Voice Based Classification of Healthy Subjects and Patients with Amyotrophic Lateral Sclerosis

Suhas BN, Deep Patel, Nithin Rao, Yamini Belur, Pradeep Reddy, Nalini Atcharyam
Ravi Yadav, Dipanjan Gope and Prasanta Kumar Ghosh

SPIRE Lab, Electrical Engineering, Indian Institute of Science, Bengaluru 560012, India
National Institute of Mental Health and Neurosciences, Bengaluru 560029, India

PROBLEM STATEMENT

To determine the choice of task and device in developing an assistive tool for detection and monitoring of ALS

MOTIVATION

- **Need for detection of ALS**: Early detection helps better management
- **Monitoring of ALS**: Following the progress of the condition [1]

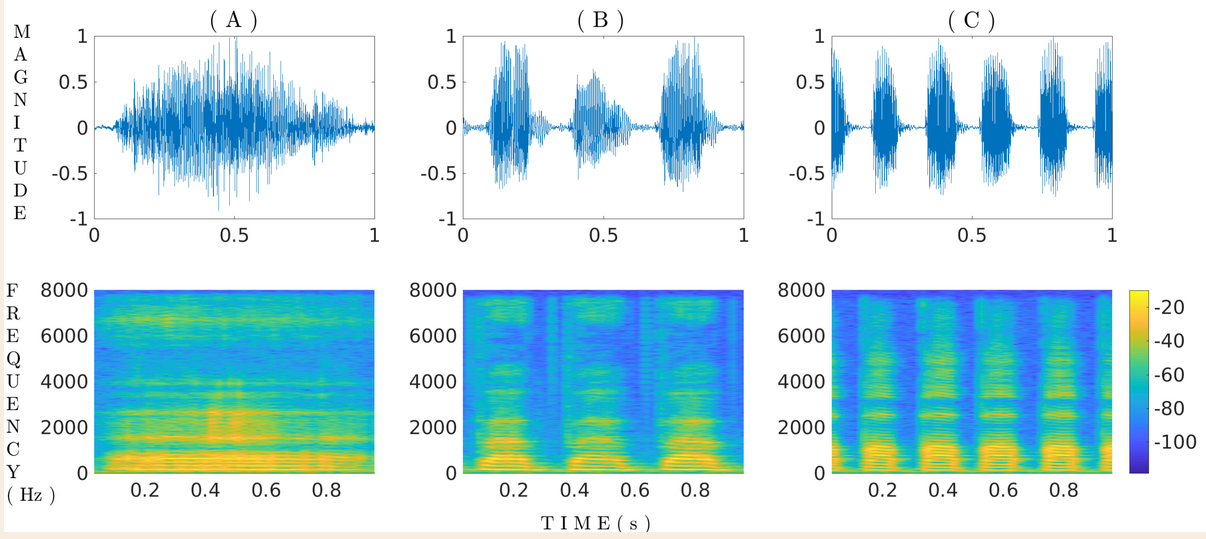
CHALLENGES

- Identifying speech cues that help in better diagnosis [2].
- Access for people of different socio-economic backgrounds

DATASET

Recording setup	Moto G5 Plus (MOT), Zoom H6 X/Y recorder (ZOO), Xiaomi Redmi 4 (XIA), Dell XPS 15 (LAP), Apple iPhone 7 (IPH).
Sampling freq	44.1 Khz
Speech Stimuli	Spontaneous (SPON), Diadochokinetic rate (DDK), Sustained Phonation (PHON)
Native Language (count)	Bengali (5 subjects each), Hindi (5 subjects each) , Kannada (5 subjects each), Odiya (3 subjects each), Tamil (3 subjects each), Telugu (4 subjects each).
Total duration	SPON (7 hours), DDK (5.36 hours), PHON (7.9 hours)

