



# Speech Recognition with Smartphone Accelerometers

CHENG ZHANG

CSCI 6331

FINAL PROJECT

# Smartphone Sensors



## ▶ **Permission Required:**

- ▣ Voice sensors (Speaker)
- ▣ Image sensors (Camera)
- ▣ Navigation Sensors (GPS, Compass)

## ▶ **No Permission Required:**

- ▣ Motion sensors (Accelerometer, gyroscope)
- ▣ Light sensors

# Motion Sensors Threat

In 2014, a research shows that the smartphone gyroscope can pick up surface vibrations incurred by an independent loudspeaker placed on the same table (Michalevsky *et al.* USENIX Association)



Smartphone	Speakers	SVM	GMM	DTW
Galaxy S III	Mixed female/Male	20%	19%	17%
	Female speakers	30%	20%	29%
	Male speakers	32%	21%	25%

Table. Speaker recognition results

Source: Michalevsky, Y., Boneh, D., & Nakibly, G. (2014). Gyrophone: Recognizing speech from gyroscope signals. In 23rd {USENIX} Security Symposium ({USENIX} Security 14) (pp. 1053-1067).

# Feasibility: Accelerometer

## ► Effectiveness: Sampling Rate

### □ Voice Frequency

Human Voice Frequency Range	
Male	Female
80-180 Hz	165-300 Hz

### □ Accelerometer Sampling Rate

Smartphone	Year	Maximum Sampling Rate
Moto G4	2016	100 Hz
LG G5	2016	200 Hz
Huawei Mate 9	2016	250 Hz
Google Pixel 3	2018	410 Hz
Huawei Mate 20	2018	500 Hz

