# INDOOR LOCATION SENSING USING SMARTPHONE ACOUSTIC SYSTEM


**FILIP BILIŃSKI** F.K.Bilinski@student.tudelft.nl

TU DELFT FACULTY OF ELECTRICAL ENGINEERING, MATHEMATICS, AND COMPUTER SCIENCE

Supervisor: Qun Song


## INTRODUCTION

Localization services are widely used in mobile app development. While outdoor localization already provides precise localization with GPS, using GPS indoors comes with challenges[1] that result in accuracy to low for room recognition. There are however real word applications where indoor localization would prove helpful.

**Use cases:**
- Indoor way-finding
- Hospital patient localization
- Automated museum tour guides
- Smart-building automatization

## BACKGROUND

Since there are possible applications for an indoor localization service over the past years there were multiple proposed solutions that make use of a range of different technologies and algorithms.

**Multiple classes of solutions:**
- Acoustic, WIFI based, mixed, etc.
- Infrastructure-free / Infrastructure-dependant
- Passive sensing[2] / Active sensing[3]

## RESEARCH QUESTION

Can the robustness of the system against music containing environment be improved?

**Sub-questions:**
- How is the system affected by the presence of music in the environment?
- Can deep learning methods be used to improve robustness against music in the environment?

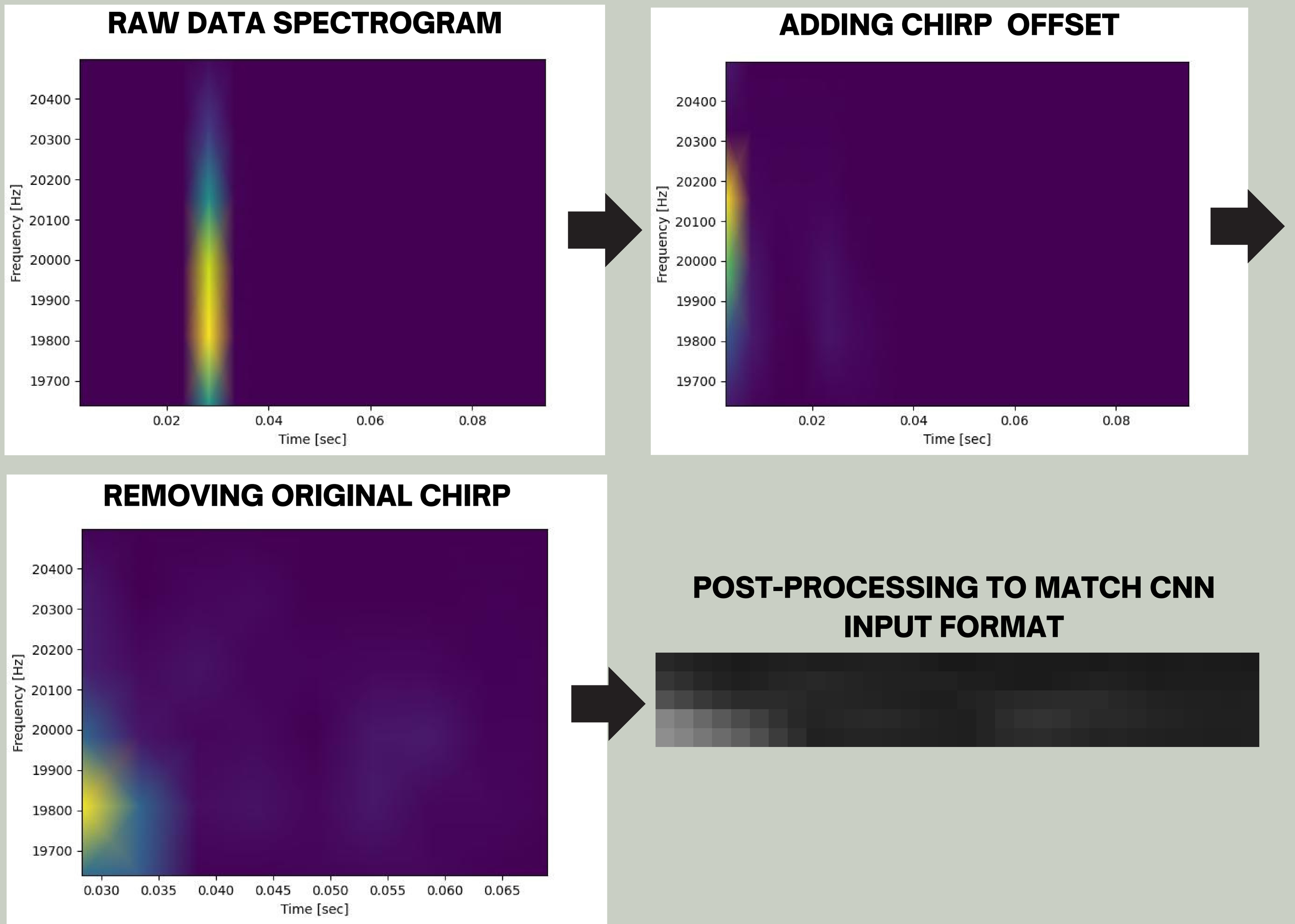


FIG 1: SPECTROGRAM CREATION FOR CNN INPUT STEP BY STEP

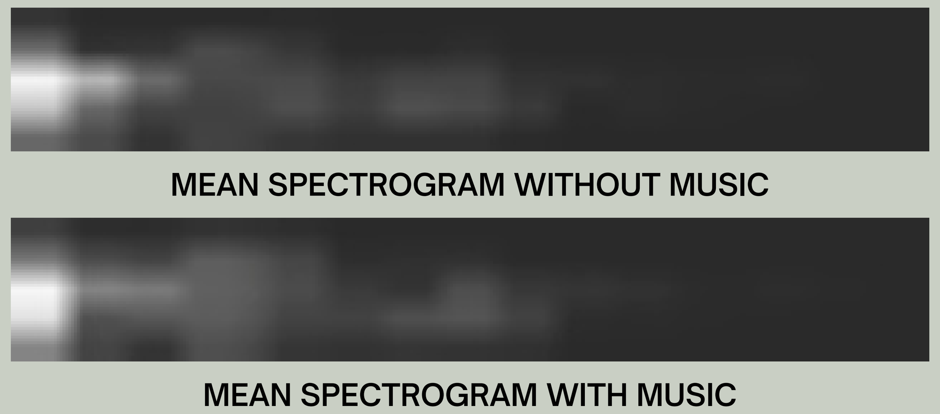


FIG 3: MEAN SPECTROGRAMS SHOWING HOW MUSIC IMPACTS THE SYSTEM

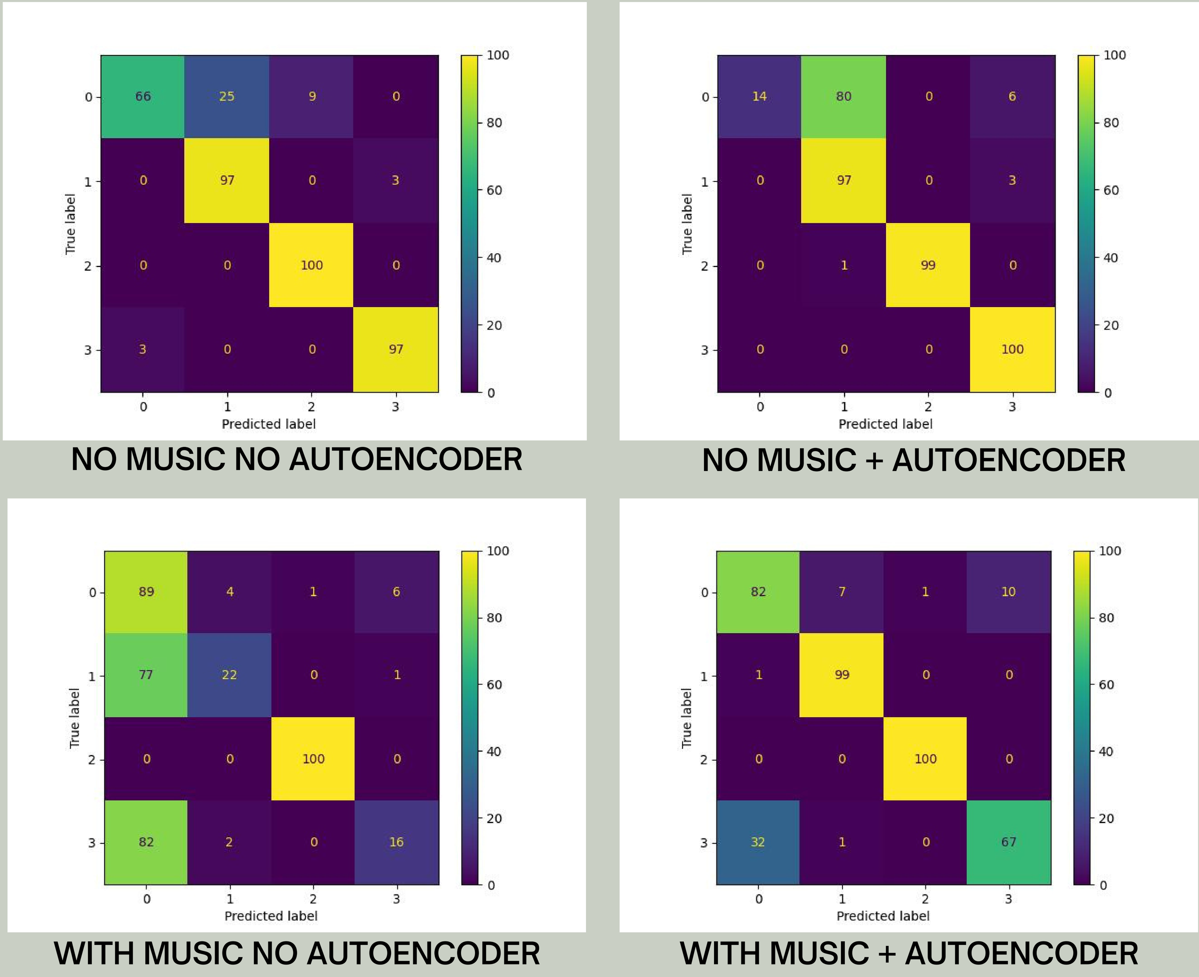


FIG 4:EXPERIMENT RESULTS

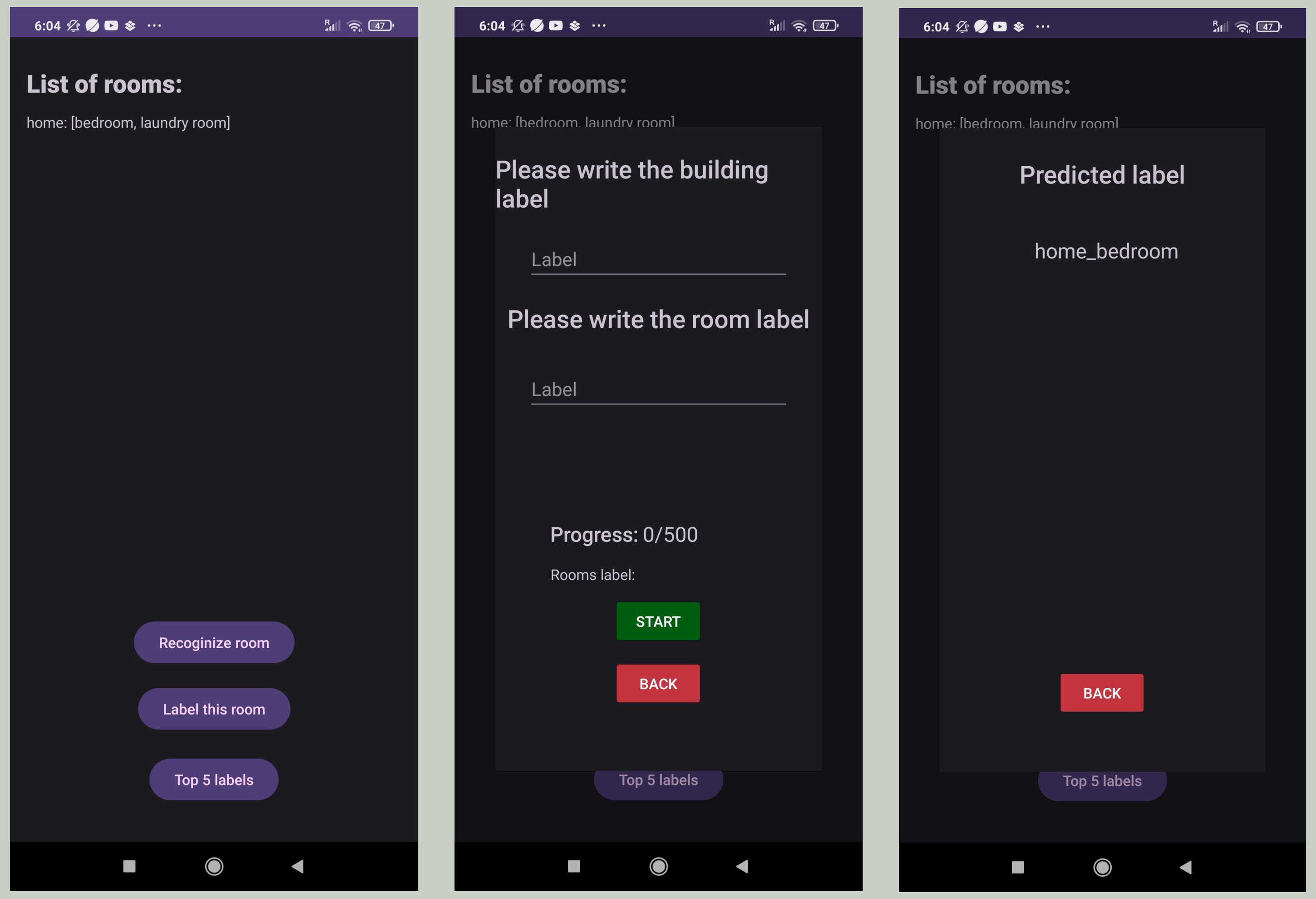


FIG 2: MOBILE APPLICATION DESIGN