WAVEFORMS AND SPECTROGRAMS



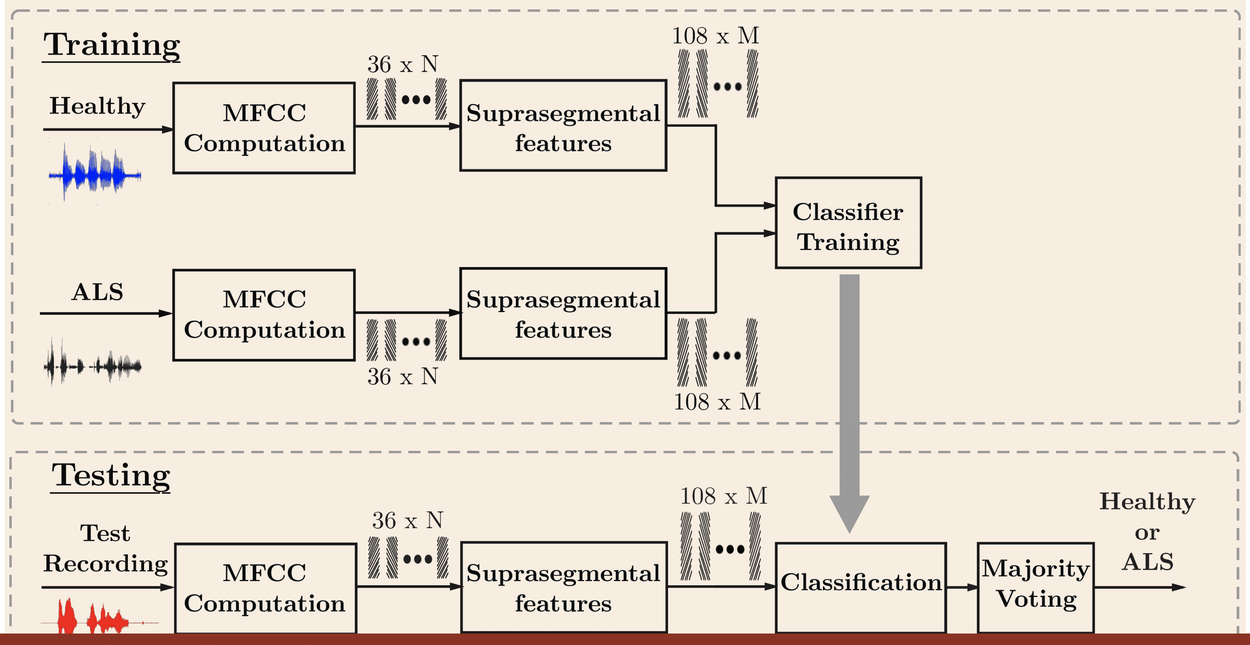
(A) subject with ALSFRS-R 0 (B) subject with ALSFRS-R 2 (C) A healthy subject

ALSFRS-R RATING FOR SPEECH

Severity	0	1	2	3	4
Finding	Loss of useful speech	Speech combined with nonvocal communications	Intelligible with repeating	Detectable speech disturbance	Normal