## ► Robustness: Human Activities

- The frequency of human activities is below 80 Hz

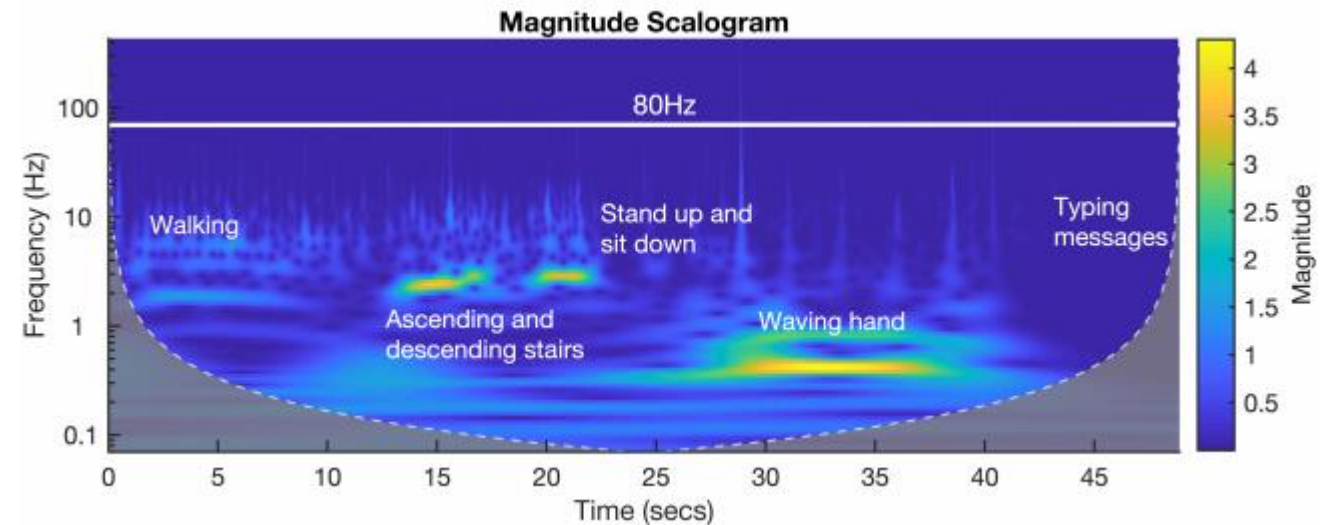


Fig. The response of a smartphone accelerometer to five human activities

# Accelerometer

## Data Collection



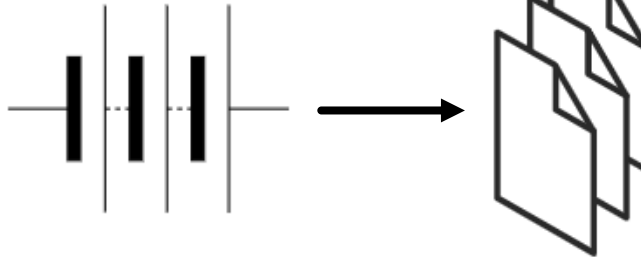
Microphone make  
a sound



Collect  
Accelerometer data



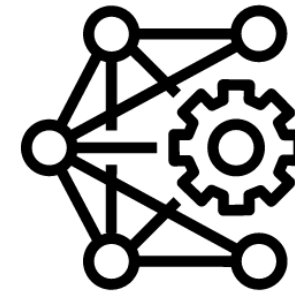
## Preprocessing



Signal-to-spectrogram  
conversion



## Train the model



Speech  
recognition



Zero  
One  
Two  
Three

Speech  
information

# Data Collection

## ► Dataset: Audio MNIST

- The dataset was used to play the audio on the App
- Contain 30,000 audio samples of spoken digits
- Range from 0 to 9

## ► Platform

- Smartphone: Huawei Mate 10
- Develop a simple Android App to collect data
- Play one speech signal on the smartphone and collect accelerometer data at the same time

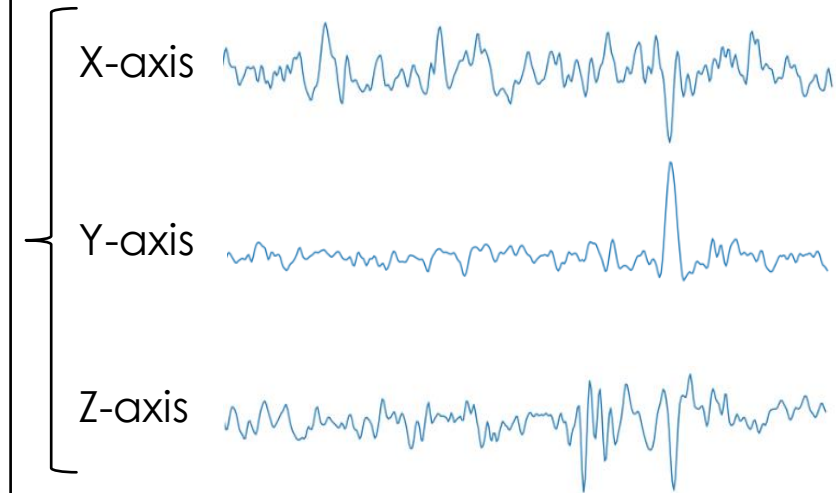
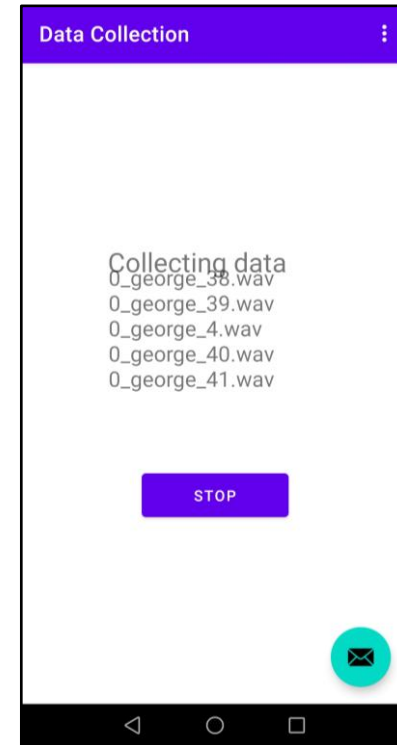


