Using Gesture Recognition to Navigate Google Chrome

Yuki Adams
Earlham College Department of Computer Science

Motivation and Background
- Humans frequently use hand gestures and body language to aid in communication [1]
- Developing a static gesture recognition system is a complex challenge in computer vision
- Main question: using just a laptop, how can I develop a gesture recognition system that works in real time?
- Minimal work in using gesture recognition with the browser
  - Chrome Gesture Control (has issues using webcam) [2]
  - Hand-Web-Browser [3]
- Other hardware options are more robust, but at an expense

Methodology
- Live video is passed to the preprocessing stage
- Video is converted into grayscale, blurred, overlayed with a binary mask, and keypoints are found
- After each frame is preprocessed, keypoints are drawn on the image
- The largest white area is the hand
- Circularity, centroid, and convex hull are features being extracted from the keypoints
- These features are passed to the classifier, and the gesture is predicted based on the training set

Tools
- I utilized an open-source library called OpenCV because it could handle
  - real-time video
  - object detection
  - image processing
- Using Scikit-learn's KNN classifier to predict gestures
- Google Chrome API to make extension

Preliminary Findings
- KNN is computationally expensive and requires a significant amount of memory during the prediction stage
- Latency is expected during preprocessing and prediction stages

Next Steps and Future Work
- Finish tweaking classifier
- Google Chrome extension
- Increase functionality
  - Allow for mouse control
  - Open/Close browser
  - Pair with facial recognition
  - Use different classification method

References

Acknowledgements
I am grateful for the advising I received from David Barbella and Xunfei Jiang.