

GestHome: User-defined Postures Detection for Smart Home

Christopher Albert Priatko Theodore Fabian Rudy

Supervisor: Prof. Michael R. Lyu
Computer Science and Engineering
The Chinese University of Hong Kong

May 22, 2023

Table of Contents

- 1 Introduction
- 2 Implementation
- 3 Evaluation
 - Stage 1
 - Stage 2
 - Back-end Framework
- 4 Demo
- 5 Conclusion
- 6 Q and A session

Introduction - Recap

From the last semester, with this project, we have managed to achieve these goals:

- Research different aspects of computer vision that is viable for the projects
- Compare the performance between those computer vision
- Implement a working prototype of action recognition with pre-determined gestures

Introduction - Objective

This semester, we aim to improve the prototype we have made in the last semester in these ways:

- Combine stage 1 (Face Recognition for Login) and stage 2 (Action Recognition) together to create a seamless smart home system
- Improve the performance of our system by applying more refined models

Table of Contents

- 1 Introduction
- 2 Implementation**
- 3 Evaluation
 - Stage 1
 - Stage 2
 - Back-end Framework
- 4 Demo
- 5 Conclusion
- 6 Q and A session

Literature Review



Face Detection

Dlib HOG, One-shot learning



Pose Estimation

PYSKL, ST-GCN++, etc



Backend Framework

Flask, MongoDB, Jinja, etc

Table of Contents

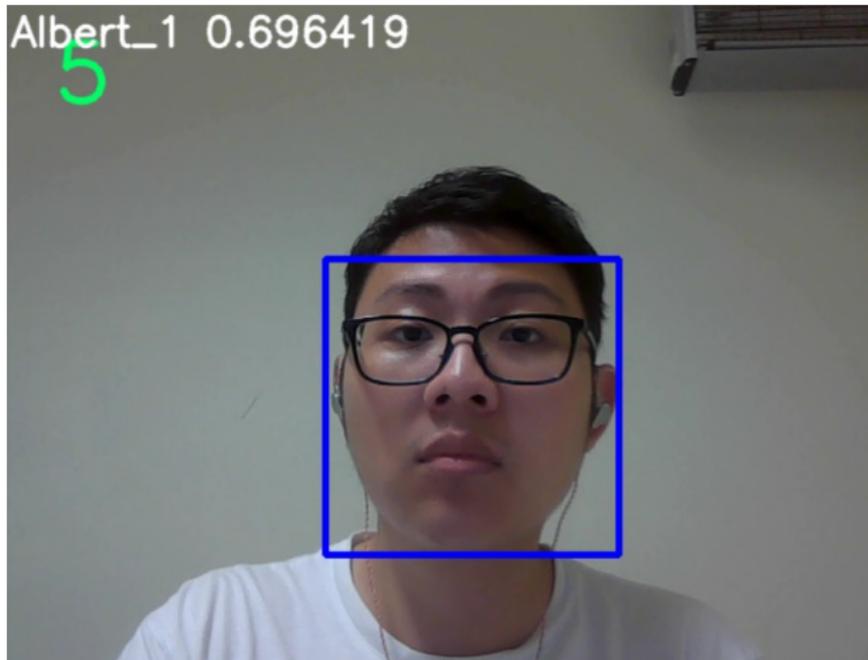
- 1 Introduction
- 2 Implementation
- 3 Evaluation**
 - Stage 1
 - Stage 2
 - Back-end Framework
- 4 Demo
- 5 Conclusion
- 6 Q and A session