Fig. The accelerometer signal of one audio sample

# Preprocessing

## ► High-pass Filter

- Minimize distortions caused by hardware and human activities
- Eliminate frequency components below 80 Hz
- Apply to each axis of acceleration signal

## ► Signal-to-spectrogram Conversion

- Use short time Fourier transform (STFT) to calculate signal's spectrogram
- Each speech signal has three spectrograms
- The shape of spectrogram is  $24 \times 17 \times 3$  (Time  $\times$  Frequency  $\times$  Axis)

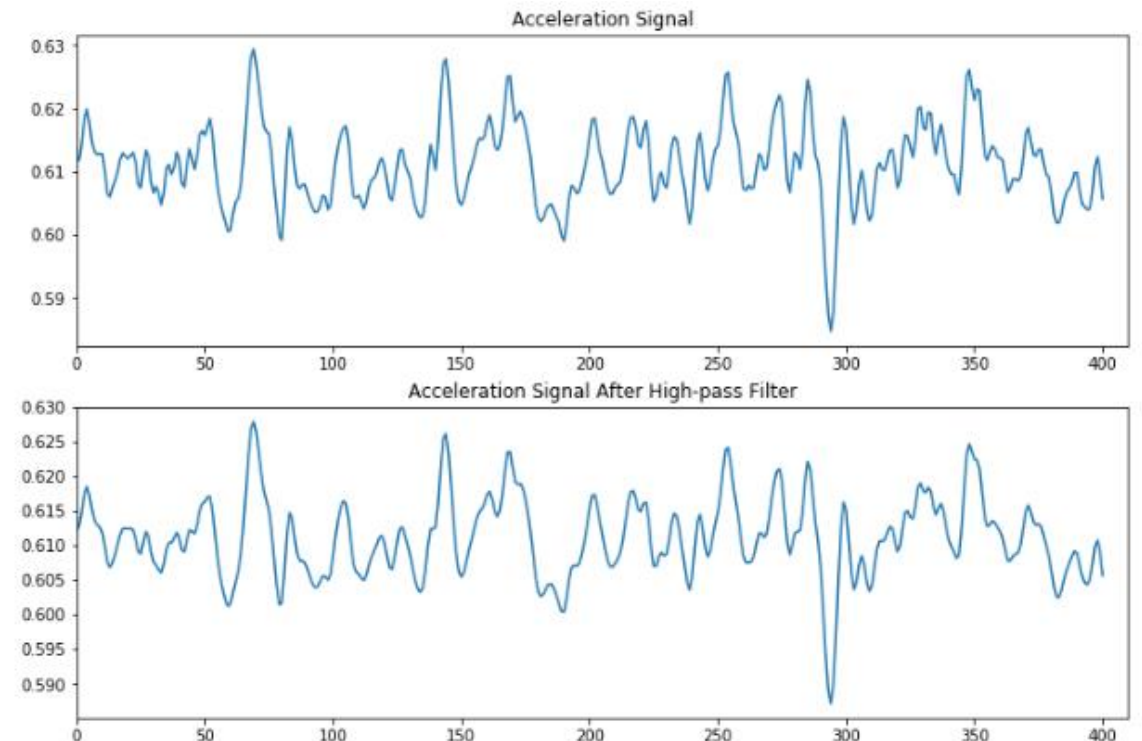


Fig. Acceleration signals processed with high-pass filter

# Preprocessing

## ► High-pass Filter

- Minimize distortions caused by hardware and human activities
- Eliminate frequency components below 80 Hz
- Apply to each axis of acceleration signal

## ► Signal-to-spectrogram Conversion

- Use short time Fourier transform (STFT) to calculate signal's spectrogram
- Each speech signal has three spectrograms
- The shape of spectrogram is  $24 \times 17 \times 3$  (Time  $\times$  Frequency  $\times$  Axis)

