Health Information Systems Lab Challenge
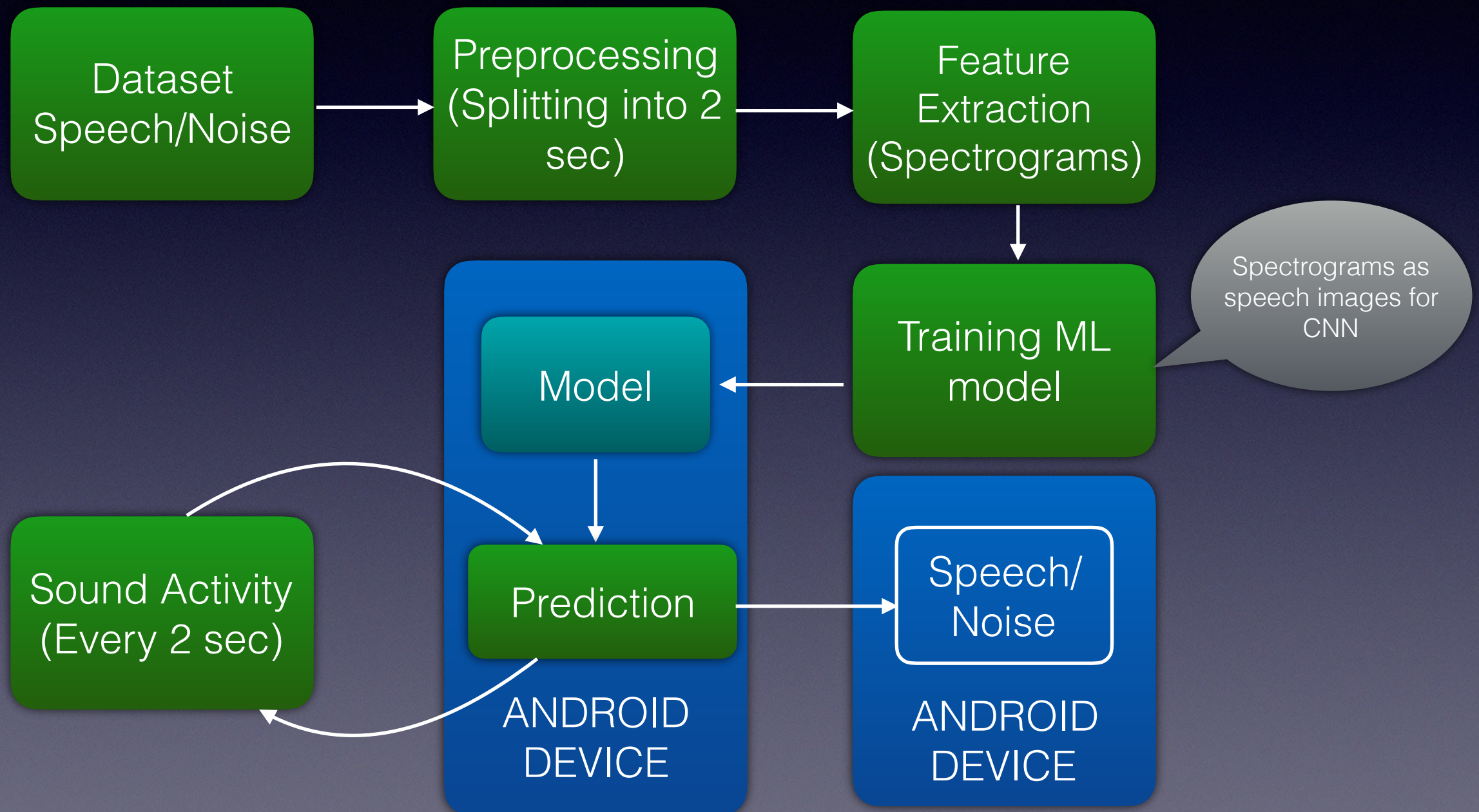# (with Tensorflow)

