

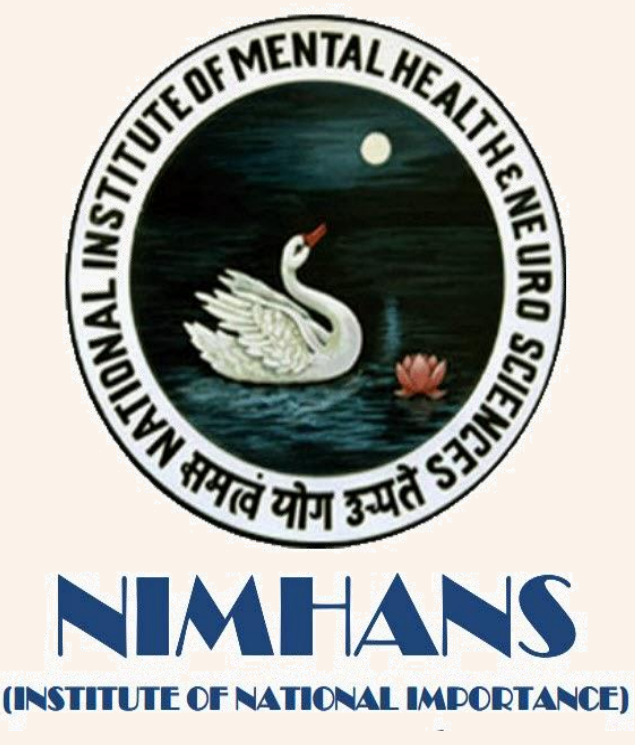
Low frequency characteristics are the key differentiators between dysarthric speech in ALS and healthy speech

Jhansi M*, Aravind I*, Yamini BK**, Nalini A\$, Seena V\$, Saraswathi N\$, Ravi Y\$, Prasanta Kumar Ghosh*

*SPIRE Lab, Electrical Engineering, Indian Institute of Science (IISc), Bangalore, India

**Dept. of Speech Pathology and Audiology,

\$Dept. of Neurology, National Institute for Mental Health and Neuro Sciences, Bangalore, India



Objective

To learn representations through neural network based approach for classification of dysarthric speech in Amyotrophic Lateral Sclerosis (ALS) and healthy (HC) speech

Motivation

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

Dataset

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)

Mean age	ALS(M)	ALS(F)	HC(M)	HC(F)
	58.60	56.02	44.21	46.93

M: Male
F: Female

Recorder: Zoom H-6 recorder

Speech tasks: Spontaneous speech (SPON), Diadochokinetic rate (DIDK), Sustained phonation (PHON), and Image description (IMAG)

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)

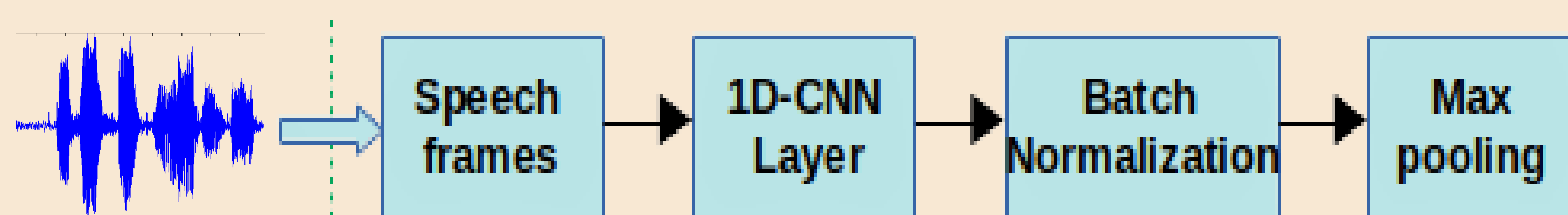
Evaluation metric: Classification accuracy

$$\text{accuracy\%} = \frac{\# \text{ correctly identified}}{\# \text{ total}}$$

Classification using **Mel Frequency Cepstral Coefficients (MFCCs)** as features with the proposed BLSTM network is considered as **baseline**

