# Acoustic Sensing for the Detection of Facial Expressions [Project -1 CS47007]

**Supervisor:** Prof. Sandip Chakraborty **Mentored By**: Miss. Pragma Kar

**Submitted By:** Avijit Mandal (18CS30010)

# Acoustic Sensing and Applications

- Acoustic sensing is the analysis of inaudible sound signals which is reflected from human parts.
- Recent studies had demonstrated technical feasibility and effectiveness of Acoustic Sensing. Examples,
  - Acoustic Sensing to Measure Heartbeats[6].
  - Measure Finger Movements[11].
  - Construct an Indoor Floor Map[1][9].
  - Eye Blinking Detection[5] for Safety in Driving[5].

# Project Idea

1. The project aims at analyzing smartphone-generated acoustic signals for facial expression detection.
2. We aim at developing a smartphone application that can produce audio signals of different frequency ranges and record the reflected multipath signal.
3. Since ultrasound signals with higher frequency range do not interfere with the audible ranges, we use high-frequency audio signals of 20 kHz and capture the unfiltered reflected signal for analysis.

# Problem Motivation

- Existing Challenges of Image-Based Detection
  - **Facial Occlusion**, eg Sunglasses, Scarves, masks.
  - Enacted training **Datasets**, different from natural.
  - **Privacy Concerns** for keeping Video On all the time
- Advantages with Acoustic Sensing
  - **Ultrasound signals** are not affected by audible Signals.
  - Smartphones **support** generation high-frequency chirp Signals.

# Literature Survey

| No | Paper Title | Author | Year | Problem Statement | Solution Proposed | Results |
|---|---|---|---|---|---|---|
| 1 | Acousticcardiogram: Monitoring Heartbeats using Acoustic Signals on Smart Devices | Kun Qian, Chenshu Wu, Fu Xiao, Yue Zheng, Yi Zhang | 2018 | Measuring heartbeats without direct contact. | Modelling Chest Movement with phase of Frequency Modulated Continuous Wave | average error of only 0.6 beats per minute |
| 2 | BackDoor: Making Microphones Hear Inaudible Sounds | Nirupam Roy, Haitham Hassanieh, Romit Roy Choudhury | 2017 | Recording ultrasound signals without excessive hardware and keeping inaudible to human ear | The core idea lies in exploiting non-linearities in microphone hardware. | Played 100 different sounds and verified its inaudibility with 7 different individual. |

# Literature Survey

| No | Paper Title | Author | Year | Problem Statement | Solution Proposed | Results |
|---|---|---|---|---|---|---|
| 3 | BatMapper: Acoustic Sensing Based Indoor Floor Plan Construction Using Smartphones | Bing Zhou, Mohammed Elbadry, Ruipeng Gao, Fan Ye | 2017 | Developing indoor location-based services (LBS) based on mobile sensing data. | Developed sound signals that can be detected by heterogeneous microphones on commodity smartphones followed by acoustic sensing to detect objects. | BatMapper is accurate to a distance of 1–2 cm in ranges up to around 4m |
| 4 | FingerIO: Using Active Sonar for Fine-Grained Finger Tracking | Rajalakshmi Nandakumar, Vikram Iyer, Desney Tan, Shyamnath Gollakota | 2016 | Tracking finger movements using acoustic sensing with possible occlusion. | Used Orthogonal Frequency Division Multiplexing (OFDM). Converted smartphones into an active sonar system that achieves sub-centimeter level accuracy. | The evaluation shows that FingerIO can achieve 2-D finger tracking with an average accuracy of 8 mm |

