A Robot Based on Jetson Nano and Machine Learning

—— AI Flower

Student Name: Tan Shuyao
Student ID: 16251423

Outline

Introduction

Percentage of people with mental disorders in selected countries (2009)

Objective

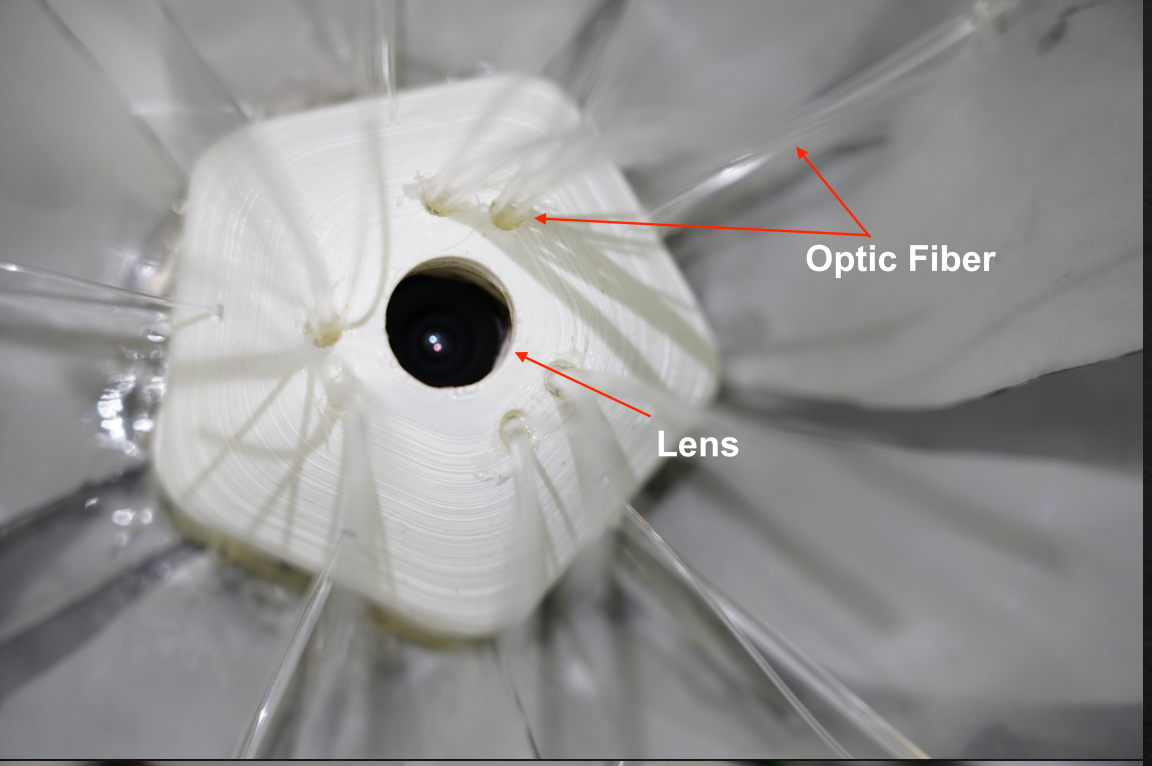
- Creating a product which could help people with stress
- The product should have following features:
  - Look like flower
  - Identify people’s expressions
  - Make responses