Proposed approach

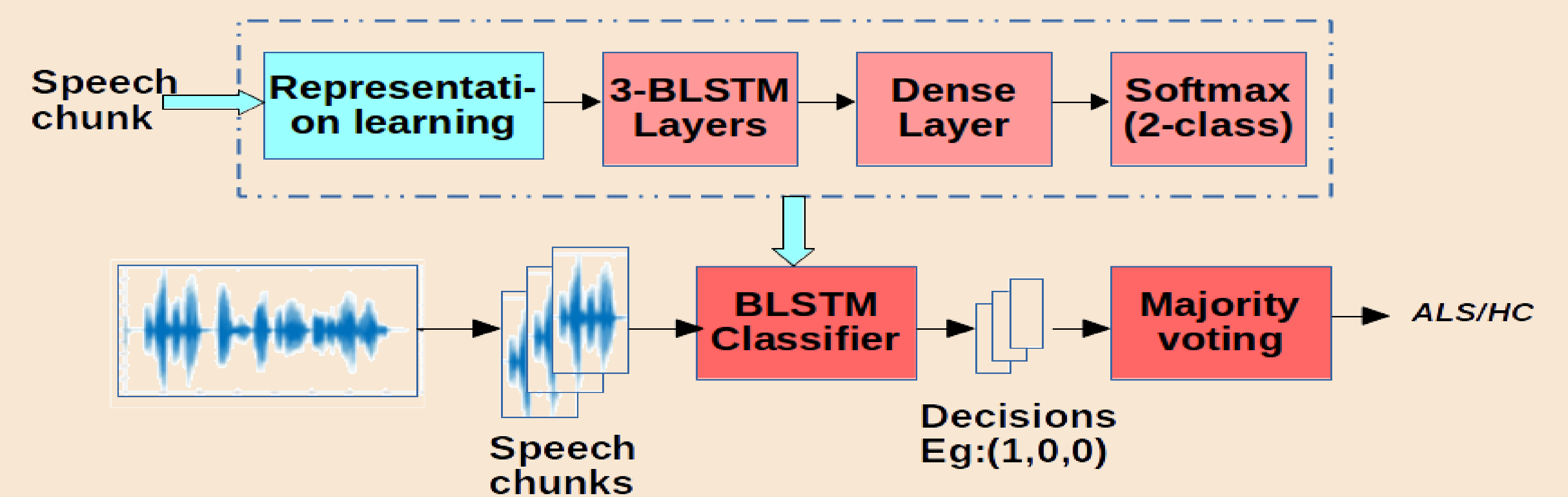
Learning representations from speech



CNN- Convolutional neural network

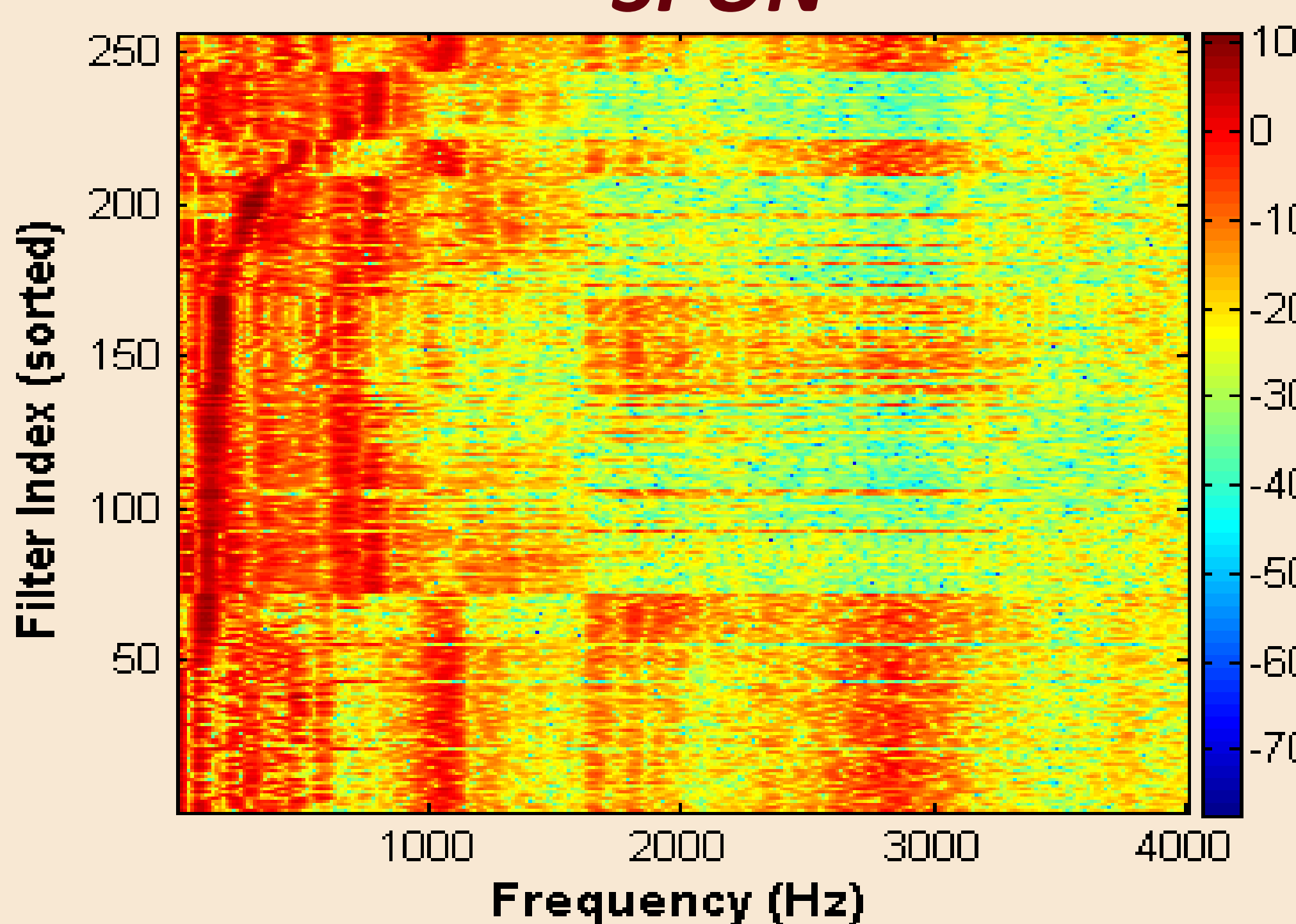
BLSTM- Bidirectional Long Short Term Memory