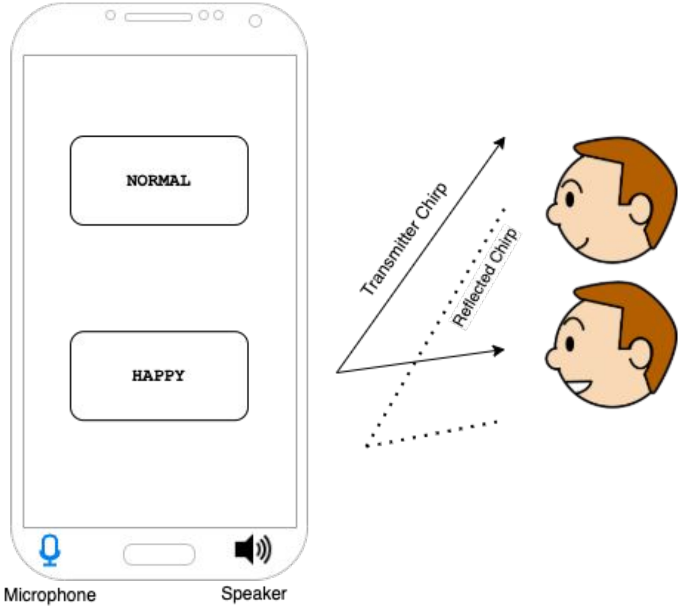
# Literature Survey

| No | Paper Title | Author | Year | Problem Statement | Solution Proposed | Results |
|---|---|---|---|---|---|---|
| 5 | BlinkListener: “Listen” to Your Eye Blink Using Your Smartphone | Jialin Liu, Dong Li, Lei Wang, Jie Xiong | 2021 | Camera based blinking detections suffers from privacy concerns, strict lighting conditions, and line-of-sight requirement. | Authors first quantitatively model the relationship between signal variation and the subtle movements caused by eye blink and interference. | on experimental results, BlinkListener provides robust performance with a 95% detection accuracy. |

# Proposed Model

## Overview

1. Current implementations supports generation of audio signals of all freqs
2. Record the reflected sound from surroundings
3. Analysis of the reflected sound – Anger, Happy etc. using Deep Neural Network



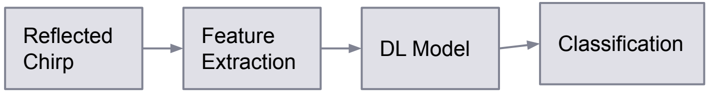
# Proposed Model

## Architecture

1. Generation of Chirp Signal
2. Feature Extraction – Using JLibrosa Library
3. Deep Learning Model for classifications of facial expression.

$$x(t) = \sin\left[\phi_0 + 2\pi(\frac{ct^2}{2} + f_0 t)\right]$$

Equation for generate chirp signal

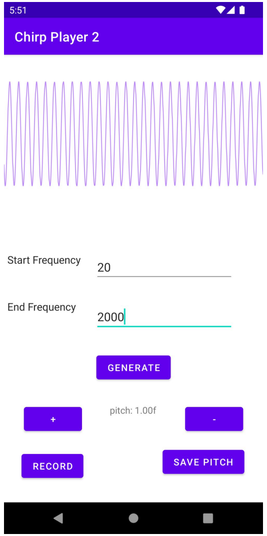



The proposed Architecture

# Implementation

## Proposed Model Implementation

1. Chirp Sound Generation – Using AudioTrack API
2. Pitch Controller and Visualization
3. Recording of Reflected Sound
4. Audio Features Extractions using JLibrosa Library.
5. Storage of Audio Features in SQLite Database