Scan here for audio samples

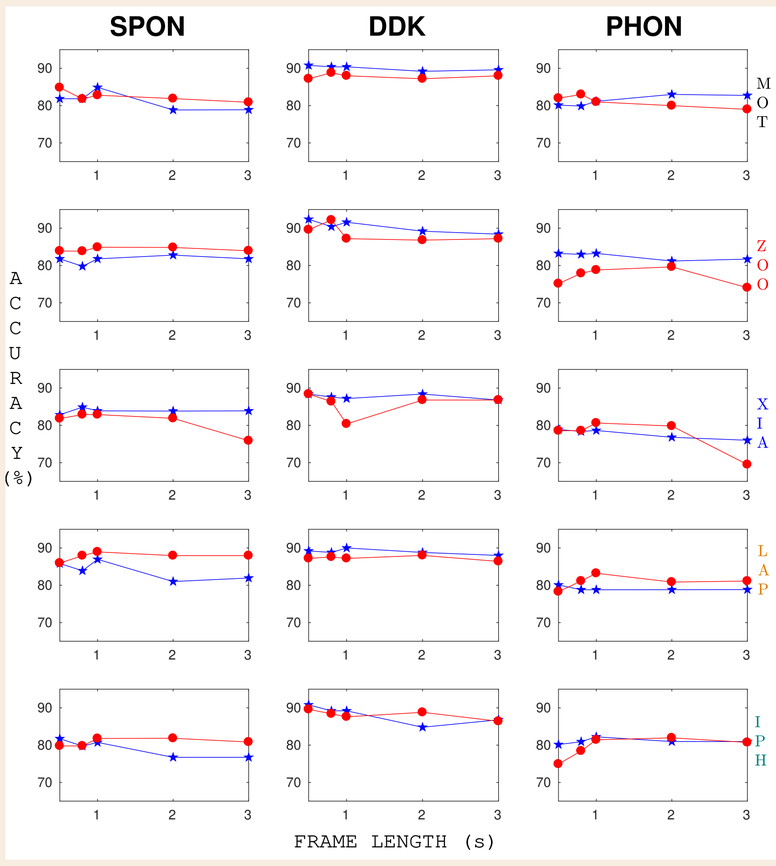
METHODOLOGY



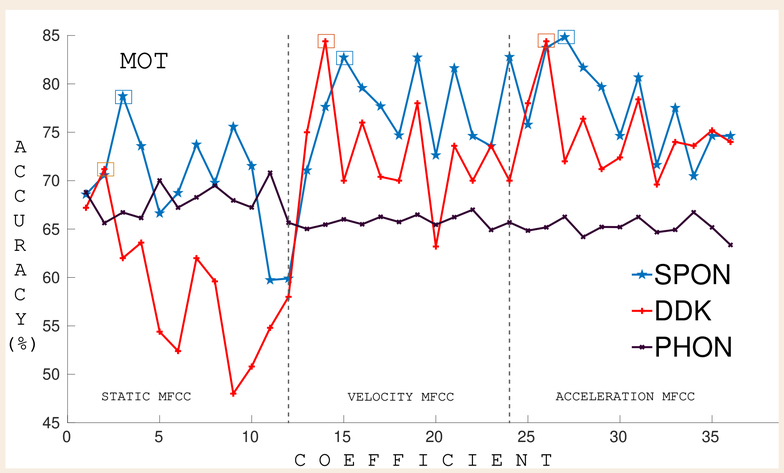
EXPERIMENTS

- 50 subjects, 5 fold cross validation setup.
- Per fold : 5 ALS & 5 Healthy Controls balanced for ALS severity, age and language.
- **Classifiers** : Support Vector Machines (**SVM**) & Deep Neural Networks (**DNN**)
- **Parameters**: Choices of N_w = {0.5, 0.8, 1, 2, 3}s with N_{sh} = 0.1s. Optimal N_w=0.8s.
- **SVM**: C & γ selected by maximizing the performance on the validation set
  DNN: Activation Functions = {'sigmoid', 'tanh', 'relu'}
  Hidden Layers = {1, 2, 3}
  & No. of Neurons={64, 128, 256, 512} for which the validation loss is minimized.

RESULTS



Classification accuracy by varying N_w. ★ SVM, ● DNN.



Classification results of (3xN) coefficients of suprasegmental features of each mean, median and S.D of one MFCC.

- A comparative study on speech classification of ALS vs Healthy controls using 3 speech tasks & 5 recording devices.
- **DDK task** consistently performs better than other tasks while MOT performs the best across all devices.
- SPON: 3rd, 15th & 27th dim of MFCC yield highest accuracy : 78.7%, 82.7% and 84.8% among 36-D MFCC : correspond to 4th DCT basis function.
- DDK: 2nd, 14th & 26th dim of MFCC yield highest accuracy : 71.2%, 84.4% and 84.4% among 36-D MFCC : correspond to 3rd DCT basis function.

REFERENCES

1. B.K. Yamini and A. Nalini, 'Measures of maximum performance of speech related tasks in patients with amyotrophic lateral sclerosis,' Amyotrophic Lateral Sclerosis, vol. 9, 2008.
2. Aravind Illa, Deep Patel, BK Yamini, N Shivashankar, Preethish Kumar Veeramani, Kiran Polavarapui, Saraswati Nashi, Atchayaram Nalini, Prasanta Kumar Ghosh, et al., 'Comparison of speech tasks for automatic classification of patients with amyotrophic lateral sclerosis and healthy subjects,' in 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2018, pp. 6014-6018.

ACKNOWLEDGEMENT

Authors thank the Pratiksha Trust for their support.

Ongoing work :
Developing an assistive tool for detection and monitoring of ALS.

bnsuhas96@gmail.com, prasantg@iisc.ac.in