BLSTM Classifier

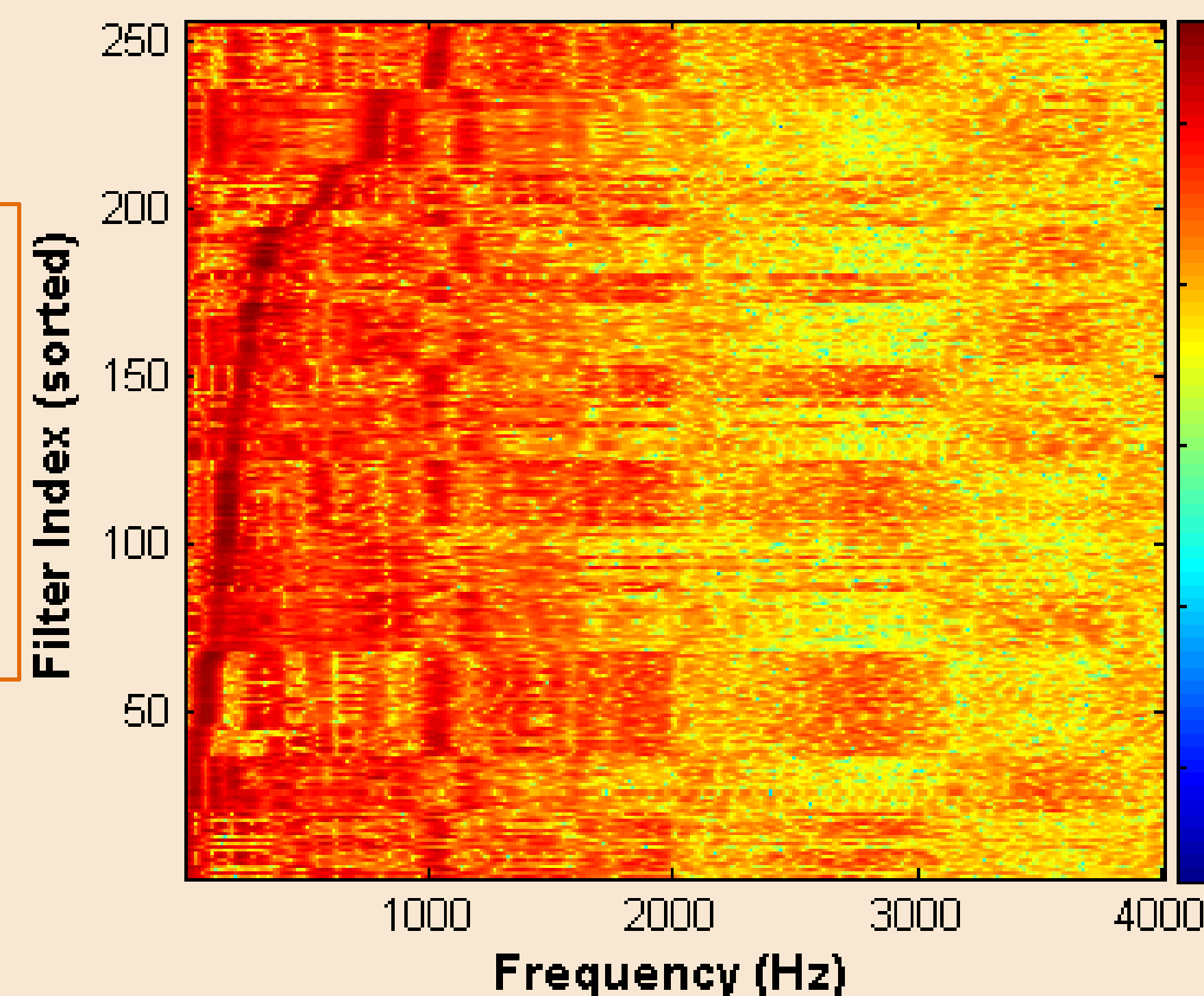


Magnitude response of CNN filters and classification accuracies

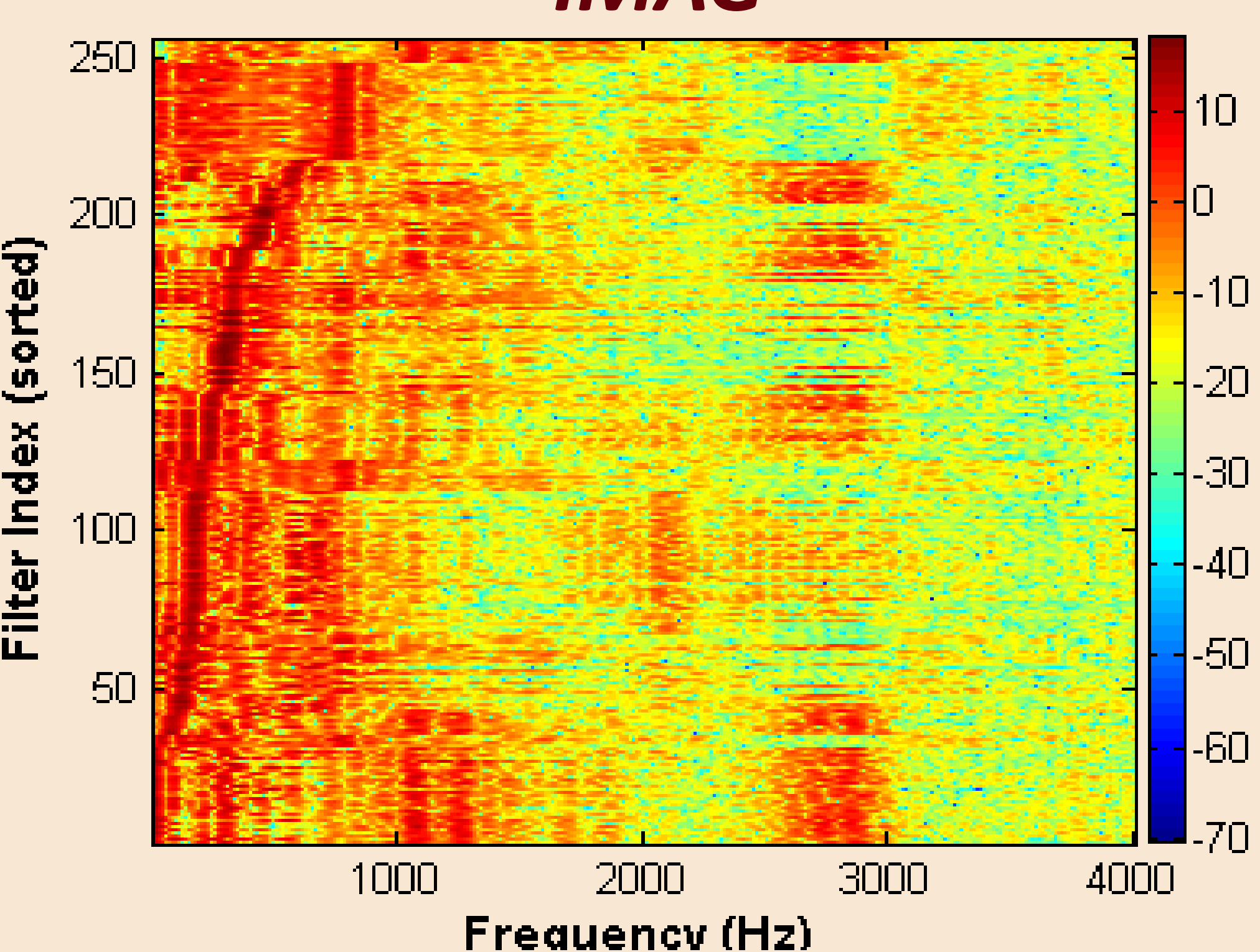
SPON



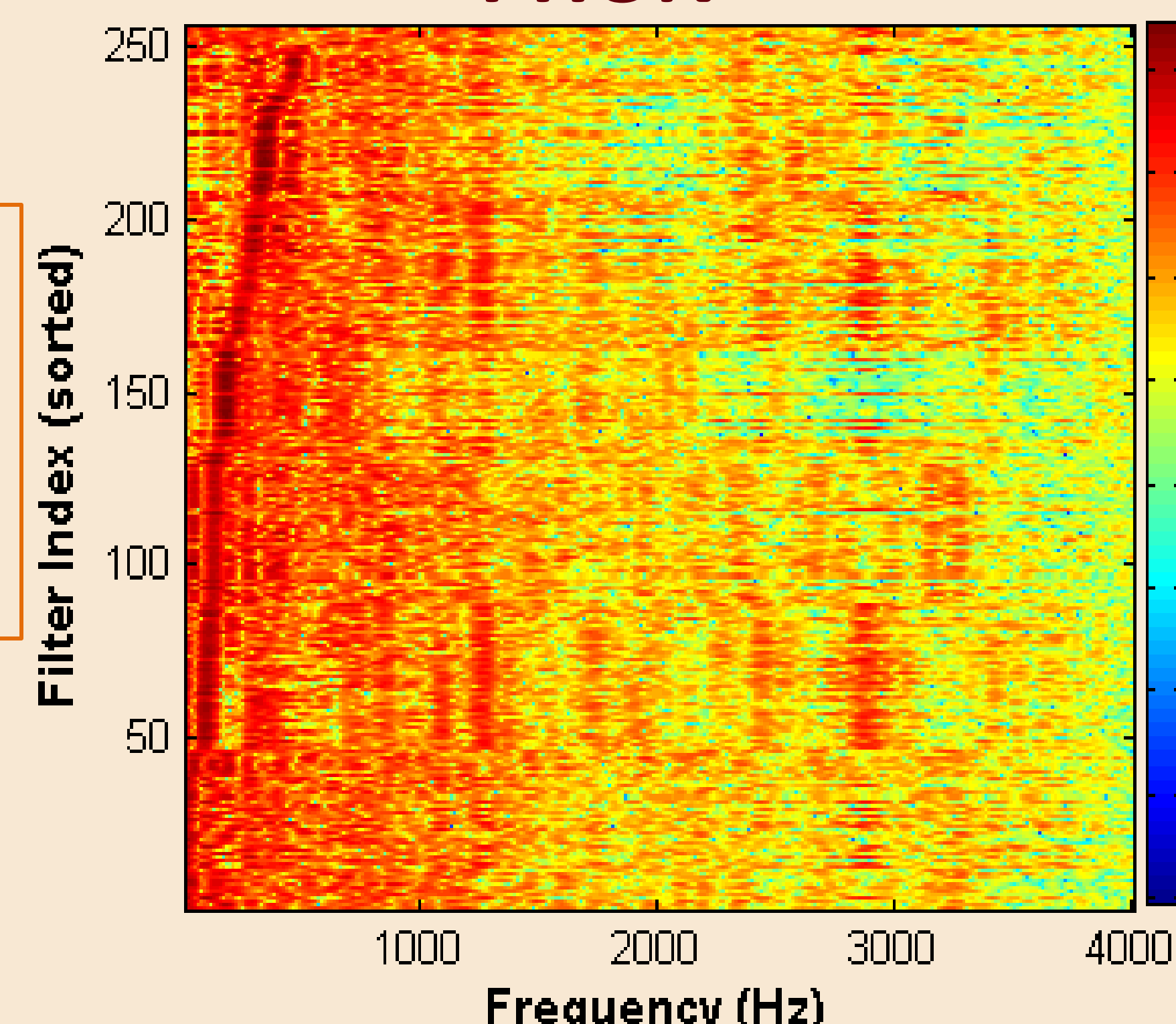
DIDK



IMAG



PHON



Key findings

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

Future work

To investigate the proposed approach for severity prediction of ALS patients

Acknowledgement

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

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

Jhansi M*, Aravind I*, Yamini BK**, Nalini A\$, Seena V\$, Saraswathi N\$, Ravi Y\$, Prasanta Kumar Ghosh*

*SPIRE Lab, Electrical Engineering, Indian Institute of Science (IISc), Bangalore, India

**Dept. of Speech Pathology and Audiology,

\$Dept. of Neurology, National Institute for Mental Health and Neuro Sciences, Bangalore, India

