Kavan Mehta

Mrs. Secord

ISM 1

18 March 2022

Research Assessment #21

Date: 18 March 2022

Subject: Google's approach for hand gesture recognition

MLA citation(s):

Zhang, Fan, and Valentin Bazarevsky. "On-Device, Real-Time Hand Tracking with MediaPipe." *Google AI Blog*, Google, 19 Aug. 2019, https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html.

Assessment:

As I am currently creating my ISM Final Product, a hand gestures recognition model that can perform simple tasks on a device, I aimed to explore current solutions to hand recognition to learn about the possible algorithmic approaches I can take when creating my project with convolutional neural networks. I was hoping to find out more about the process of analyzing hand movements, current solutions in the field of AI and their approaches, and possible ideas that I can implement with the guidance of my mentor for my final product. To learn this new knowledge about final product ideas, I found a research blog from Google,"On-Device, Real-Time Hand Tracking with MediaPipe," which went over a new technology and approach to hand gesture recognition created by engineers at Google using MediaPipe and other software tools.

Through this article, I was able to understand the process and approach that engineers at Google used to solve the complex task of analyzing hand movements and gestures. I mainly gained new knowledge about the approach I could use by learning about how Google used a palm recognition model called BlazePalm to reduce unnecessary analysis of conflicting external surroundings, then applied a model to recognize 21 3D keypoints of the hand to decide which hand gesture is being shown in the image, and ultimately trains and becomes more accurate over time in recognizing the palm and gesture of a hand in real-time video (Zhang and Bazarevsky 3). Thus, I was able to see how I can possibly implement a program to first segment the hand to specifically focus only on the hand and get rid of any particular external disturbance, then perform analysis to learn what the hand gesture is through finger placement/finger poses, and ultimately improve its accuracy to get the best possible results. Then, I plan to use python libraries and frameworks to enable my machine learning model to perform tasks on a device and demonstrate the application of my final product.

I also learned that MediaPipe could be a great resource for me to learn and explore the algorithmic approaches to hand gesture recognition and utilize it to help create my final product. Furthermore, in the article, I learned that MediaPipe has “modular components, called Calculators…to solve tasks like model inference, media processing systems, and data transformations" on multiple devices (Zhang and Bazarevsky 5). It also helps the efficiency of the overall project by running separate modules only at necessary times leading to “high performance and optimal throughput of the ML pipeline” (Zhang and Bazarevsky 5). I plan to look into MediaPipe for its multiple use cases and potential benefits for my final product.

I am currently in the process of programming and deciding on specific algorithmic approaches for my final product. I also am looking forward to meet my mentor to discuss the

ways I can implement these strategies in my project and am really hopeful to obtain a deeper understanding about image recognition through my mentorship and project. I will continue to explore any other articles and tutorials to create my project successfully in this semester of ISM with this new knowledge about Google's approach to hand gestures and MediaPipe.