

Dōkō

General Usage Robot



General
Utility
Neuro-network
Driven
Advanced
Mobile-Robot

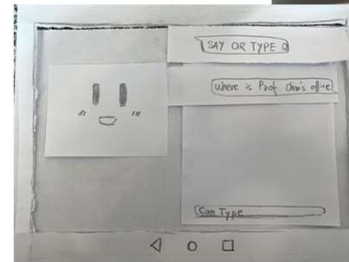
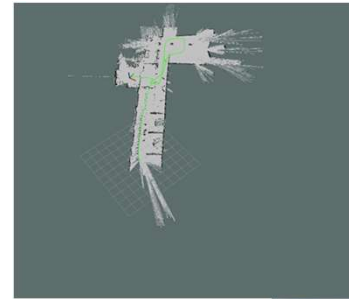
E. Tan, X. Wang, J. Zhong, L. Zhu
Advisor: Prof M. Chuah
Lehigh University

In collaboration with



Overview -- What are we doing

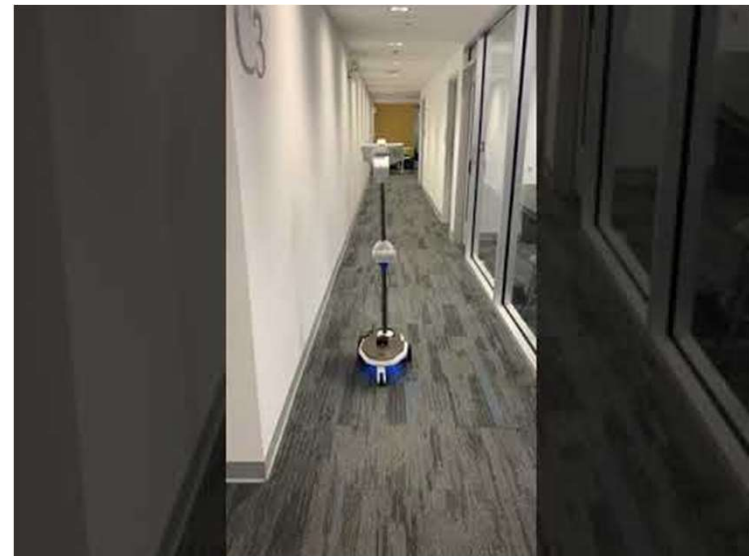
- **Navigation Module**- be able to navigate in an indoor environment and handle complex environment.
- **Vision Module**- recognize up to two hundred faces and be able to extract emotion features.
- **Voice Module**- cloud computing powered voice assistant and start interactive conversations based on the users reactions.



Images all rights reserved

Navigation: See that which is unseen.

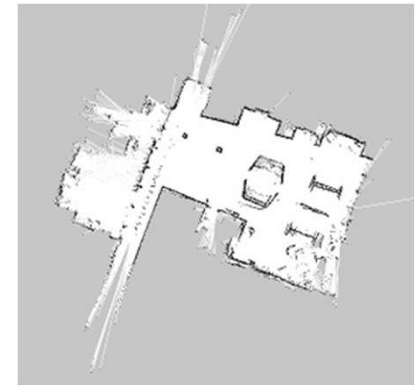
- A fine tuned sonar module
- Handles glass walls as to normal walls
- Low cost yet efficient



Navigation: See that which is unseen.

Some experiments we did:

- Refine map to correctly display glass walls
- Test different angles for the sonar to be set



Emotion Detection: Keras Deep Learning Model

- Categorize emotions into 6 classes: Angry, Disgust, Happy, Neutral, Sad, Surprised.
- Perform training and testing on FER-2013 [1] and RAVDESS [2] dataset.
- Model structure is built according to “Emotion Recognition using Deep Convolutional Network” paper by E. Correa etc. [3]
- We also tried to share the bottleneck blocks of the mobilenet structure for face recognition model with the emotion detection model in order to save computation, but did not get good performance. So we did not apply this approach.