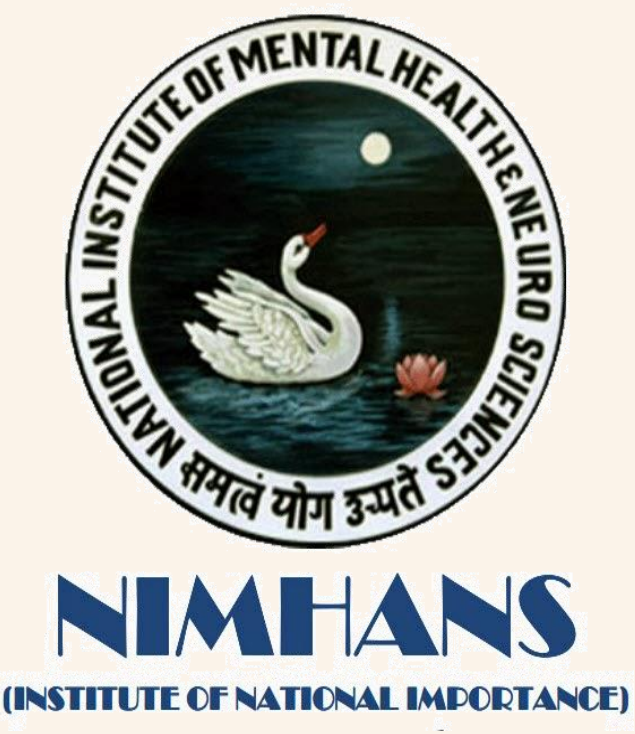
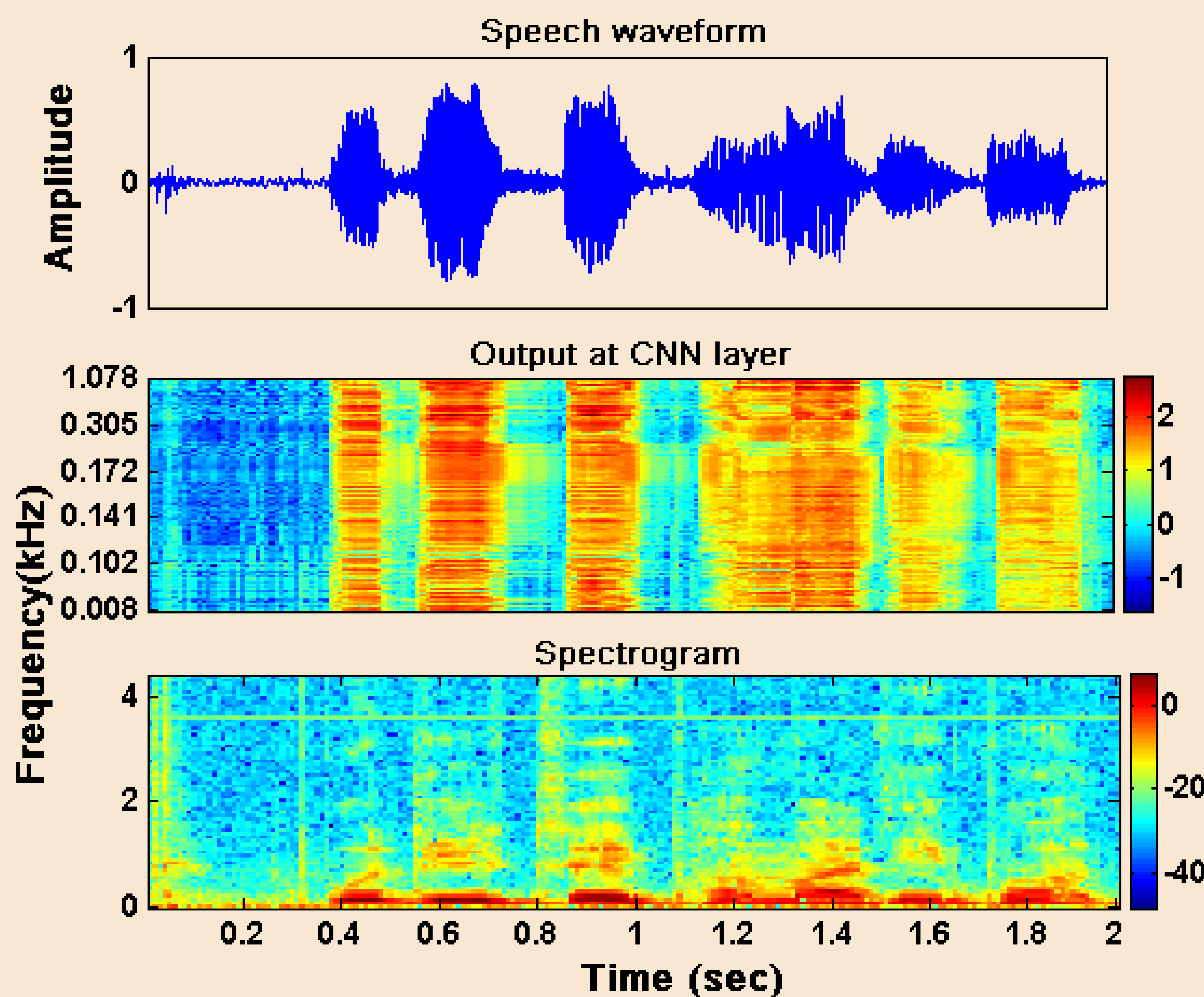