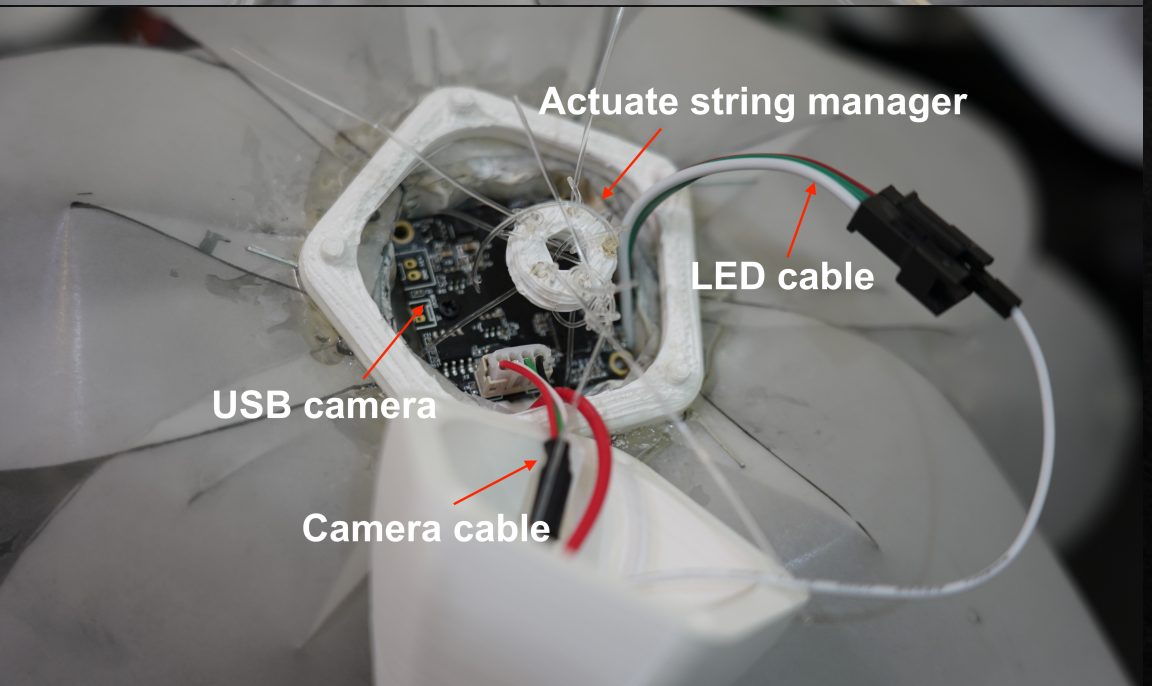
Product Overview

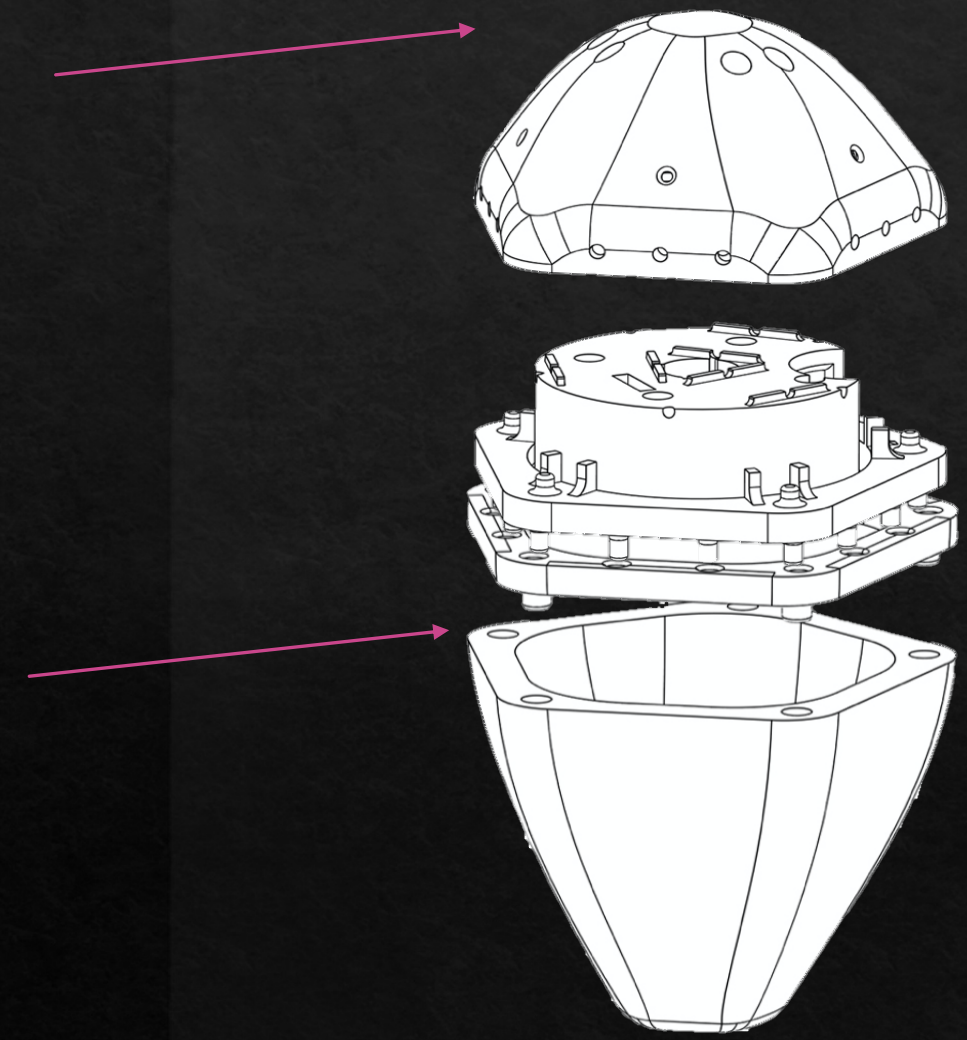
AI Flower

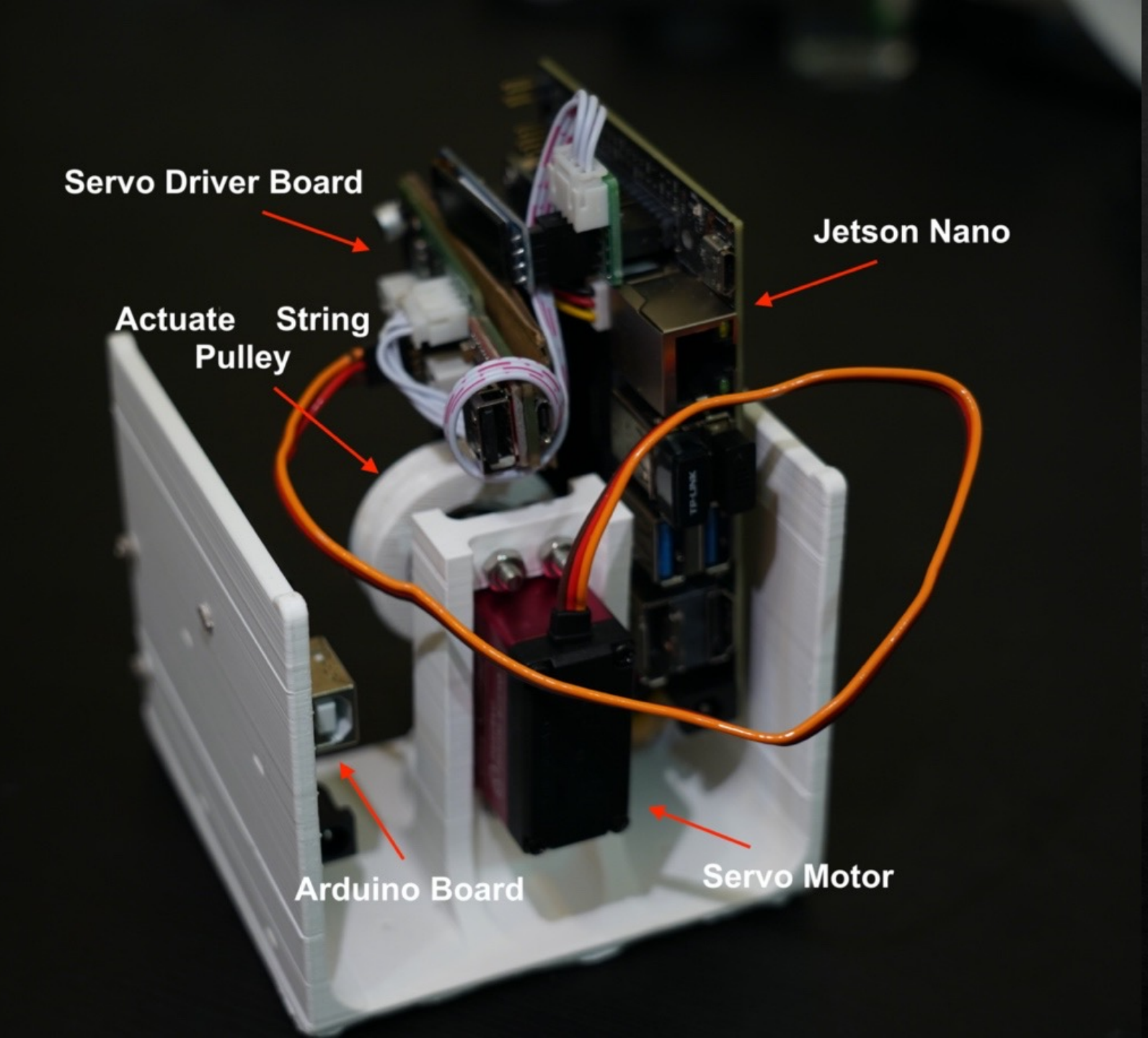
- Made of:
  - Jetson Nano
  - 2 Development Boards: Arduino board, servo driver board
  - USB Camera
  - LED strip WS2812B
  - Servo motor
  - Etc…

