Reid Glaze
Computer Vision
Dr Iona Fleming
4/25/22

Final Project Report

Project Idea and Goals

The goal of this project was to write a program that can detect the number of fingers that the user is holding up in front of the webcam. OpenCV was be used. However, TensorFlow and similar libraries were not. Although it is not perfect, the program was made to be as robust as possible and can work in different lighting environments. Additionally, it was written to work with different skin colors.

Algorithm / Approach

The first goal in this project was to segment the hand correctly. The original idea was to use binary thresholding. Originally, this seemed like it would be a relatively simply task. However, it ended up being much more complicated because of the variable lighting and texture of the hand.

The second goal was to count the number of segments that would represent the number of fingers. This means that it was necessary to have clear segmentation within the fingers and a clear separation of finger segmentation. This proved to be a much more difficult problem than anticipated because segmentation includes the wrist and other parts of the hand.

Data Analysis and Results

The project was largely a success. It can detect the number of fingers that the user is holding up (0-5). Furthermore, it works in variable lighting and with different skin colors. However, the program is not without several flaws. It does not always segment the number of fingers correctly and may require the user to move their hand for the segmentation to work. Additionally, the hand segmentation does not appear to look like a normal hand because of the of the blurring, contrasting, smoothing, and thresholding techniques used. However, the program provides visual ques that can help the user diagnose the problem easily.

Computer Vision Methods used for Segmentation

<u>Accumulated Weight for Background</u>

The program takes the weighted average for the region of interest bounded by the red box in Figure 1. This is done over time and warns the user. This is done so that a variety of backgrounds can be used. It is important that the user keeps the background in the region of interest consistent over this time.

<u>Absolute Difference</u>

The difference between the user's hand and the measured weighted background is taken. This is done so that the program can account for different colors and lighting.

Reid Glaze
Computer Vision
Dr Iona Fleming
4/25/22

Noise for Color Image

A very small amount of noise was added at the very beginning. It could have little to no effect, but the idea was to make texture in the hand less significant.

Median & Gaussian Blurring for Color Image

Median and gaussian blurring where applied several different times to the color image. This was also done to make texture in the hand less significant. When fingers cross over the palm, this often causes problems with the segmentation. This is shown in figure 1. However, too much blurring can make segmentation more difficult. The kernel size and frequency of blurring was experimented with to minimize both problems.

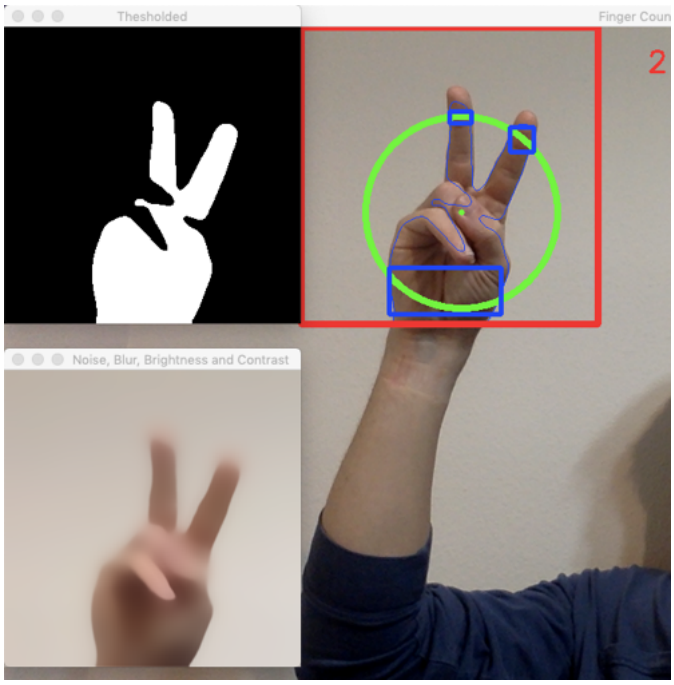

Figure 1 – Two Fingers

Bilateral Filter for Color Image

The bilateral filter was found to be a very important tool in performing segmentation and was used 3 times. The bilateral filter is effective at blurring an image while keeping the edges sharp. The downside of this filter is that it slows down the program a lot. However, the improvement in segmentation that resulted from this filter was worth the downsides.