Prabhakaran Santhanam
MSc Computer Science ETH

Noise/Speech Classification

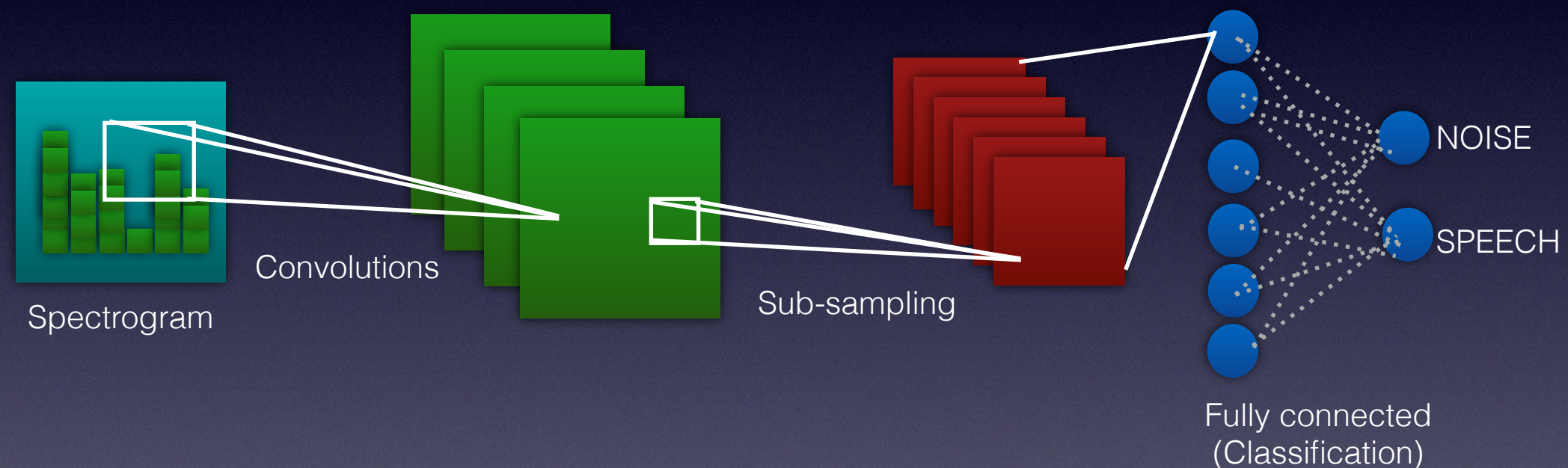

APPROACH

Dataset

- MUSAN: A corpus of music, speech, and noise. <u>Link</u>.

- Each .wav file is broken to 2 sec using <u>ffmpeg</u>.

