

MonkAI Gesture Detector

James Jiang and Argie Cunanan

(ai ai captain)

Team Members and Roles

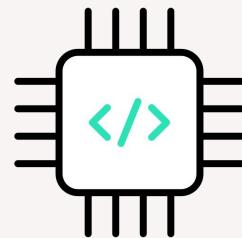
James Jiang

CNN Architecture & Feature Recognition



Argie Cunanan

OpenMV Integration & Firmware



Overview

Project Goals

- Deploy a lightweight, neural network based gesture classifier on embedded hardware using the **OpenMV H7** microcontroller to recognize hand gestures in real-time.
- Classify **4** distinct hand gestures
- Run inference on-device (edge) with minimal latency
- Demonstrate with visual feedback



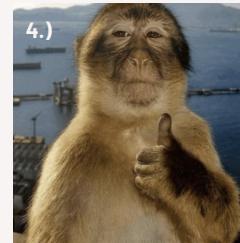
Neutral / no hands



One finger up



Finger on mouth



Thumbs up

Workflow

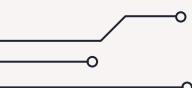
Capture: Use OpenMV H7 camera to grab a grayscale frame.

Features: Compute simple on-board features (ROI averages / thresholded pixels).

Model: Model: Apply a hand-coded neural network (exported Edge Impulse) to classify 1 of 4 gestures.

Output: Print the class over serial and show a matching icon.

Host demo: Laptop script listens to serial and displays the picture for that gesture.



Current Progress

1.) Model + Approach

- Switched from **decision tree** → **neural network** running directly on OpenMV H7.
- Built a **small dataset** (400) of 4 gestures and preprocessed into **fixed-size grayscale** images.
-

2.) Data Collection

- Wrote **OpenMV script** for consistent image capture
- Captured **initial dataset** for
 - Neutral
 - One Finger Up
 - Finger on Mouth
 - Thumbs Up.

Current Progress (CONT.)

3.) Feature Pipeline

- Implemented **on-board preprocessing**: grayscale, downsampling, thresholding.
- Verified **consistent frame rates** and stable **feature extraction**.

4.) Initial Training

- **Built and exported** a tiny neural network (4 layers) via edge impulse.
- Verified export → inference matches the model outputs on sample frames.

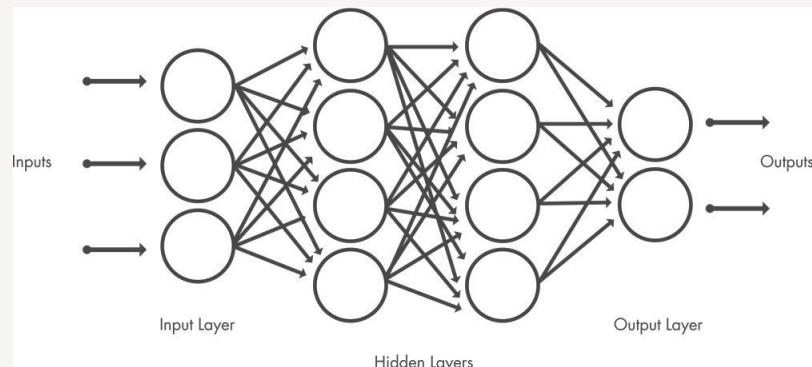
Convolutional Neural Network

Accuracy (Early Stage)

- ★ Model classifies clear poses reasonably well. 70% with live demonstration via Edge Impulse
- ★ Struggles with inconsistent lighting + hand distance

Performance

- ★ ~18-20 FPS on OpenMV H7.
- ★ Latency low enough for real-time gesture feedback.



Current Setbacks

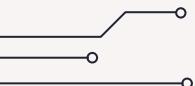
- Lighting variance causes inconsistent preprocessing.
- Model size must stay tiny to run on-board without frame drops.
- Gesture dataset needs more consistency (angle, hand distance).



Updated Plan

Model Improvements

- Collect larger + more consistent dataset.
- Tune neural network: add 2-3 hidden layers,
adjust small weights.
- Add rejection label ("uncertain").



Deliverables

Final Submission

- Full real-time gesture classifier [demo video](#).
- Comparison**: Python vs OpenMV performance + accuracy.
- Short technical [explanation](#) of the neural network executed on-board.
- [Presentation](#) slides

GitHub Submission

- Complete [OpenMV code](#) (capture, preprocessing, model inference).
- Python training + export [scripts](#).
- [Dataset](#) samples + instructions to retrain.
- [README](#) with architecture diagram and deployment guide.