DEEPEAR: ROBUST SMARTPHONE AUDIO SENSING IN UNCONSTRAINED ACOUSTIC ENVIRONMENTS USING DEEP LEARNING

By
Sayed Khushal Shah(20)
&
Zeenat Tariq(22)
Team 6-1

SUMMARY

Challenges to audio sensing presented by diverse acoustic environments. Experimentally validate the design of DeepEar in two ways.

First, compare the model accuracy and robustness of DeepEar, under unconstrained environments, against existing audio sensing systems designed for mobile devices

Second, measure the energy and latency of a prototype DeepEar system designed for modern phone hardware

CONTRIBUTION

1. Deep Learning for Audio-based Sensing of Behavior and Context

DeepEar represents the first time computational models with a deep architecture have been developed to infer a broad set of human behavior and context from audio streams

2. Large-scale Study of Acoustic Environment Diversity

The challenge of performing audio sensing in diverse environments using real-world audio datasets spanning four common audio sensing tasks, along with audio captured from 168 place visits

3. Low-energy Smartphone Prototype

By directly utilizing the DSP, present in most smartphone sold today, it was demonstrated that DeepEar can perform continuous sensing with acceptable levels of energy and latency

ASPECTS

Positive

The feasibility of executing DeepEar-style modeling directly on mobile devices

A growing number of wearable incorporate closely related system-on-a-chip devices, to the one used in the Snapdragon 800 that was targeted in design

DeepEar has the feasibility to be deployed on mobile devices directly rather than the old applications which has to be trained for many hours using good PC's

Negative

Limitation in existing experiment results is the emphasis on comparing DeepEar with existing mobile audio sensing systems only

Neglects a number of techniques for offline (i.e., server side) audio analysis that are designed to combat diversity in acoustic environments

Further experiments are needed to understand how it compares to the latest in shallow learning for speech and general audio tasks

QUESTIONS

As we know from the paper that DeepEar was the first Deep Learning audio sensing application. What features were extracted from the audio data and trained?

What type of real World dataset were used to train DeepEar Model. What will be the affect on accuracy if we use Google Audio dataset for the same model?

In this paper they have used DNN, if CNN or RNN is used in the same pattern what will be the impact on energy efficiency?

Where the filtration technique is applied in the model?

How we can improve the accuracy and behavior of an audio for deep learning using small dataset where data has some environmental noise?


Thank
you!!