Product Structure - Head

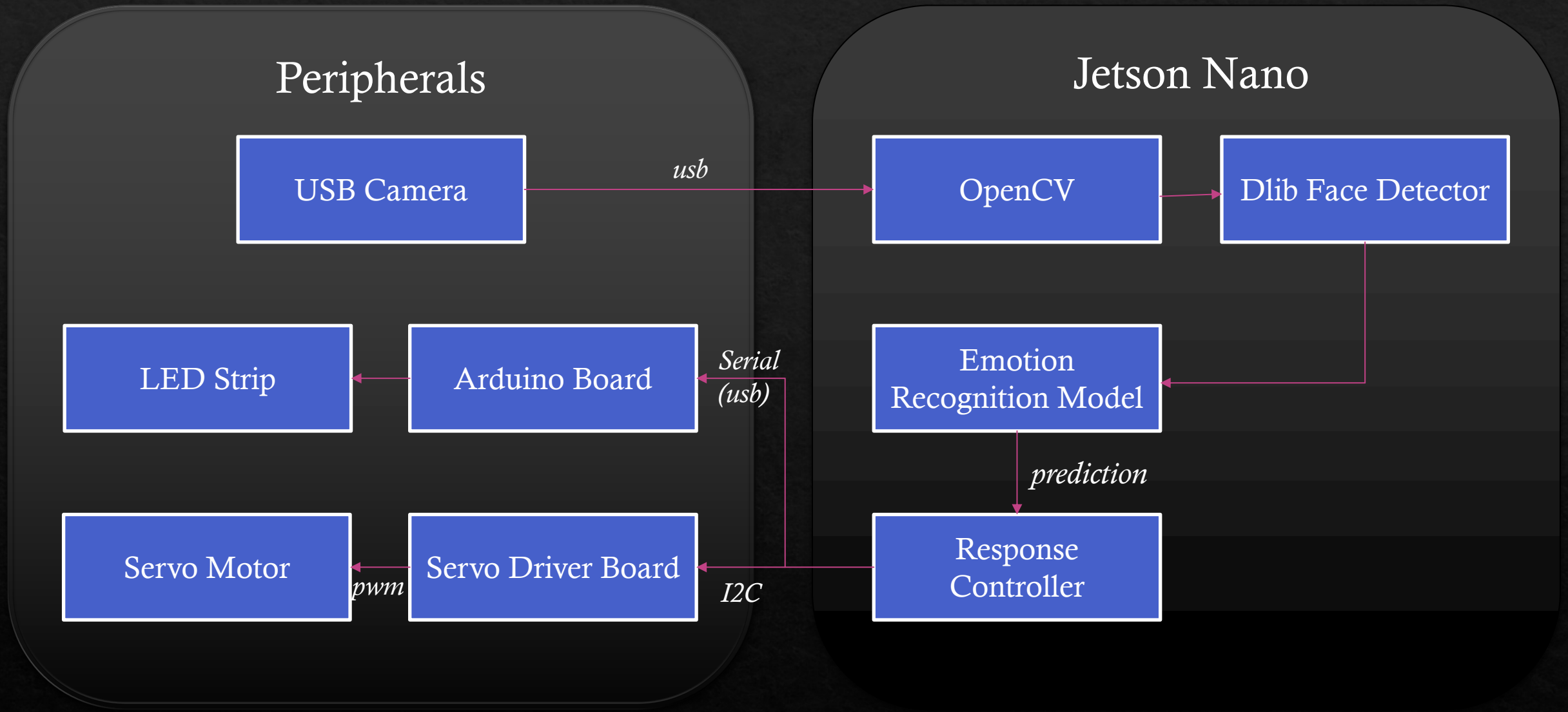






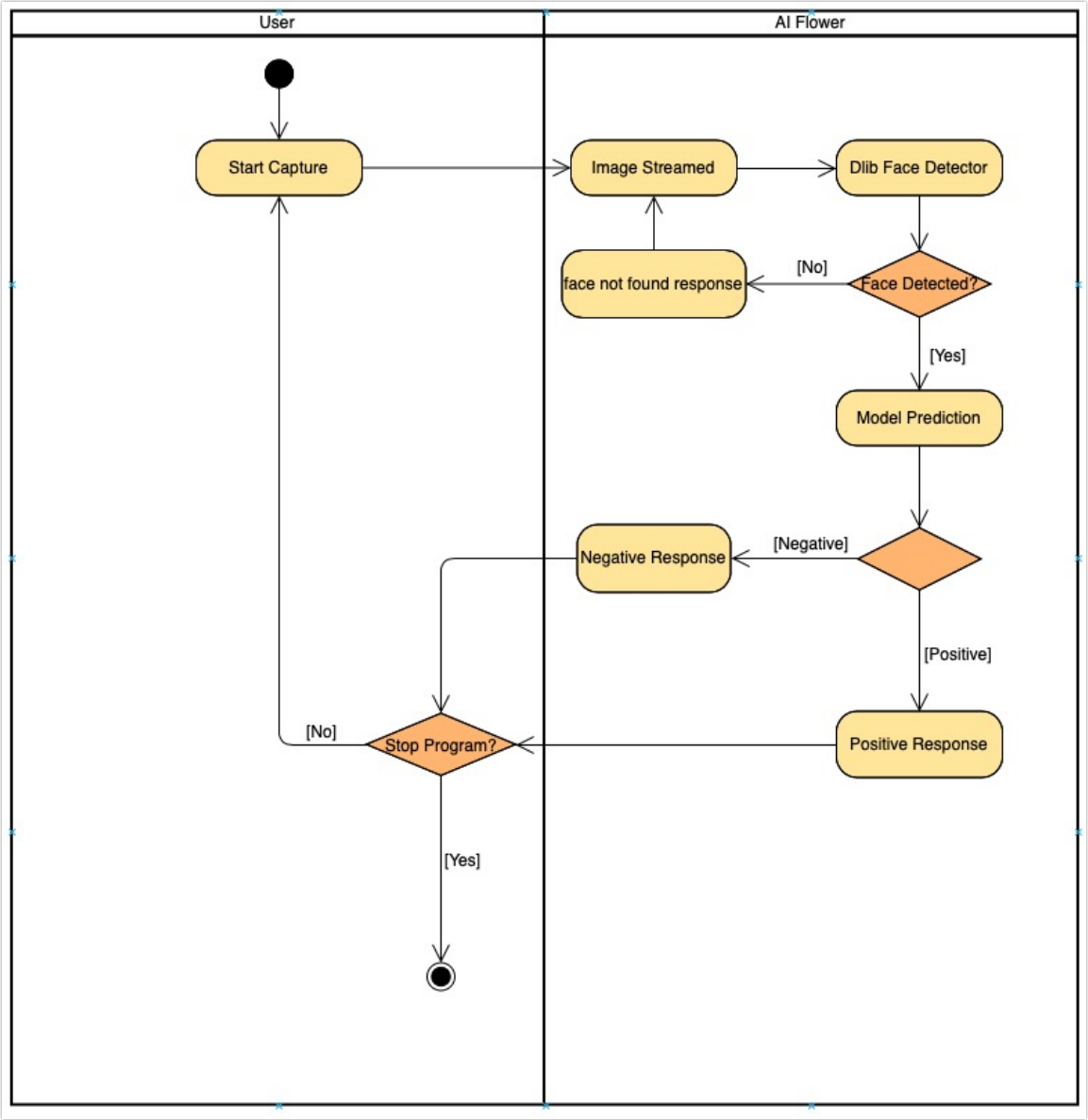


Product Structure - Base



Technical Overview

Human Computer Interaction



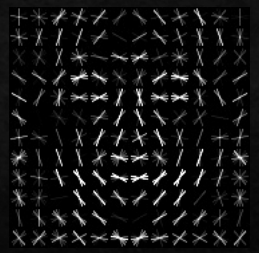
Proposed Method

Outline

- Dlib Face Detector
- Dataset
- Emotion Recognition Model
- Servo Motor Control
- LED control

Dlib Face Detector

- Dlib is a powerful toolkit supporting C++ and Python
- Rationale of Dlib Face Detector:
  - Histogram of Oriented Gradient (HOG)



  - Linear Classifier

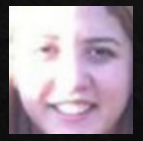
Reasons to use Dlib Face Detector

- Comparing to Haar cascade, DNN Face Detector in OpenCV, CNN Face Detector in Dlib, Dlib HOG Face Detector is:
  - Fastest method on CPU
  - Works very well for frontal and slightly non-frontal faces