Number of Samples	Total	Speech	Noise
Train	2000	1000	1000
Validation	400	200	200

Model
Convolutional Neural Networks

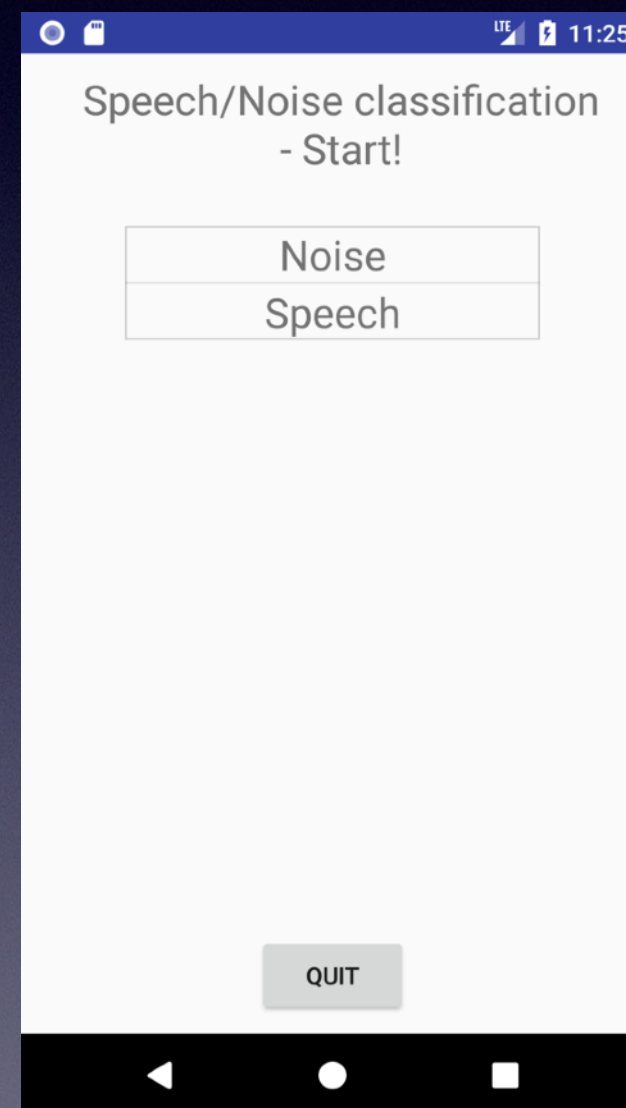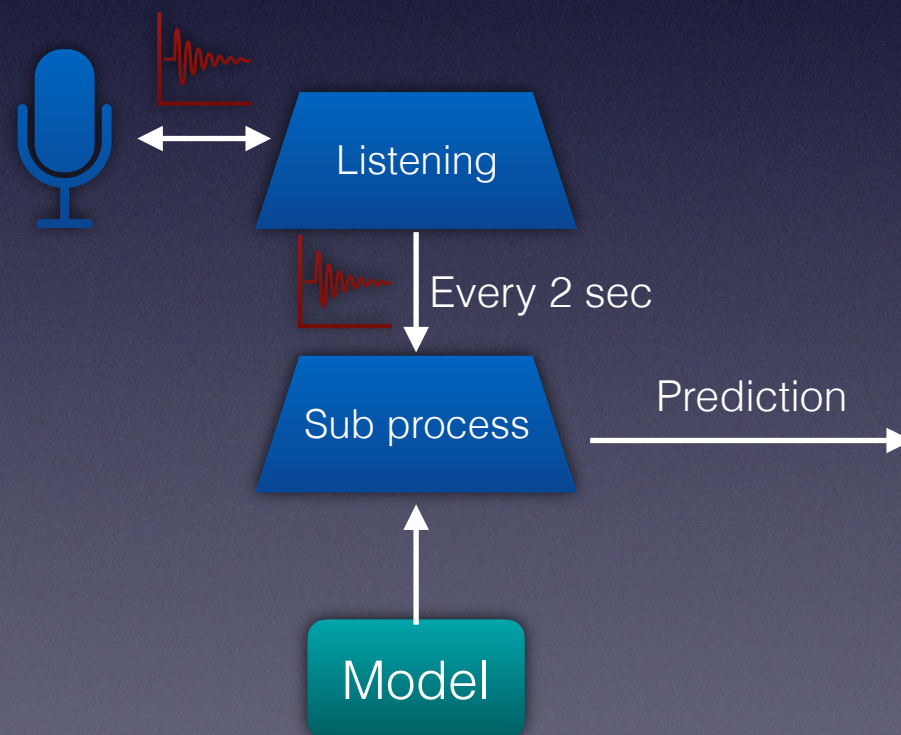Spectrogram   Convolutions   Sub-sampling   NOISE   SPEECH   Fully connected (Classification)

Sainath, T.N. and Parada, C., 2015. Convolutional neural networks for small-footprint keyword spotting. In Sixteenth Annual Conference of the International Speech Communication Association.Vancouver

Android UI and Prediction

Android Implementation

- Permission for writing on external storage and record Audio.

- Load the Model.

- Get the Audio Activity for every two seconds.

- Independently process the Audio and make prediction.

- TensorFlowInferenceInterface class and libandroid_tensorflow_inference_java.jar (platform dependent) for connecting Tensorflow to android.

Future Work

- Dataset: it needs to be large and robust. Then it can be trained with larger models.

- Models: Recurrent Neural Networks and other deep neural networks can be used.

- Preprocessing/Feature Extraction: Mel-Frequency Cepstral Coefficients (MFCC) can be used instead of Spectrograms.

References

- Chang, C.C. and Lin, C.J., 2011. LIBSVM: a library for support vector machines. ACM transactions on intelligent systems and technology (TIST), 2(3), p. 27.

- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H., 2009. The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), pp.10-18.

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. and Ghemawat, S., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv: 1603.04467.

- Han, W., Chan, C.F., Choy, C.S. and Pun, K.P., 2006, May. An efficient MFCC extraction method in speech recognition. In Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on (pp. 4-pp). IEEE.

- Vogel, L., 2013. Android development tutorial. Based on Android.