THE ANDREW & ERNA VITERBI FACULTY OF ELECTRICAL ENGINEERING

Work in Progress

SIPL
Signal and Image Processing Lab

TECHNION Israel Institute of Technology

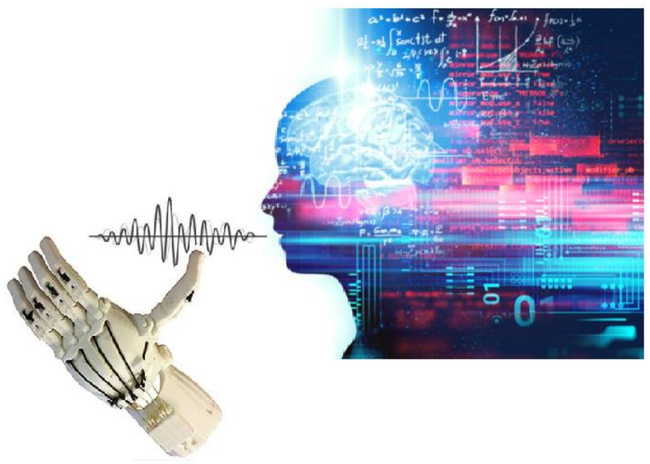
# User Specific Speech Recognition For Controlling a 3D Printed Prosthetic Hand


**Noa Tykochinsky and Itay Wengrowicz, Supervised by Shunit Polinsky**

**In collaboration with**


## Introduction

- Most of the current prosthetics available require muscle movement
- We would like to improve the user's everyday life by applying voice control to the prosthetic hand



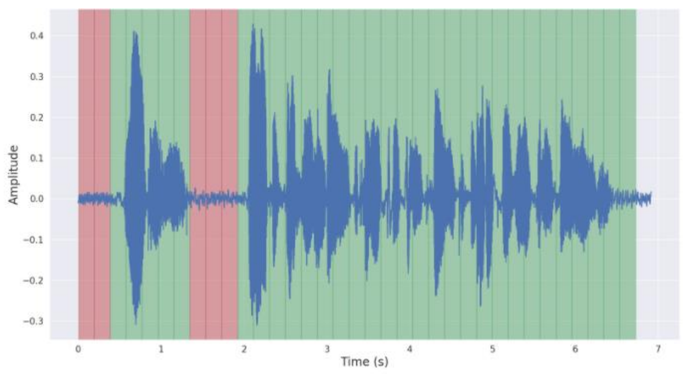
## Goals

- Algorithm that verifies the user's specific voice
- Recognize voice commands
- Send commands to prosthetic hand accordingly
- Work in real time and real-life environments

## Challenges

- Efficient algorithm for fast run time
- Identify speaker and commands in non ideal environments
- Israeli accent affect the speech to text results

## Voice Activity Detection

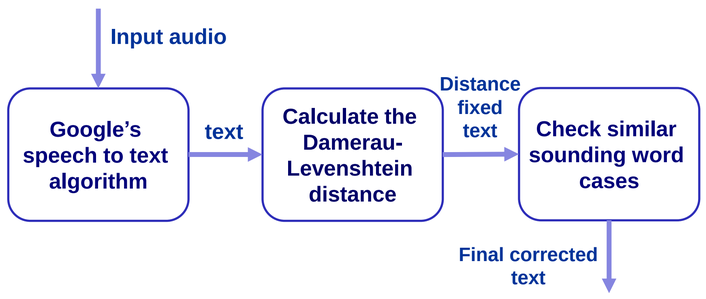



Voiced and quiet audio segments

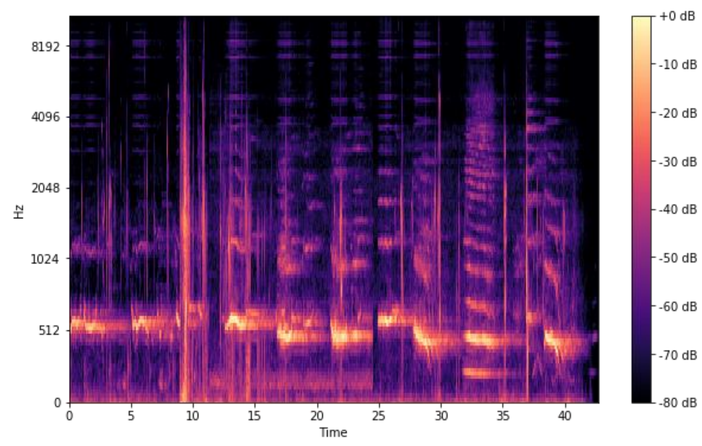
- VAD classifies audio data as voiced or unvoiced
- Detect if there is audio input the algorithm should start processing
- We use VAD that Google developed, which is:
  - Fast
  - Modern
  - Free

