Report Date: 11/18/2022
To: ematson@purdue.edu, ahsmith@purdue.edu, lee3450@purdue.edu
From: The Team Gangsture (Gesture Drone Control)

- Hyeongbin Park (The Leader / mtanger@kw.ac.kr)
- Seunghwan Kim (franz0602@stu.jejunu.ac.kr)
- Yujin Lee (yj5878@kw.ac.kr)
- Soeun Lee (thdms8477@jejunu.ac.kr)

Summary

Every part of development is going to be integrated into the iOS application. It is developed to put landmarks of the gestures on the trained machine learning model in real time. The mobile controller, streamer, and uploader were developed for the iOS environment. Additionally, the introduction and literature review part of the paper were revised. For the methodology part, the dataset was collected with comparing factors.

What "Gangsture" completed this week:

- **Studied and revised the introduction and literature review of the draft**
  - Revised the previous introduction draft with Purdue students.
  - Further research using extra sensors to detect hand gestures [1] - [3] and pre-trained model framework were [4], [5] added to the literature review.
- **Collected dataset with varying conditions for methodology part**
  - Before starting the collecting process, several condition factors to compare were documented and features of calculating a brightness value and the number of images were updated for the dataset in the previous code [6].
  - The number of people, hand type (left or right), and hand size were used as comparing factors for hand gesture datasets and environmental elements were implications of face detection datasets.
- **Completed to put landmarks into the trained ML model as input in real time**
  - The coordinate values of landmarks were put into the trained machine learning model and the direction result was obtained through the iOS platform in real time.
  - The hand type and coordinates are obtained through the MediaPipe and the relative coordinates are calculated by Swift and put into the trained ML model.
  - The input value of trained model is a total of 127. One is the hand type indicating the left or right hand and the others are 126 input values consisting of relative coordinates and 3D coordinate values of 21 landmarks.

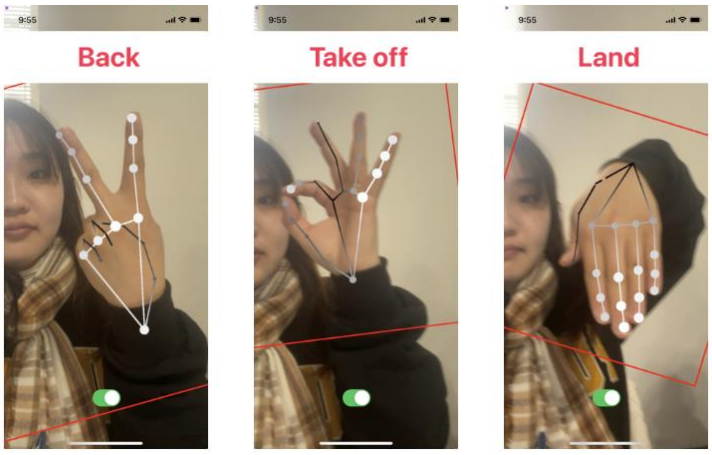

Fig. 1. Control signal results of trained model in real time

- **Completed development of saving a video to Firebase using Swift**
  - There were problems with developing the code because Firebase provides some non-updated documents for saving a file [7].
  - When the button on a mobile application was pushed, the video file from the bundle path was read. The metadata from the Firebase storage was loaded. Thus, the file was uploaded.
  - It is not an automatic function yet.
- **Tested controlling drones with mobile and streaming video in real time**
  - Thread adjustments of drone control class and video streaming class were made.
  - Utilizing AVFoundation, VideoToolBox, and TelloSwift libraries, UDP communications, video streaming, and drone controllers were developed [8].
  - Using UIKit, the front of the application demo was produced.

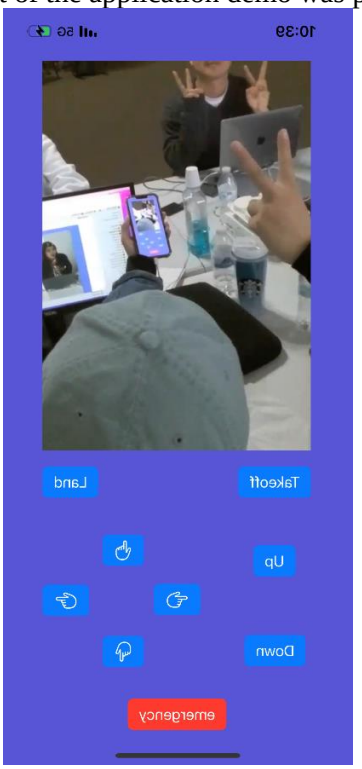

Fig. 2. Interface screen in the demo application

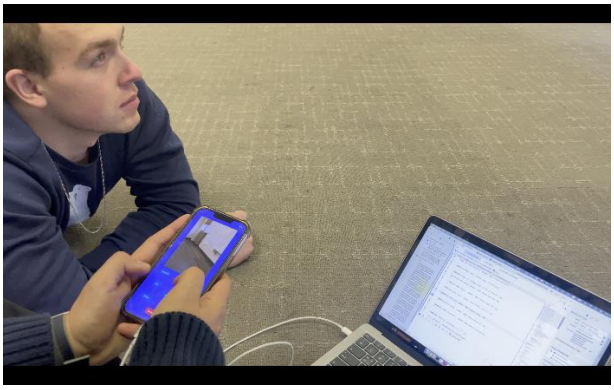

Fig. 3. Mobile drone controller Fig.