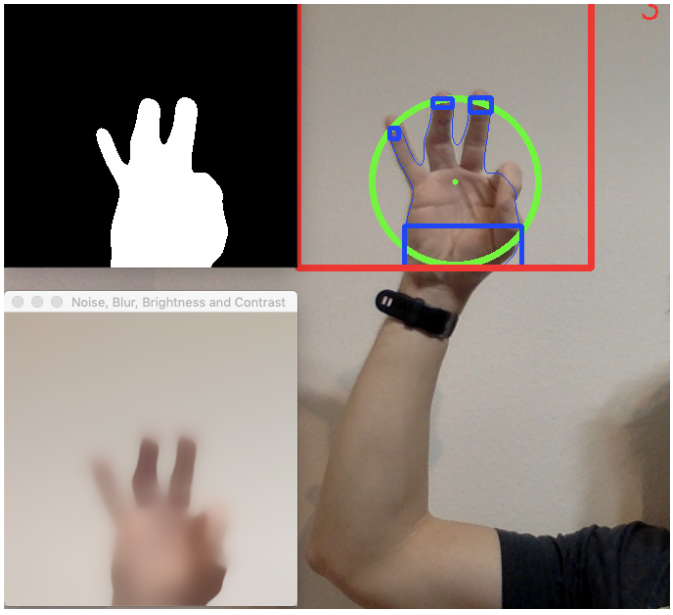

Figure 2 – Three Fingers

Contrast Adjustment for Color image

A very small increase in contrast was applied to the region of interest twice. The increase in contrast was done in between the times when the bilateral filter was used. The point of this was to increase the contrast of the edges. It was found that a large adjustment in

contrast caused problems with segmentation inside of the hand. Therefore, this was used sparingly.

Color to Grayscale Conversion

Color to grayscale conversion is done before applying the binary threshold. This is a necessary task for performing binary thresholding. Additionally, it was important to use various image processing methods before performing the conversion because the inside of the hand can be complicated and difficult to segment.

Gaussian Blurring for Grayscale Image

Gaussian blurring was used for the grayscale image. This was done to minimize any abnormalities in the segmentation that may occur inside of the hand.

Edge Smoothing for Grayscale Image

Edge smoothing was applied to the grayscale image. This was done mainly to make segmentation of the fingers clearer. A kernel size of 3 was used and 3 iterations were performed.

OTSU Binary Thresholding

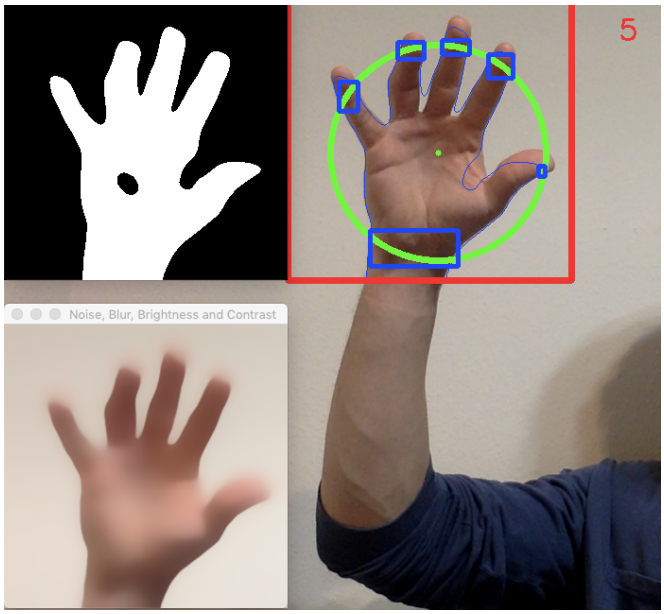

Figure 3 – Five Fingers

Reid Glaze
Computer Vision
Dr Iona Fleming
4/25/22

OTSU binary thresholding was used to finally segment the grayscale image into a binary image. This method sets a custom threshold for each kernel used. The size of the kernel used was 11 by 11. This method was used to account for different colors and levels of brightness. It also helps segment variability inside of the hand. Finger counting and segmentation in a dark environment is shown in figure 3. Although the segmentation of the hand looks much different, it is still able to count all 5 fingers.

