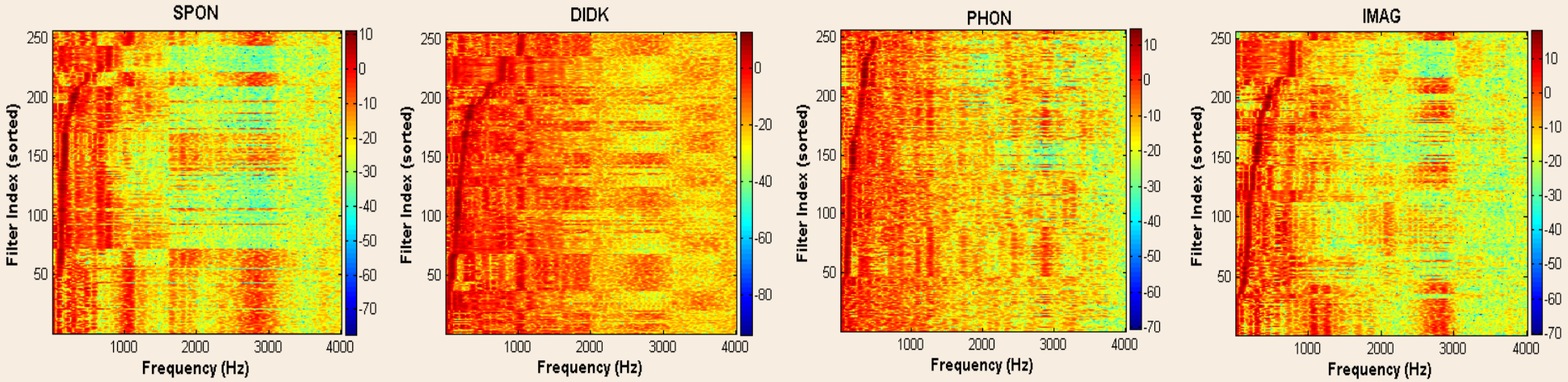
Experimental setup:

- **Input features**: Raw speech waveform
- **Number of CNN filters**: 256 (size: 12X1)
- **Number of BLSTM layers**: 3 (each with 150 units)
- **Activation function**: ReLU log (softmax @ output)
- **Evaluation metric**: Classification accuracy
- **Baseline**: Mel Frequency Cepstral Coefficients (MFCCs)

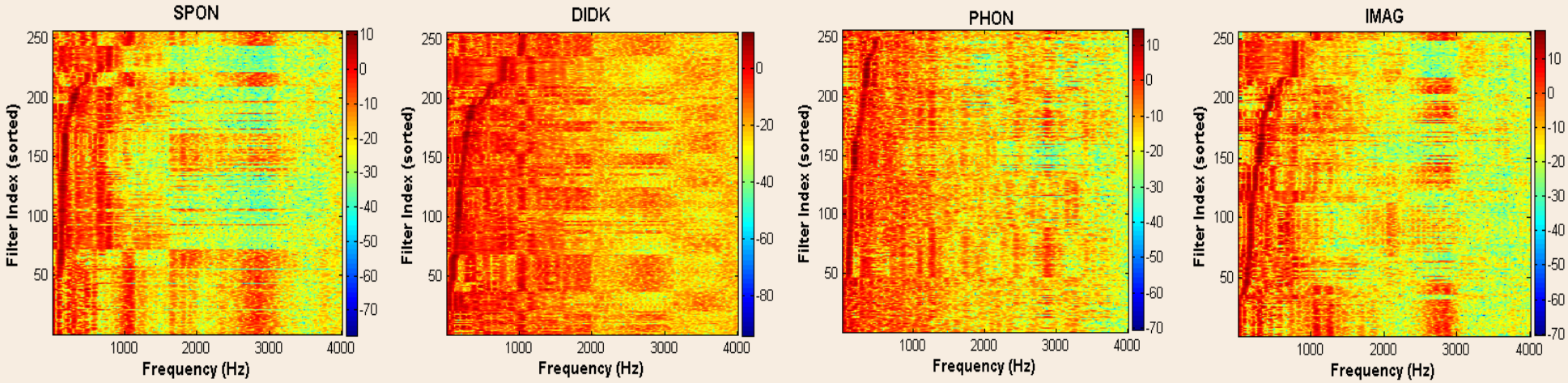
Data collection:

- Speech data is collected from National Institute for Mental Health and Neuro Sciences (NIMHANS), Bangalore, India
- **Number of subjects**: 60 ALS (30M, 30F), 60 HC(30M, 30F)
- **Recorder**: Zoom H-6 recorder
- **Speech tasks**: Spontaneous speech (SPON), Diadochokinetic rate (DIDK), Sustained phonation (PHON), and Image description (IMAG)

Magnitude response of CNN filters



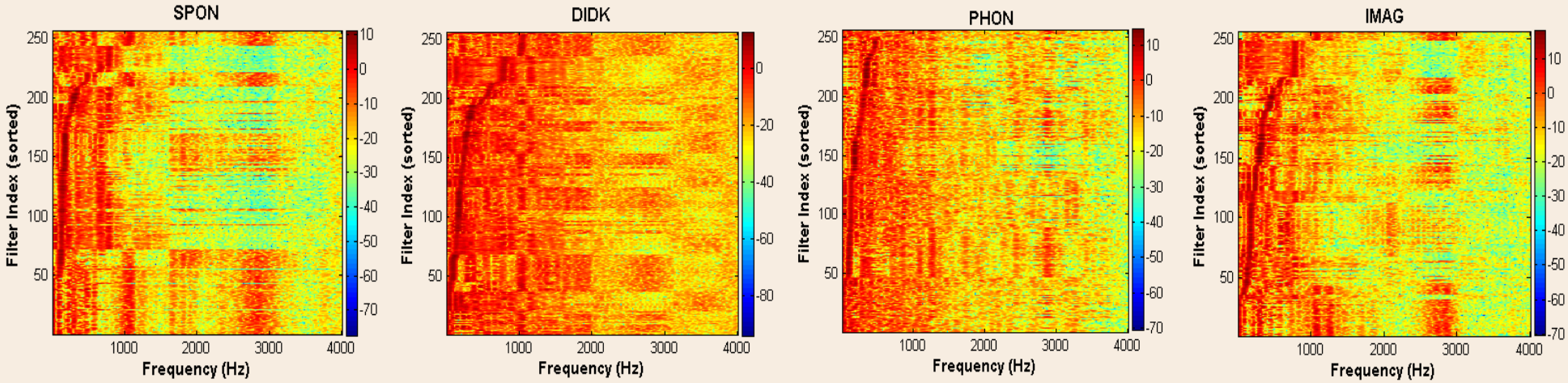
Magnitude response of CNN filters



Key findings

- Center frequencies of most of the learned filters are less than 400Hz leading to an average classification accuracy of 94.83%

Magnitude response of CNN filters



Key findings

- Center frequencies of most of the learned filters are less than 400Hz leading to an average classification accuracy of 94.83%

Acknowledgement

- We thank the Department of Science and Technology, Govt. of India, for their support in this work