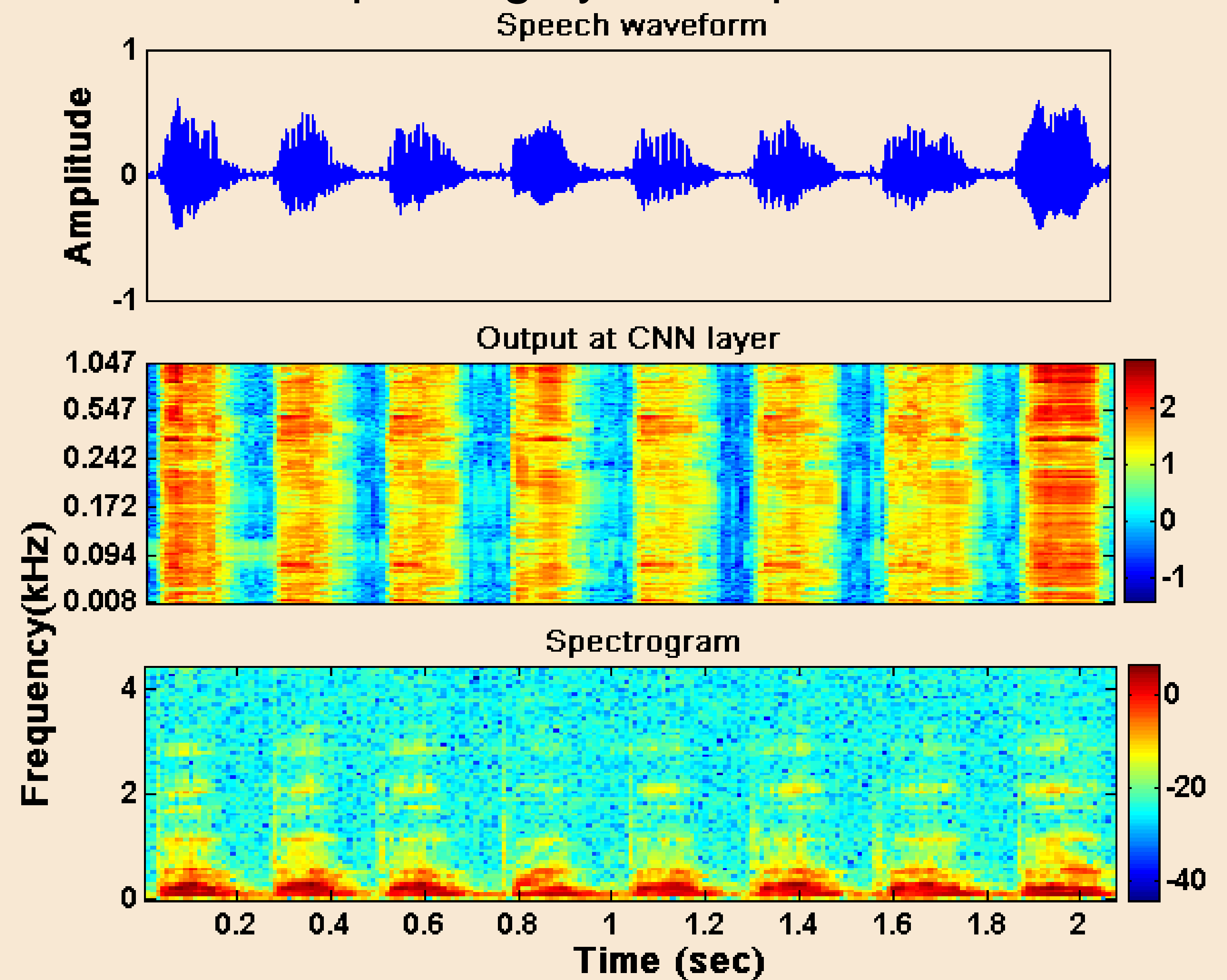