  - Light-weight
  - Works well under small occlusion

Dataset

- Final Dataset contains 9178 images
  - 4633 positive images + 4545 negative images
- Mixture of images from three datasets:
  - FER2013 - gray scale, 48 * 48 pixels

  - Expw - Expression in the Wild, cropped and screened



  - Self-collected data - grabbed from Bing, cropped and screened

Dataset

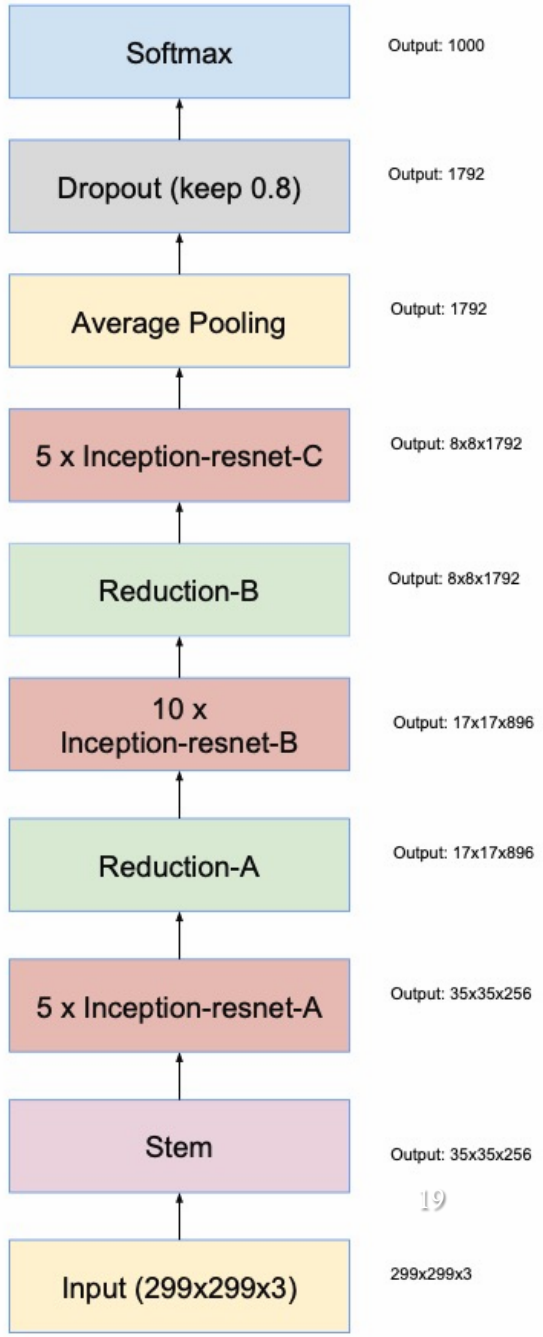
- 7 categories → 2 categories
  - Happy, surprise → Positive
  - Disgust, sad, fear, angry → Negative
  - Neutral discarded

Emotion Recognition Model

- Requirement of model:
  - Low inference time
  - Relatively high accuracy
  - Parameter sizes are small
  - Use limited computing power
- Transfer learning from ImageNet for better result