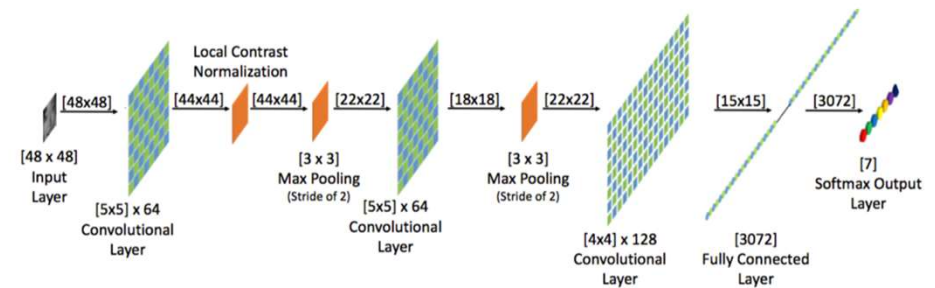
[1] <https://datarepository.wolframcloud.com/resources/FER-2013>

[2] <https://zenodo.org/record/1188976#.XfpkNdZKgWp>

[3] <https://github.com/atulapra/Emotion-detection/blob/master/ResearchPaper.pdf>

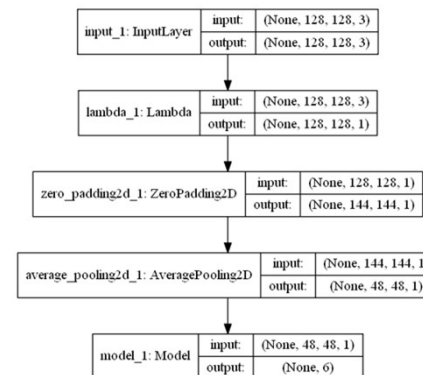
Emotion Detection: Model Structure

- The same image for face detection is passed to the emotion detection model, with size equals to 128x128x3.
- The image first goes through the wrapping model and then the base model.
- The output layer is changed from 7 to 6.
- Training parameter:
 - Optimizer: Adam
 - Learning rate: 0.001
 - Decay: 1e-5



Base Model

<https://github.com/atulapra/Emotion-detection>



Wrapping Model

Emotion Detection: App Deployment

- Converted to TFLite model and deployed on the Doko Android app.
- demo



Vision

- Use emotion detection results to enhance conversation.



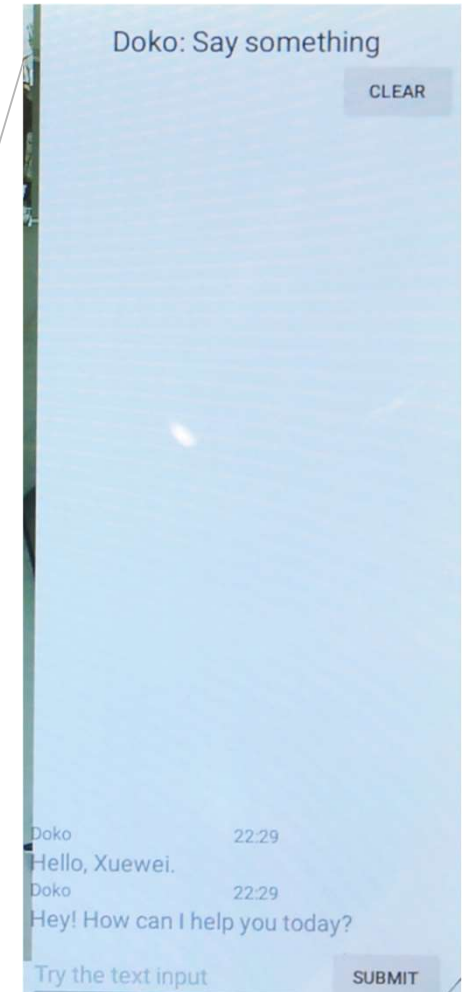
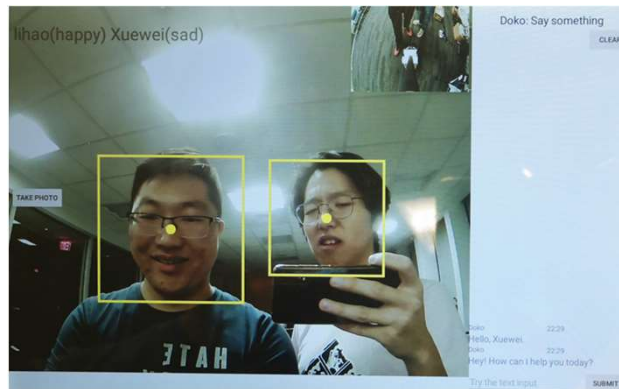
A sad face detected.