Stage 1 - Overview



Face Detection

Stage 1 - Face Recognition

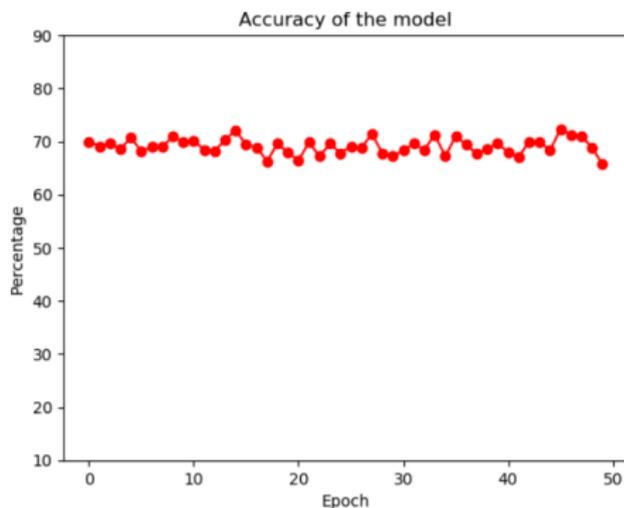
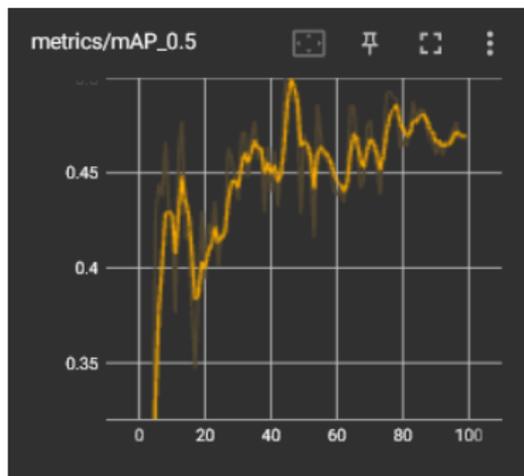


Face Recognition using Dlib with Siamese Neural Network

Face Recognition using Siamese Neural Network

- Realized that Dlib's face recognition accuracy is lower than expected
- Utilized Siamese Neural Network for face recognition
- Found that Siamese Neural Network is better than Dlib's face recognition

Accuracy Comparison



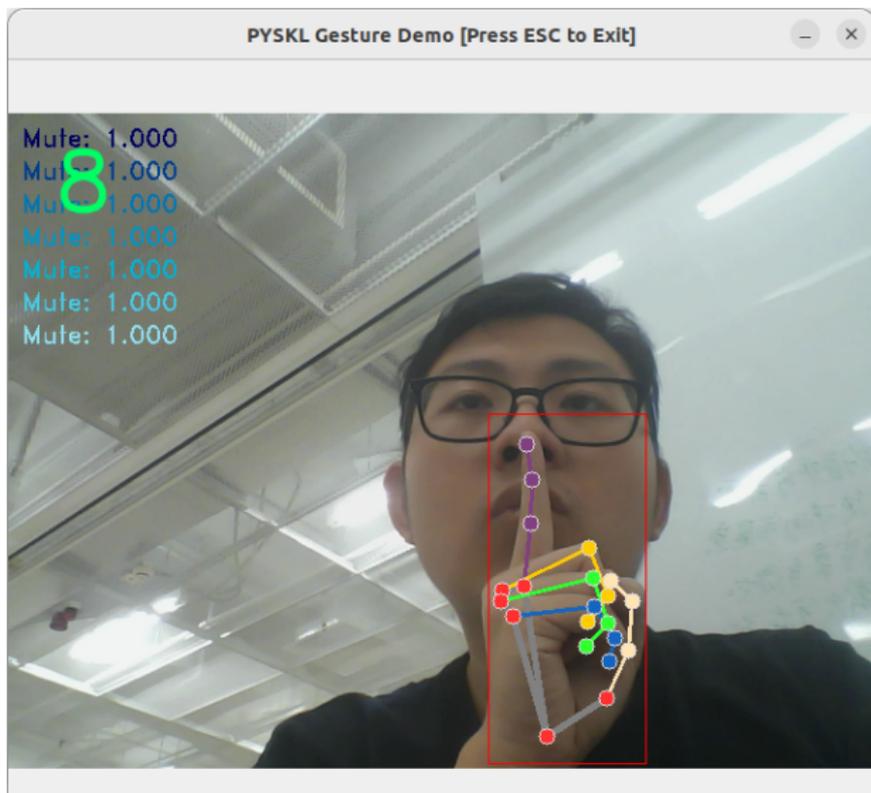
Accuracy of Dlib's Face Recognition (Left) vs Dlib with Siamese Neural Network (Right)

Stage 2 - Overview



Action Recognition

Stage 2 - Action recognition



Inference run using PYSKL

Stage 2 - Action recognition (cont.)