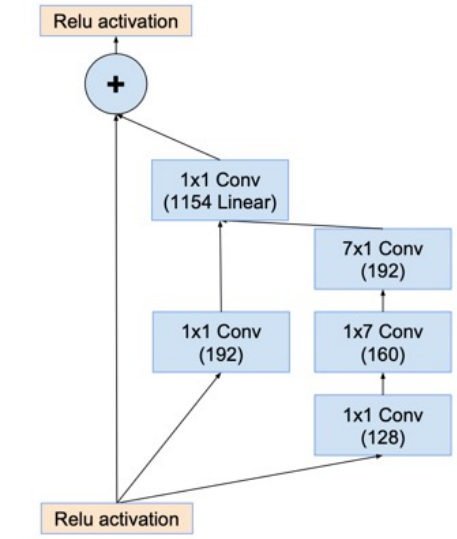
Inception-ResNet-v2

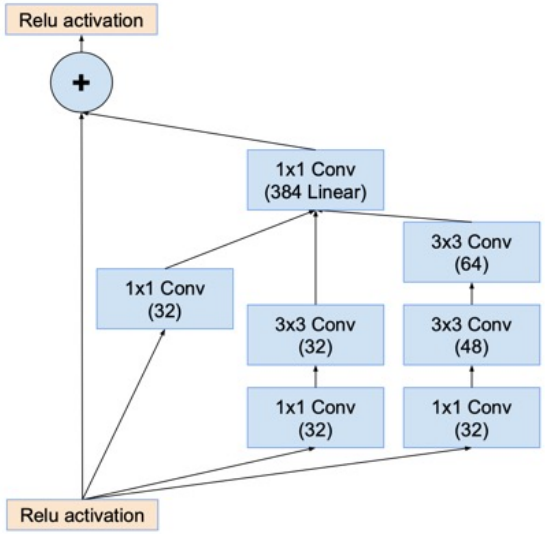
- Very deep ConvNets of Inception architecture + Residual Network
  - The use of residual network improves training speed greatly