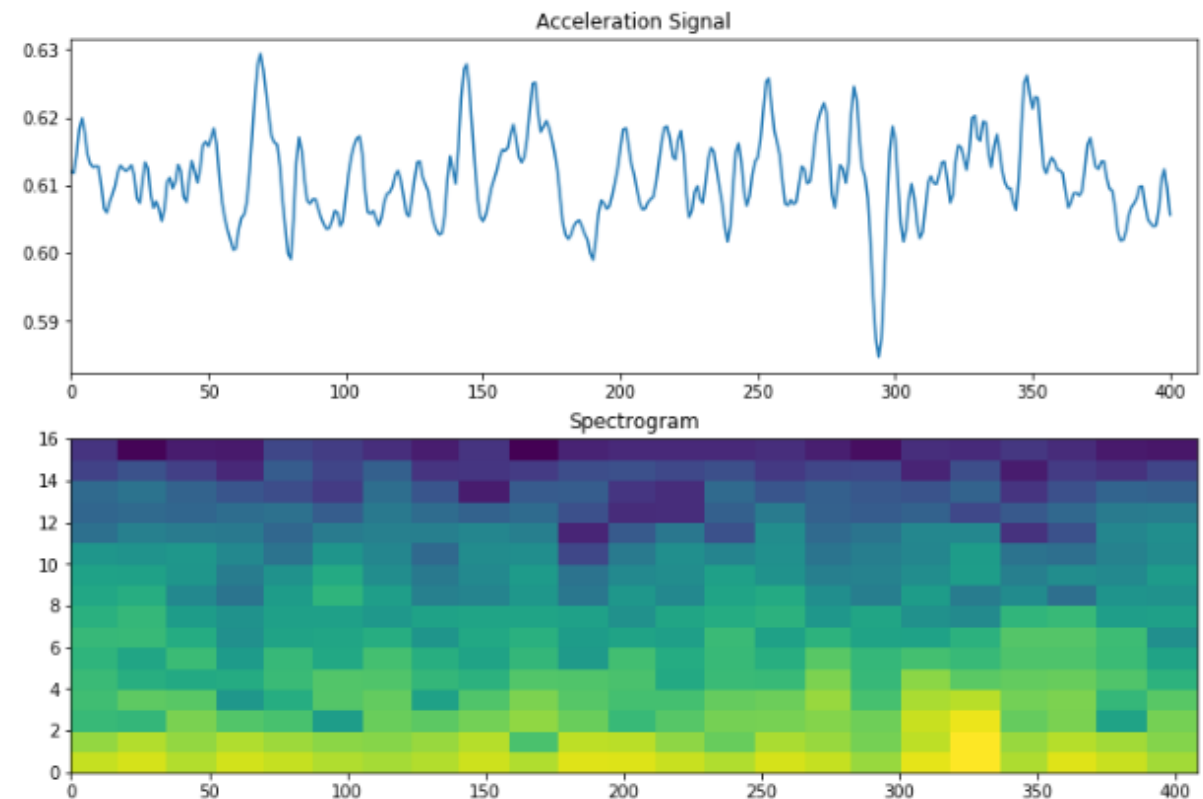
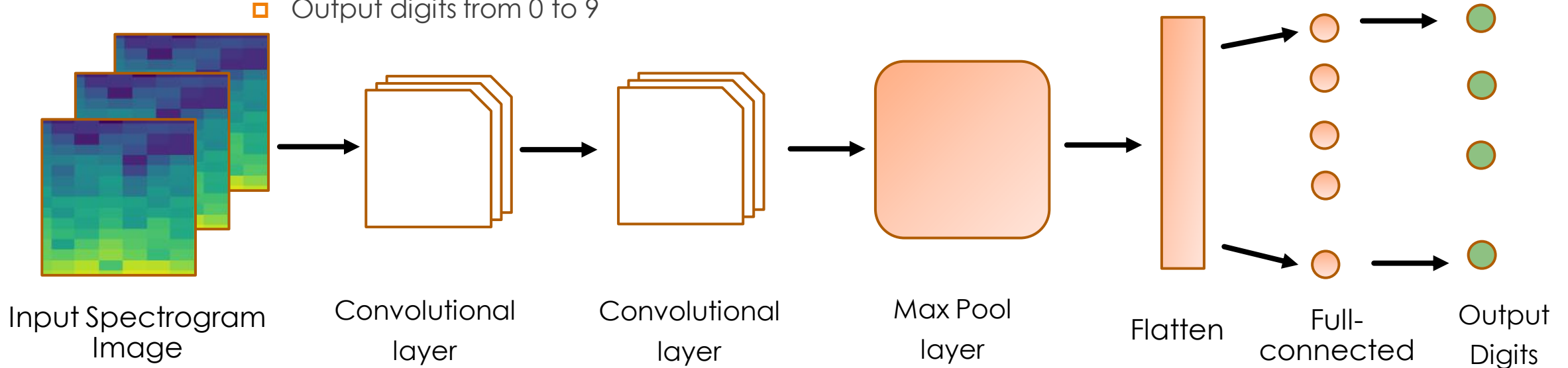