Fig. 4. Flight via UDP socket communication

Things to do by next week

- Write the methodology part of the draft and revise the literature review
- Optimize and develop the process of face recognition
- Develop a function connecting to the Internet and make an automatic save function
- Save video taken by drones automatically to the iOS device

Problems or challenges:

- **Insufficient documents about Swift**
  - Firebase is made by Google. The Firebase tutorial using Swift provides was out of date. It needs more research about each module and input of functions.
- **Security policy of Apple**
  - Some functions need an Apple developer account which caused errors in building the application.
  - For example, it is not available to connect to the Internet through the iOS application. This function will be developed with an Apple developer account next week.
- **Intermittent thread crashes during video streaming**
  - In the process of thread alignment, memory collision problem occurs intermittently in the process of decoding due to an image codec problem.
  - It is necessary to find a video decoder process to suit the mobile environment by checking the thread collision part.
- **Detailed development of natural drone movement control**
  - Since there are some unnatural controls on the continuous movement of drones, parameter adjustments are needed to make them more natural.
  - It is necessary to study how to convert altitude, speed, and coordinates into parameters in more detail for the drone to respond more closely to continuous movements.
- **Normalize z-coordinate value**
  - Due to the difference between the MediaPipe version of the trained model and the iOS app, the Z-coordinate value is not normalized, so it must be normalized in iOS [9].
  - If possible later, the recent version of MediaPipe will be applied on the iOS app.

References

[1] C. -C. Tsai, C. -C. Kuo and Y. -L. Chen, "3D Hand Gesture Recognition for Drone Control in Unity," 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE), 2020, pp. 985-988, doi: 10.1109/CASE48305.2020.9216807.

[2] J. Jun, B. Zhao, P. Zhang, F. Jia and P. Fang, "Interaction with UAV Formation based on Gesture Recognition with Leap Motion," 2021 33rd Chinese Control and Decision Conference (CCDC), 2021, pp. 3275-3280, doi: 10.1109/CCDC52312.2021.9602101.

[3] B. Hu and J. Wang, "Deep Learning Based Hand Gesture Recognition and UAV Flight Controls," 2018 24th International Conference on Automation and Computing (ICAC), 2018, pp. 1-6, doi: 10.23919/IConAC.2018.8748953.

[4] F. Naseer, G. Ullah, M. A. Siddiqui, M. Jawad Khan, K. -S. Hong and N. Naseer, "Deep Learning-Based Unmanned Aerial Vehicle Control with Hand Gesture and Computer Vision," 2022 13th Asian Control Conference (ASCC), 2022, pp. 1-6, doi: 10.23919/ASCC56756.2022.9828347.

[5] A. Findelair, X. Yu and J. Saniie, "Design Flow and Implementation of a Vision-Based Gesture-Controlled Drone," 2022 IEEE International Conference on Electro Information Technology (eIT), 2022, pp. 320-324, doi: 10.1109/eIT53891.2022.9813802.

[6] Da NioDa Nio 18255 silver badges1515 bronze badges, Mark SetchellMark Setchell 179k2727 gold badges249249 silver badges398398 bronze badges, Prachi VaishnavPrachi Vaishnav4877 bronze badges, and Dhanashree DesaiDhanashree Desai 74055 silver badges55 bronze badges, “Get brightness value from HSV - python,” *Stack Overflow*, 01-Sep-1966. [Online]. Available: https://stackoverflow.com/questions/57114114/get-brightness-value-from-hsv-python. [Accessed: 18-Nov-2022].

[7] “Upload files with cloud storage on Apple platforms | Firebase Storage,” *Google*. [Online]. Available: https://firebase.google.com/docs/storage/ios/upload-files. [Accessed: 18-Nov-2022].

[8] “Creating a custom camera ViewEdit pagepage history,” *Creating a Custom Camera View | CodePath iOS Cliffnotes*. [Online]. Available: https://guides.codepath.com/ios/Creating-a-Custom-Camera-View. [Accessed: 18-Nov-2022].

[9] Google, “Release changes in z coordinate normalization in 3D landmarks · Google/mediapipe,” *GitHub*. [Online]. Available: https://github.com/google/mediapipe/releases/tag/v0.7.6. [Accessed: 18-Nov-2022].