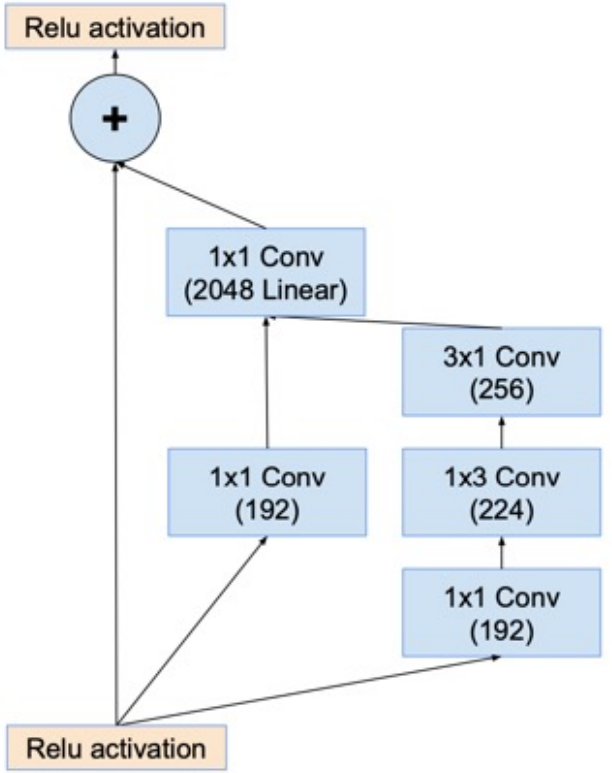


Inception-ResNet-v2

- The schema for Inception-Resnet modules







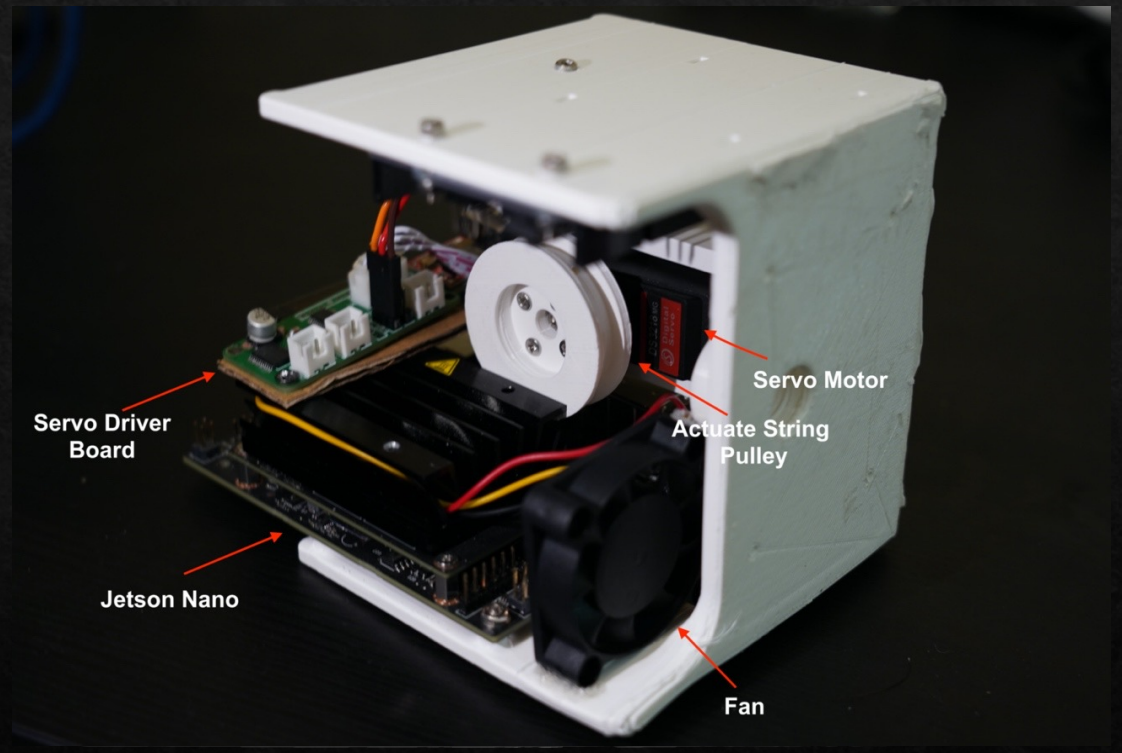
Emotion Recognition Model

- Backbone: Inception-ResNet-v2 + Last Layer: fully connected layer
- Input: 299 * 299 *3, output: 2
- Accuracy of Emotion Recognition Model: 79%, trained on dataset mentioned before

Servo Motor Control

- Jetson Nano's GPIO Library does not implement PWM software

Servo Motor ⟵ Servo Driver Board ⟺ Jetson Nano



Servo Motor Control

- Serial Protocol: Inter-integrated circuit (I^2C)
- I^2C contains 2 wires: clock line (SCL) + a single data line (SDA)
- Servo driver board connects to pin 3 and 5 of Jetson Nano; servo motor to pin 14 and 15 of the board
- Positive prediction: move 330 degrees, string loosens
- Negative prediction: move back, string tightens

Jetson Nano J41 Header

Sysfs GPIO	Name	Pin	Pin	Name	Sysfs GPIO
	3.3 VDC *Power*	1	2	**5.0 VDC** *Power*	
	I2C_2_SDA *I2C Bus 1*	3	4	**5.0 VDC** *Power*	
	I2C_2_SCL *I2C Bus 1*	5	6	**GND**	
gpio216	**AUDIO_MCLK**	7	8	**UART_2_TX** */dev/ttyTHS1*	
	GND	9	10	**UART_2_RX** */dev/ttyTHS1*	
gpio50	**UART_2_RTS**	11	12	**I2S_4_SCLK**	gpio79
gpio14	**SPI_2_SCK**	13	14	**GND**	
gpio194	**LCD_TE**	15	16	**SPI_2_CS1**	gpio232
	3.3 VDC *Power*	17	18	**SPI_2_CS0**	gpio15
gpio16	**SPI_1_MOSI**	19	20	**GND**	
gpio17	**SPI_1_MISO**	21	22	**SPI_2_MISO**	gpio13
gpio18	**SPI_1_SCK**	23	24	**SPI_1_CS0**	gpio19
	GND	25	26	**SPI_1_CS1**	gpio20
	I2C_1_SDA *I2C Bus 0*	27	28	**I2C_1_SCL** *I2C Bus 0*	
gpio149	**CAM_AF_EN**	29	30	**GND**	
gpio200	**GPIO_PZ0**	31	32	**LCD_BL_PWM**	gpio168
gpio38	**GPIO_PE6**	33	34	**GND**	
gpio76	**I2S_4_LRCK**	35	36	**UART_2_CTS**	gpio51
gpio12	**SPI_2_MOSI**	37	38	**I2S_4_SDIN**	gpio77
	GND	39	40	**I2S_4_SDOUT**	gpio78