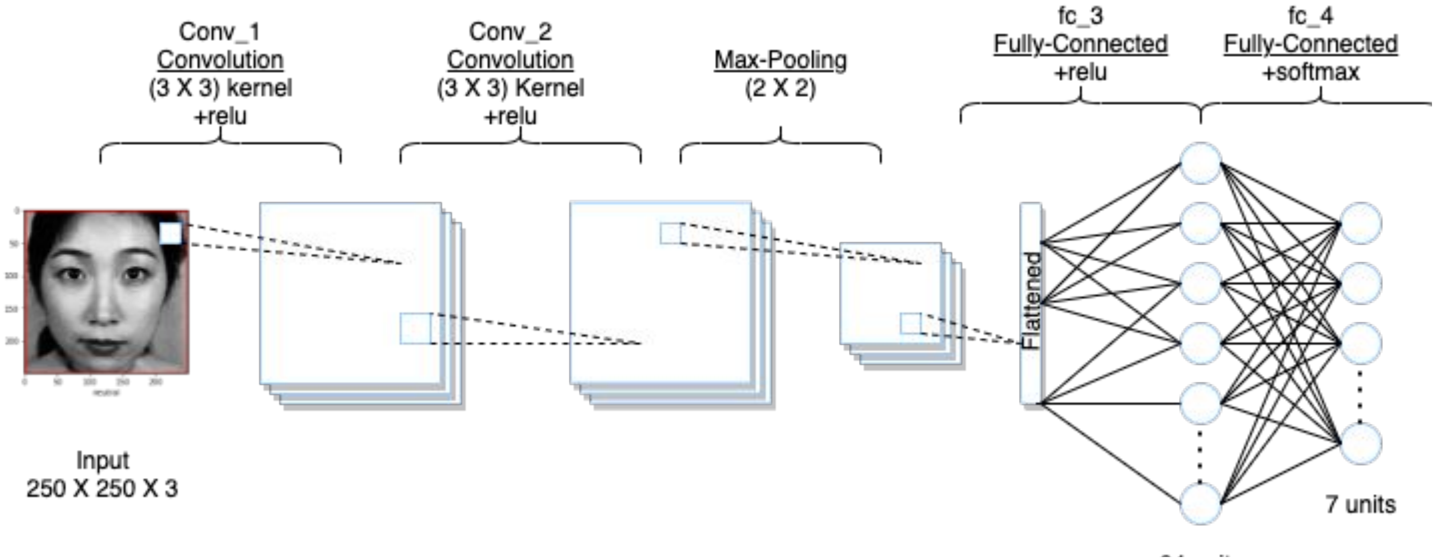



Overview of the Application

# Implementation: Baseline Model

1. Created a combined dataset using JAFFE[7](Japanese Female Facial Expression) and CK-Plus[10](Extended Cohen Kanade)
2. Detected the face for each of the images using OpenCV pre-trained haarcascade_frontalz_default_classifier
3. Cropped images around their face boundary and resize every image to exact size of 250X250
4. Trained Convolutional Neural Network using 80% processed dataset and tested it using remaining 20% of the dataset.

# Implementation: Baseline Model



# Future Work

1. The phasor diagram, based on the reflected Chirps will be analyzed, to investigate whether it can encode the expression change instances
2. A pilot study will be conducted to analyze the nature of data in real world.
3. The appropriate Deep Learning model will be developed for learning the audio features for prediction.
4. A thorough Study will be conducted for testing the accuracy of the model.

# Conclusion

1. The proposed model aims at addressing the challenges problem of facial expression detection and classification through acoustic sensing.
2. The model uses ultrasound signals transmitted and captured by commodity smartphones, without the use of any external sensors.
3. The model us currently under development and is showing promising results in terms of feasibility
4. The functionality of this application can be extended to attention estimation of learners during classes, by accessing the rate of expression change, HCI for physically disabled users, Game based learning for learners with cognitive challenges and so on

# Bibliography


1. Ruipeng Gao Fan Ye Bing Zhou, Mohammed Elbadry. Batmapper: Acoustic sensing based indoor floor plan construction using smartphones. In *MobiSys '17: Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services, 2017.*
2. Android Developer Documentation. Audiorecord, android developers. 2018.
3. Android Developer Documentation. AudioTrack, android developers. 2018.
4. Android Developer Documentation. AudioTrack, android developers. 2018.
5. Lei Wang Jie Xiong Jialin Liu, Dong Li. Blinklistener: "listen" to your eye blink using your smartphone. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2021.*
6. Fu Xiao Yue Zheng Yi Zhang Zheng Yang Yunhao Liu Kun Qian, Chenshu Wu. Acousticcardiogram: Monitoring heartbeats using acoustic signals on smart devices. In *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications, 2018.*
7. Miyuki, Gyoba Jiro Lyons, Michael, Kamachi. The japanese female facial expression (JAFFE) dataset. 2018.

# Bibliography


8. Romit Roy Choudhury Nirupam Roy, Haitham Hassanieh. Backdoor: Making microphones hear inaudible sounds. In *MobiSys '17: Proceedings of the Bibliography 23 15th Annual International Conference on Mobile Systems, Applications, and Services, 2017.*
9. Armin B. Cremers Pascal Bihler, Paul Imhoff. Smartguide – a smartphone museum guide with ultrasound control. In *The 8th International Conference on Mobile Web Information Systems (MobiWIS), 2011.*
10. Takeo Kanade Jason Saragih Zara Ambadar Patrick Lucey, Jeffrey F. Cohn. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In Disney Research, 4615 Forbes Ave, Pittsburgh, PA 15213, 2010.
11. Desney Tan Shyamnath Gollakota Rajalakshmi Nandakumar, Vikram Iyer. Fingerio: Using active sonar for fine-grained finger tracking. In *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 2016.*

# Thankyou!