- We would like to use LSTM since we manage to make it work
- Problem encountered: LSTM does not provide satisfactory performance
- We wanted to do a few-shot recognition, which is not possible using LSTM

Candidate tools

- MotionBERT
- PYSKL (ST-GCN++) (selected)
- HyRSM

PYSKL (ST-GCN)

+

- (very recently) Implemented a lightweight model for CPU user
- State of the art in both 2D and 3D skeletal based action recognition
- Helpful documentation

-

- Is not supported in Windows (unless using WSL, but will need to experience some performance issue)
- CPU version only support one hand gesture

Action Mapping

- Besides using more sophisticated action recognition, we have decided to map the actions to a predetermine function
- These functions reflect the functions used in daily life (Weather, Air Quality, etc)
- There are 15 recognizable actions, which has been mapped to its' respective function

Stage 2 - Performance comparison

- Unlike LSTM where we build our own dataset, PYSKL (ST-GCN) is pretrained on HaGRID dataset
- 40 GB per move. Accuracy of detection is much better compared to what we have.

Action recognized



Figure: Gestures trained in PYSKL, according to HaGRID (Hand Gesture database)

Note: some movements are not able to be used, since Mediapipe is not able to determine inverse hand position

Back-end Framework

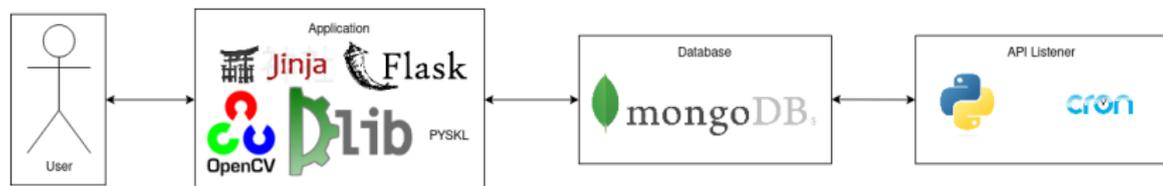


Back-end Framework

Background Reason

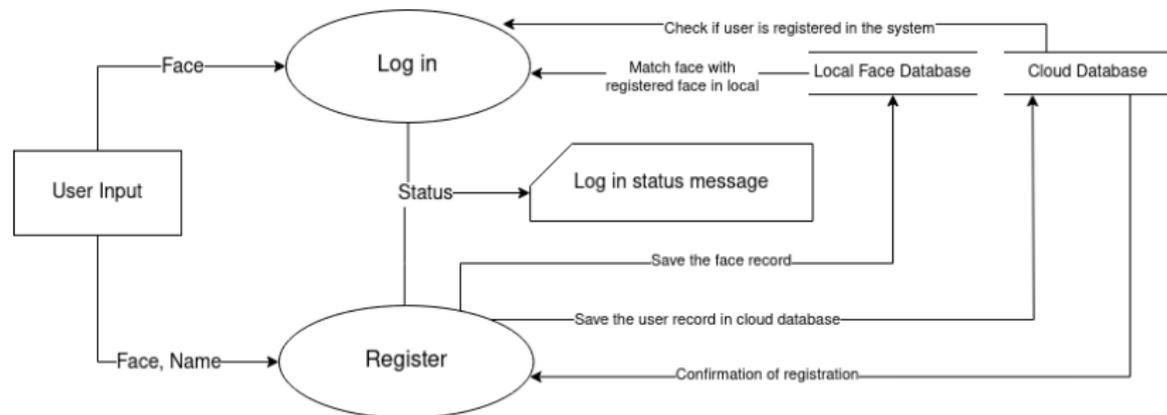
- We realize that we need to build a back-end framework to combine both stages
- We have decided to utilize Flask as the framework of our program
- We have also included Jinja and MongoDB to build our framework