## Word Detection

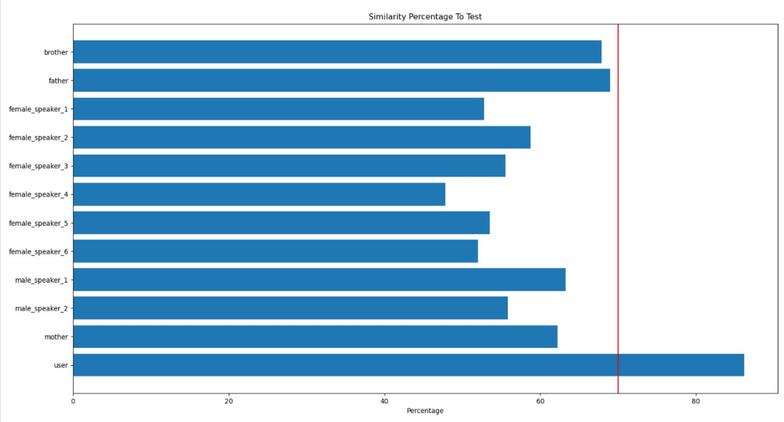
- Google's free Speech Recognition algorithm transcribes input audio to text
  - Uses a pre-trained neural network
  - Able to detect words with slight accent
- Calculate DL distance between text and known words we are looking for
  - If the distance < 0.5, receive closest word
- Similar sounding word bank for our keywords




## Speaker Verification




- Feature extraction using Mel spectrogram
  - Has the best performance for training neural networks
- Generalized end to end loss for speaker verification.
  - Pretrained model
  - Accuracy Threshold is set at 0.7
  - Result dependent on other speakers: we added data and finetuned the model

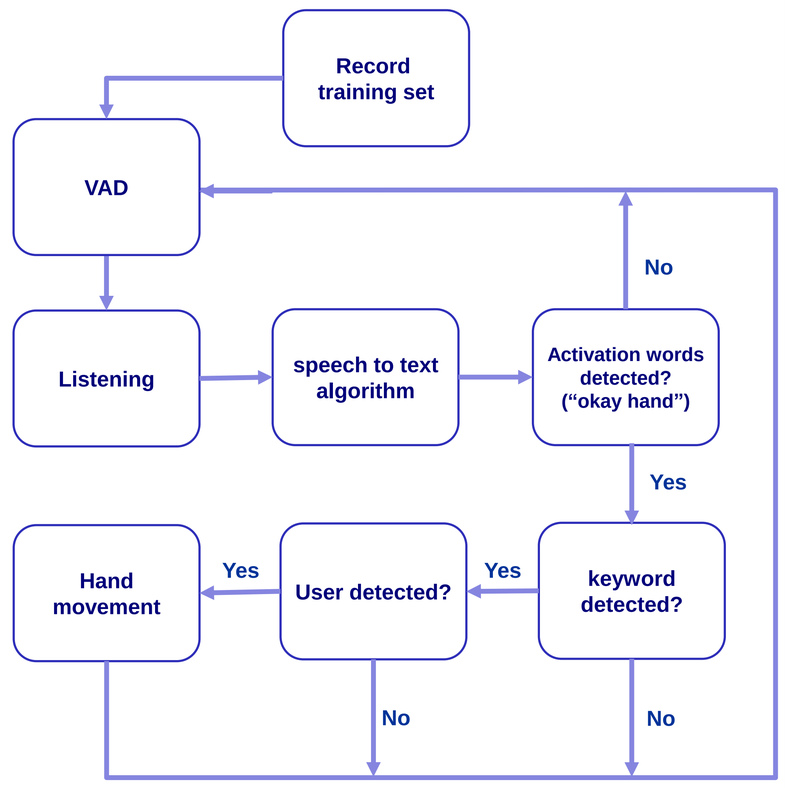



Verification accuracy compared to multiple speakers

- User will be recognized if he has the highest accuracy rate and if Rate > Threshold

## Complete Solution

- The algorithm's diagram:

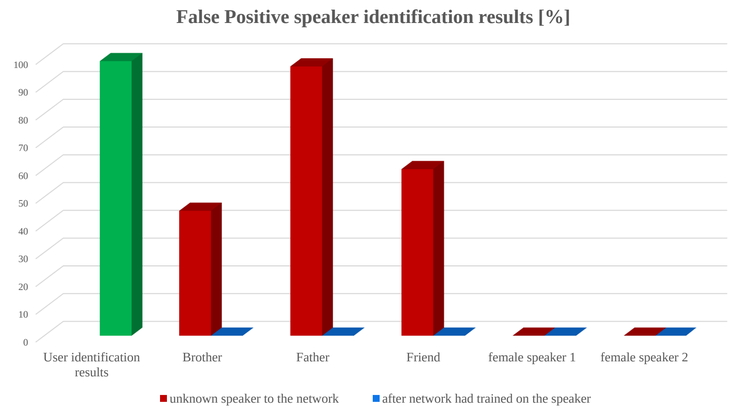



project block diagram

- After activated, for 10 seconds- only a keyword will be needed to activate the hand

## Results

- Percentages out of N=40 tests




## Conclusions

- Successful speaker verification in quiet environment
- 1.5 seconds process time
- Low False Positive rate
- Tradeoff between false-positive results and user identification

June 2021