Pin configuration of Jetson Nano

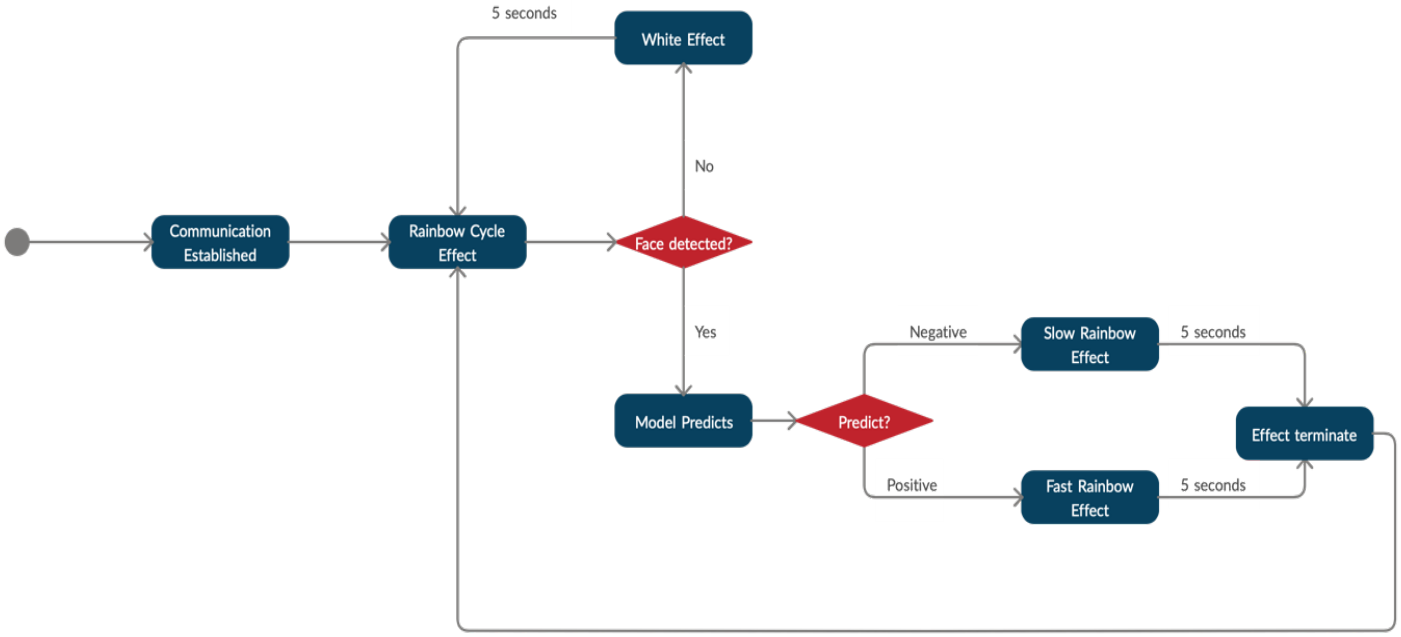
LED Strip Control

- Serial Peripheral Interface (SPI) of Jetson Nano is disabled by default, and hard to enable it

LED strip ⟵ Arduino Board ⇆ Jetson Nano

LED Strip Control

- LED Effect changes show in the chart
- Type of data transmitted: byte
- Default transmission speed: 9600 bits/second
- Multi-threading is adopted in Arduino code
  - Task 1: communication
  - Task 2: change of LED effects



Performance Results

Inference time

- "Warm up" session set for the initial preparation
- After "warm up", the inference time of Dlib face detector and Emotion Recognition Model is recorded
- Inference time:
  - *Dlib* face detector is around **2.20 seconds**
  - Emotion Recognition Model is around **0.49 second**

Fig. Ten consecutive detecting and predicting time for AI Flower

Trial	Inference Time (Dlib face detector)	Inference Time (Emotoin Recognition Model)
1	2.2135565280914307	0.64138221740722 66
2	2.200164318084717	0.4997707862854004
3	2.257528781890869	0.490254402160644 53
4	2.234680652618408	0.520592212677002
5	2.176889419555664	0.5679731369018555
6	2.1882474422454834	0.5439212322235107
7	2.1847081184387207	0.4692976474761963
8	2.2031962871551514	0.4663832187652588
9	2.188779354095459	0.4828264713287353 5
10	2.2409896850585938	0.6679956912994385

Average: **2.20 seconds** **0.49 second**

User Test

- https://drive.google.com/drive/folders/1TJXv8Hl7aW7z0kv2oGbZR5wlyHpcJ5XK?usp=sharing .

Discussion and Conclusion

Difficulties and Limitations

- Relatively few datasets available for training Emotion Recognition Models
- This project requires an integrated knowledge for hardware and software, which is a challenge for CS student
- Tradeoff between effect and budget
- The model can only classify expressions into positive or negative, thus may not have good prediction on neutral expressions
- The model can only identify expression, but not sentiment

Conclusion

- AI Flower:
  - is able to make fast predictions in real-time scenarios
  - has satisfactory accuracy in identifying people's expression, and make timely responses
  - works well with different users, with different gender and extent of face coverage
- Users are willing to interact with the flower, and feel the overall experience satisfactory
- Further Development:
  - Functions that could enable more interactions with users could be developed, e.g. using the screen to display the information inquired by the user
  - Enable SPI on Jetson Nano to have better communication with LED strip

Reference

- Building a Custom JetBot with Jetson Nano. Hackster.io: https://www.hackster.io/gatoninja236/building-a-custom-jetbot-with-jetson-nano-8d41b6
- Cadene. (2017). Pretrained ConvNets for PyTorch. Retrieved from: https://github.com/Cadene/pretrained-models.pytorch
- Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 1, pp. 886-893). IEEE.
- Food and Health Bureau. "1.4: Current Situation in Hong Kong." *Mental Health Review Report*, 2017, 239. Retrieved from: https://www.fhb.gov.hk/download/press_and_publications/otherinfo/180500_mhr/e_mhr_full_report.pdf
- Gupta, V. (2018). Face Detection – OpenCV, Dlib and Deep Learning. Retrieved from: https://www.learnopencv.com/face-detection-opencv-dlib-and-deep-learning-c-python/
- NVIDIA. Jetson-GPIO, 2019. GitHub Repository: https://github.com/NVIDIA/jetson-gpio/blob/master/README.md
- Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*.
- WHO (2009). *Addressing Global Mental Health Challenges. Geneva: Funk, M.*

Thank you

Q & A