The Framework

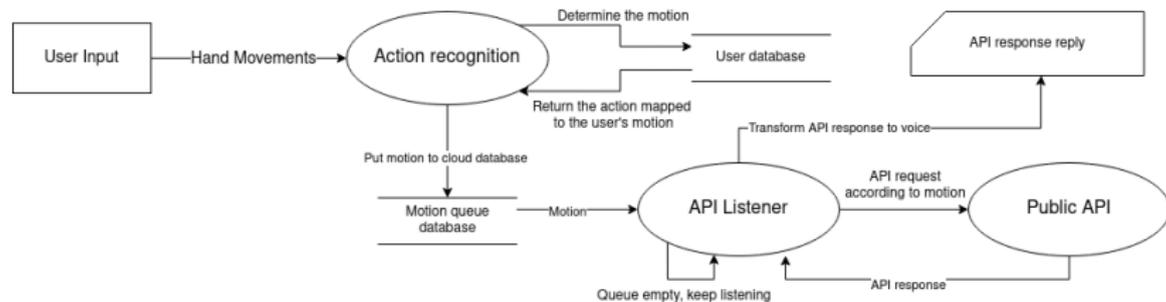


System Architecture of GestHome

Flow of the program - Login/Register



Flow of the program - Action recognition



Screenshots of the program

Welcome to GestHome!

Log In

Register

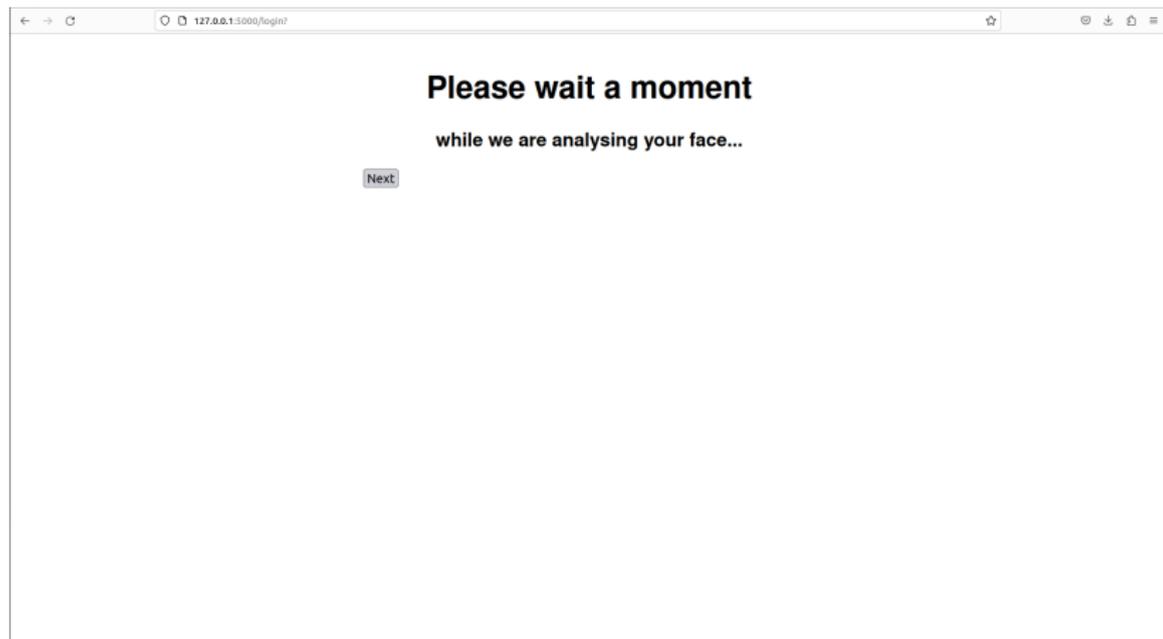
Welcome Page of GestHome

Screenshots of the program

Please enter your name:

Page for register a face

Screenshots of the program



Log in Page

Screenshots of the program

I am sorry, but I couldn't recognise you...