"Trigger sad" sent to the Dialogflow backend.

State of Dialogflow changes. In the following conversation Dialogflow will be aware of the emotion.

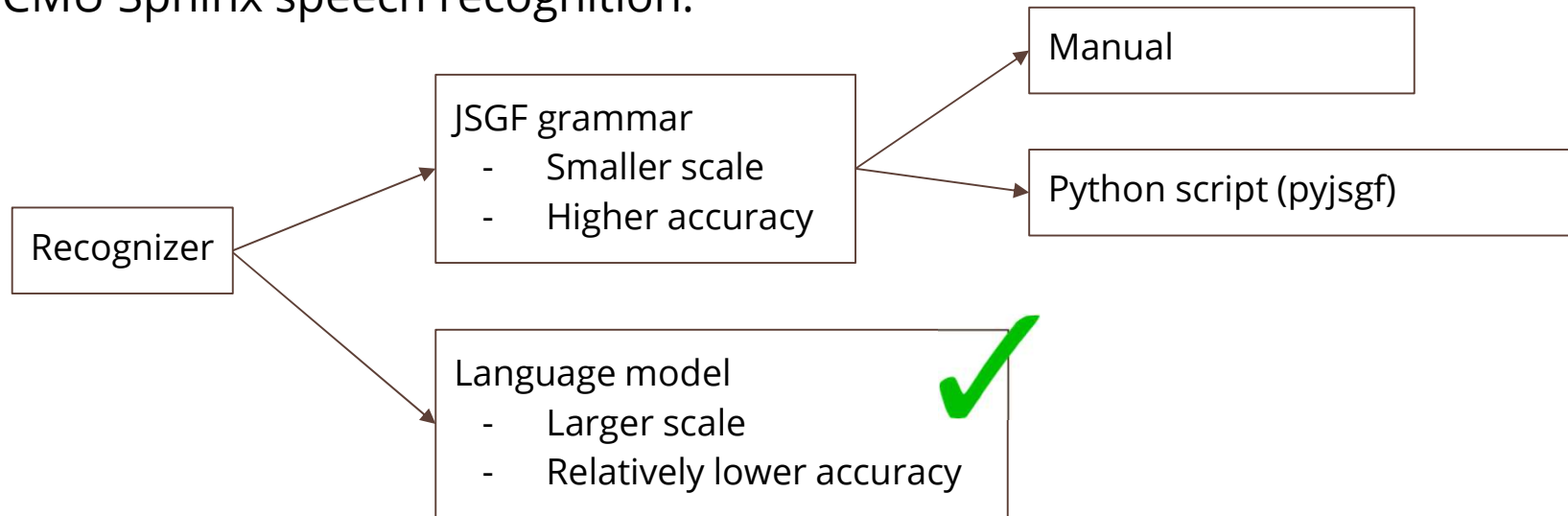
Android App

- UI Design change:
 - add the chat history record;
 - enable keyboard input.
- Dialogflow Android API update.
 - Possible to acquire non-textual results, such as websites and images.



Voice Assistant

CMU Sphinx speech recognition:



- Use **CMU Sphinx** for speech recognition, with language model compilation and a public domain pronunciation dictionary, to provide real-time speech to text transformation.

Voice Assistant

- Apply Google Dialogflow as conversational agent, supported by webhook for richer data responses.
- Connected with navigation --go to the places as user instructed.
- Contextual response based on different detected emotions.
 - Fetched medical data and built conversations that can respond to the medical request from the users.

Voice Assistant

- Here is an example of the use of voice assistant.