Illustration of 1-D CNN output

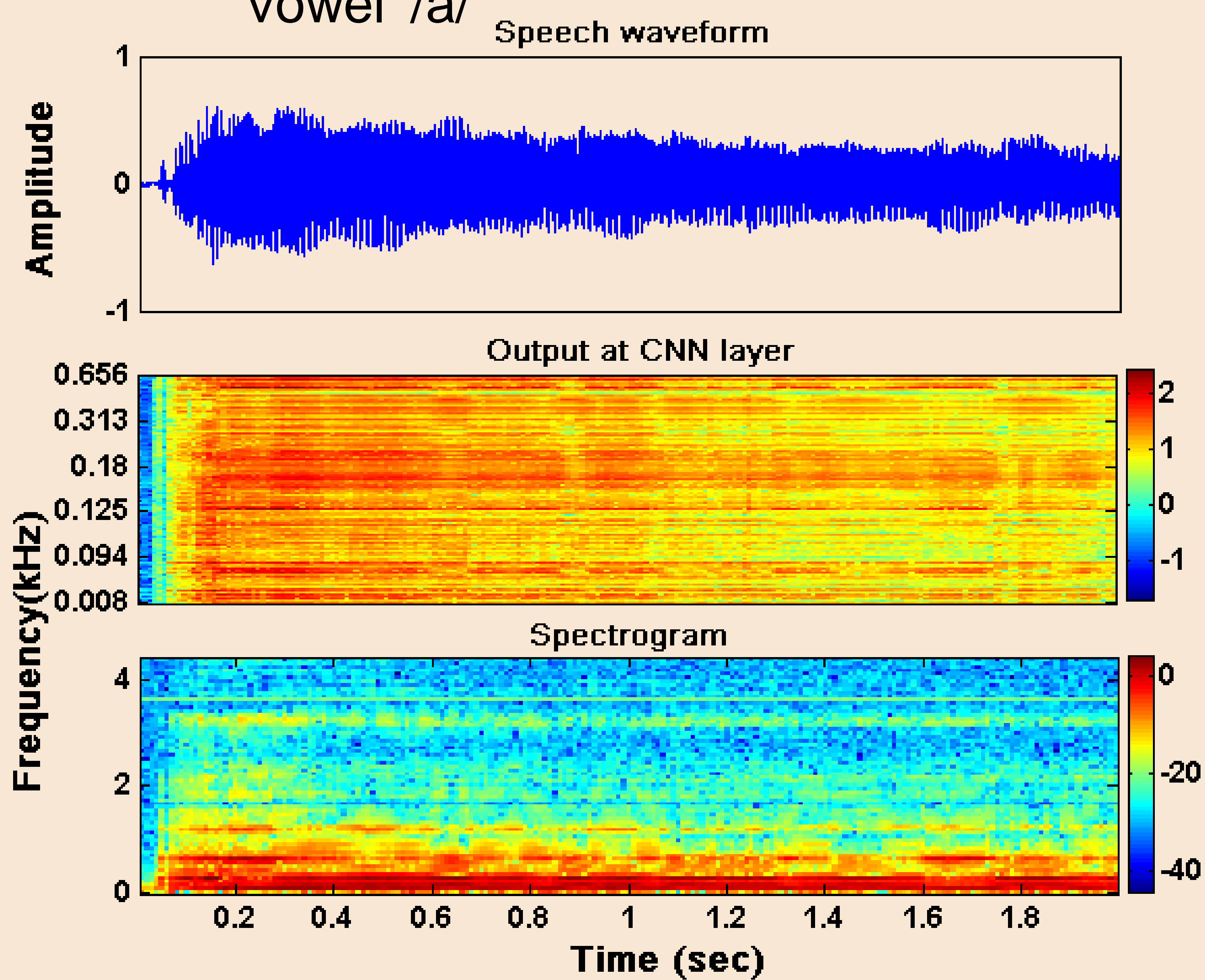
SPON : Speaking about a festival



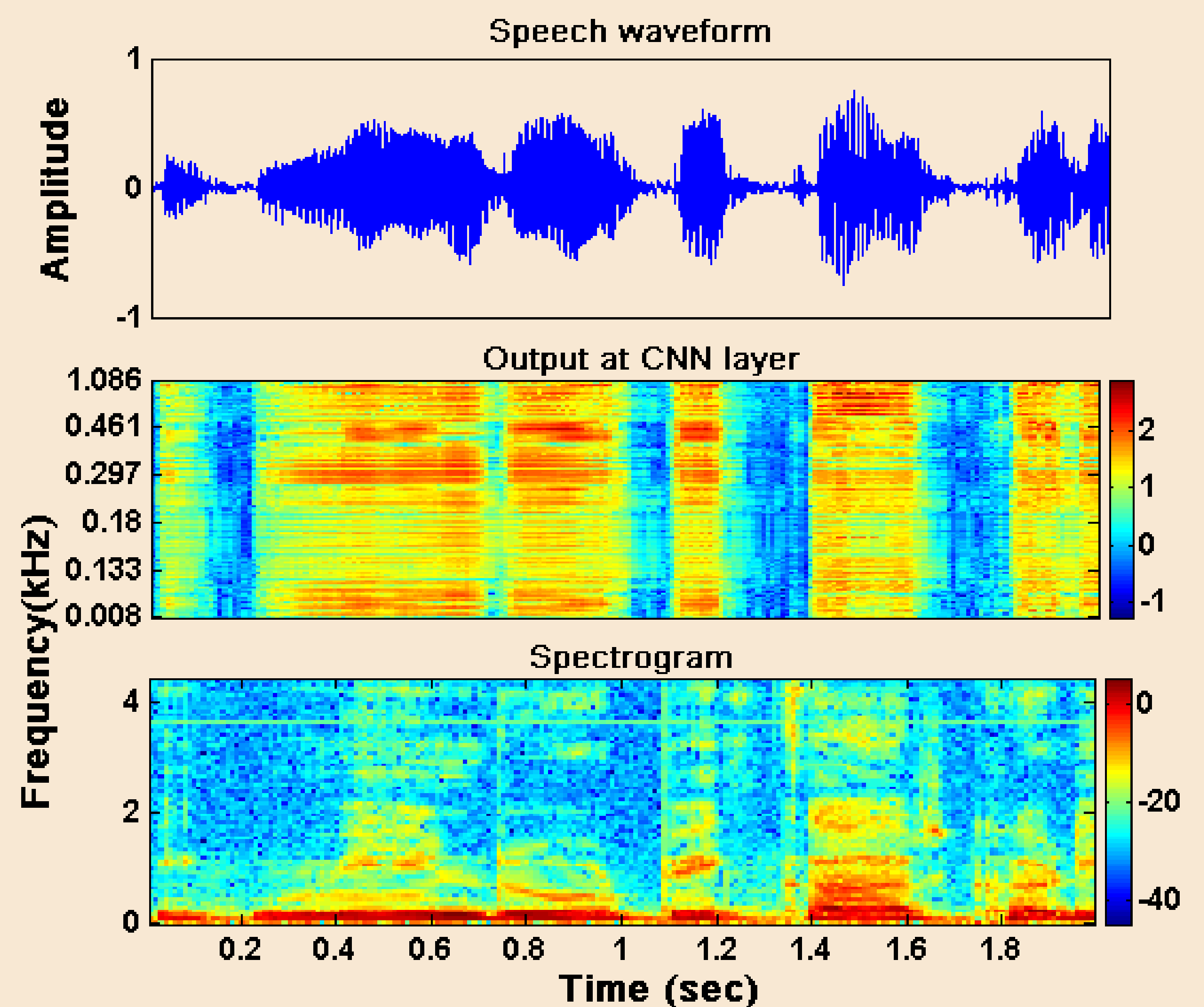
DIDK : Repeating syllable "pa"



PHON : Sustained phoneme production of vowel 'a'



IMAG : Describing an image shown



REFERENCES

1. R. D. Kent, R. L. Sufit, J. C. Rosenbek, J. F. Kent, G. Weismer, R. E. Martin, and B. R. Brooks, "Speech deterioration in amyotrophic lateral sclerosis: A case study," *Journal of Speech, Language, and Hearing Research*, vol. 34, no. 6, pp. 1269-1275, 1991.
2. "Zoom high-quality unidirectional microphone, available online: <https://www.zoom-na.com/>, last accessed:10/15/2019." [Online]. Available: <https://www.zoom-na.com/products/productaccessories/zoom-xyh-6-xy-stereo-microphone-capsule>
3. A. Nalini, K. Thennarasu, M. Gourie-Devi, S. Shenoy, and D. Kulshreshtha, "Clinical characteristics and survival pattern of 1153 patients with amyotrophic lateral sclerosis: experience over 30 years from India," *Journal of the neurological sciences*, vol. 272, no. 1-2, pp. 60-70, 2008.
4. B. Yamini, N. Shivashankar, and A. Nalini, "Vowel space area in patients with Amyotrophic Lateral Sclerosis," *Amyotrophic Lateral Sclerosis*, vol. 9, no. 1, pp. 118-119, 2008.
5. A. Illa and P. K. Ghosh, "Representation learning using convolution neural network for acoustic-to-articulatory inversion," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 5931-5935.
6. B. Suhas, D. Patel, N. Rao, Y. Belur, P. Reddy, N. Atchayaram, R. Yadav, D. Gope, and P. K. Ghosh, "Comparison of speech tasks and recording devices for voice based automatic classification of healthy subjects and patients with amyotrophic lateral sclerosis," *Proc. Interspeech 2019*, pp. 4564-4568.