What would you like to do instead?

Log in again

Register

Back to Home

Page for Failed Log in

Screenshots of the program

Welcome Back, Albert!

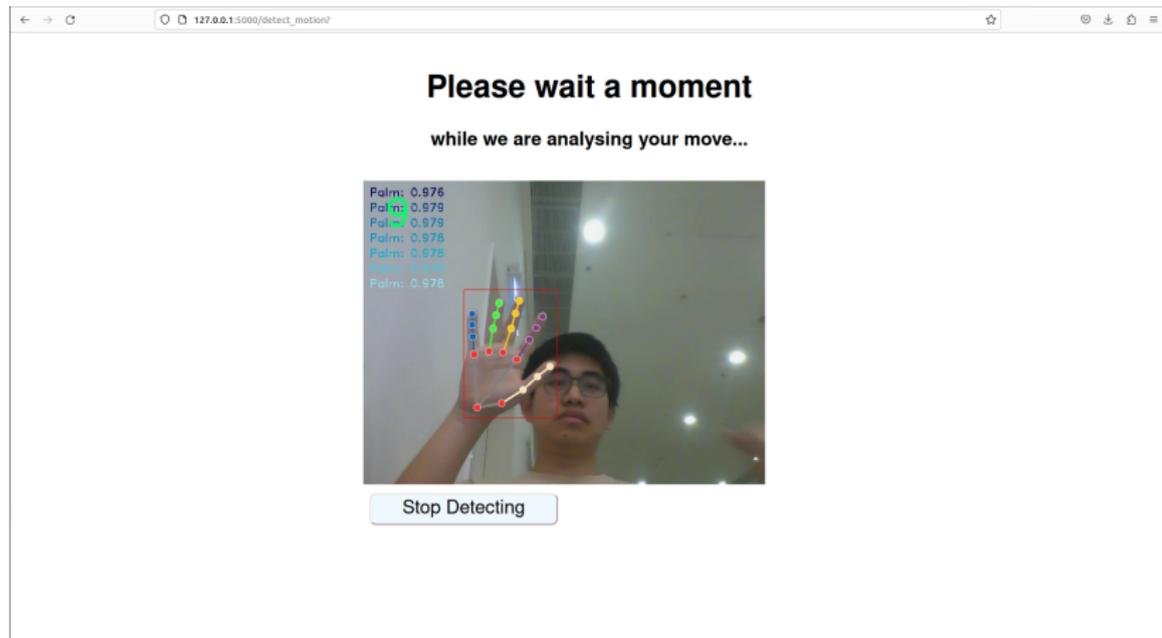
What would you like to do?

Motion detection

Log Out

Page for Successful Log in

Screenshots of the program



Page for Motion Detection

Table of Contents

- 1 Introduction
- 2 Implementation
- 3 Evaluation
 - Stage 1
 - Stage 2
 - Back-end Framework
- 4 Demo
- 5 Conclusion
- 6 Q and A session

Demo

Table of Contents

- 1 Introduction
- 2 Implementation
- 3 Evaluation
 - Stage 1
 - Stage 2
 - Back-end Framework
- 4 Demo
- 5 Conclusion**
- 6 Q and A session

Conclusion

During the past year, we have managed to:

- Researched and compared different computer vision related projects in terms of performance
- Implemented a working model of face recognition <https://www.hko.gov.hk> on and action recognition in a single, streamlined system
- Improved the performance of the models by implementing more sophisticated models

Table of Contents

- 1 Introduction
- 2 Implementation
- 3 Evaluation
 - Stage 1
 - Stage 2
 - Back-end Framework
- 4 Demo
- 5 Conclusion
- 6 Q and A session**

Q and A Session