## METHODOLOGY

- Implementing proof of concept application - Fig 1, 2
- Gathering required datasets
- Evaluating the impact of music on the system - Fig 3
- Designing the autoencoder
- Evaluating the performance of the system with and without music ith and without autoencoder

## AUTOENCODER DESIGN

**Encoder:**
- Convolutional layer - 16 5x5 filter
- Max pooling layer - 2x2 filter
- Convolutional layer - 16 5x5 filter
- Max pooling layer - 2x2 filter

**Decoder:**
- Transposed convolutional layer - 16 5x5 filters
- Transposed convolutional layer - 16 5x5 filters
- Convolutional layer - 1 5x5 filter

## RESULTS

| | No Autoencoder | | With Autoencoder | |
|---|---|---|---|---|
| | Average | std dev | Average | std dev |
| No music | 0.91 | 0.07 | 0.81 | 0.05 |
| Music | 0.70 | 0.10 | 0.76 | 0.09 |
| Mixed | 0.81 | 0.06 | 0.79 | 0.04 |

- Adding the autoencoder introduces a tradeoff between accuracy in music-containing environments and quiet environments - Fig 4
- Accuracy was improved in music-containing environments:
  - Lower accuracy drop
  - Higher accuracy
- Robustness of the system stayed relatively the same, represented by mixed dataset results

## REFERENCES:


[1] G. DEDES AND A.G. DEMPSTER. INDOOR GPS POSITIONING CHALLENGES AND OPPORTUNITIES. IN VTC-2005-FALL. 2005 IEEE 62ND VEHICULAR TECHNOLOGY CONFERENCE, 2005., VOLUME 1, PAGES 412–415, 2005

[2] STEPHEN P. TARZIA, PETER A. DINDA, ROBERT P. DICK, AND GOKHAN MEMIK. INDOOR LOCALIZATION WITHOUT INFRASTRUCTURE USING THE ACOUSTIC BACKGROUND SPECTRUM. IN MOBISYS 11: PROCEEDINGS OF THE 9TH INTERNATIONAL CONFERENCE ON MOBILE SYSTEMS, APPLICATIONS, AND SERVICES, PAGES 155–168, 2011.

[3] QUN SONG, CHAOJIE GU, AND RUI TAN. DEEP ROOM RECOGNITION USING INAUDIBLE ECHOS, 2018.