Fig. The spectrogram of one audio sample (X-axis)

# Training Process

## ► Convolutional Neural Network (CNN)

- Validate the assumptions: The feasibility of speech recognition based on the accelerometers.
- Input shape of data is  $24 \times 17 \times 3$
- Output digits from 0 to 9



# Recognition Result

## Reason for low accuracy

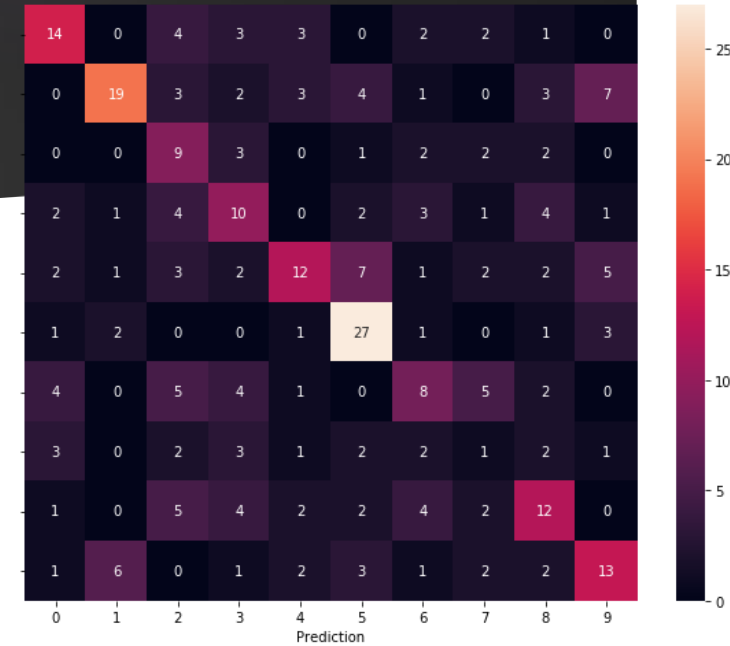
- Acoustic noise
- Model selection

Top 1 Acc	Top 3 Acc	Top 5 Acc
42%	87%	99%

Table. The accuracy of the model

Sources	Model	Purpose	Top1 Acc	Top 3 Acc	Top 5 Acc
Kaggle	CNN	Spoken digits classification	96%	-	-
NDSS	DenseNet	Accelerometer eavesdropping	78%	96%	99%

Table. Comparison with my results on the same dataset



# Future Study

## ► Attack Scenario:

- ▣ The victim makes a phone call and requests a password during the conversation

## ► Attack Process

- ▣ Hotword search
- ▣ Digits Recognition

### Hotword

Password

PIN

Security

Number

Credit

Card

Bank



Thank you!

Questions & Comments