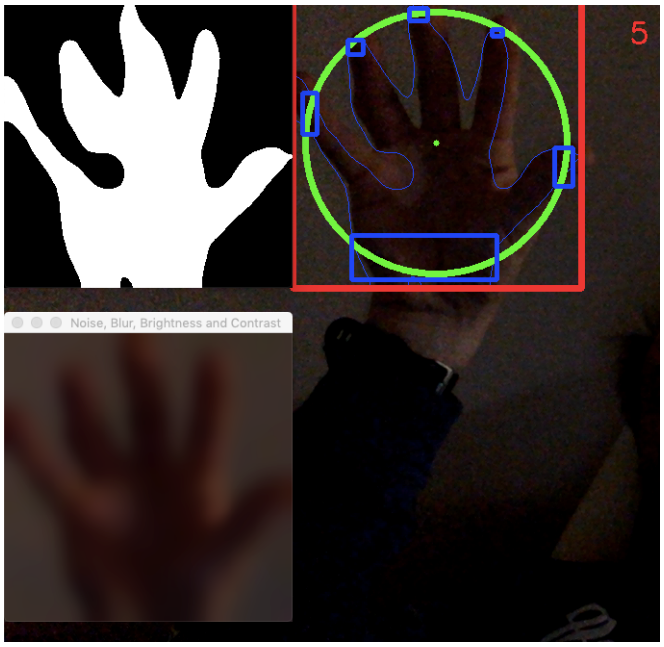

Figure 4 – Five Fingers in the Dark

Find Contours

After binary thresholding, the find contours function is used to segment the hand. The contours are depicted with the blue lines shown inside of the region of interest.

Computer Vision Methods Used for Finger Counting

Find Center of Convex Hull

The convex hull is found, using the most extreme points that lie to the right, left, top, and bottom of the hand. The averages between the right and left points as well as the top and bottom points are taken to find the center. The center is marked by the green dot in the hand.

Maximum Euclidian Distance

Four Euclidian distances are taken, measuring the distance in between the center and each of these 4 extreme points in the convex hull. The maximum of these Euclidian distances is taken before determining the radius of the circle.

Circle used for segmentation

A radius value of 0.85 times the maximum Euclidian distance is used to determine the size of the circle. It is drawn around the center of the convex hull. The 0.85 value was taken after experimentation but can be modified for different results.

Segmentation counting

Reid Glaze
Computer Vision
Dr Iona Fleming
4/25/22

The find contours function is used to find how many segmentations lie on the circle. The circular region of interest is used as a mask. Each of the segmentations are outlined by blue boxes in the images.

Excluding Trivial Segmentations

Since the wrist lies outside of the circle, it is important to exclude the contour region that lies at the bottom of the image. Additionally, forming a fist can cause a segmentation at the top of the hand when no fingers are being raised. This segmentation is excluded by dropping all segmentations that exceed 15 percent of the circumference. Exclusion of these segmentations is shown in image 3.

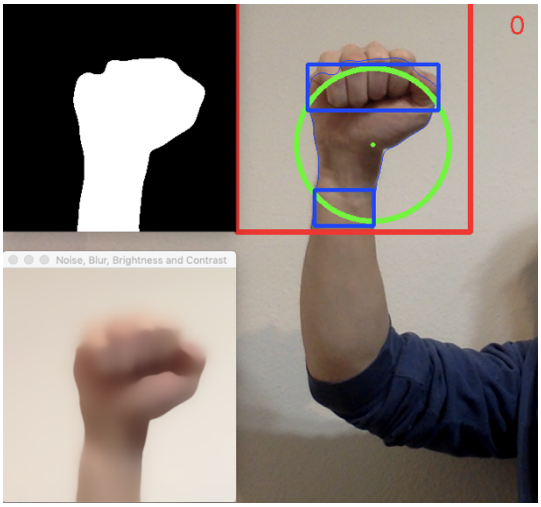

Figure 5 – Excluding Trivial Segmentations

Methods that did not work

Excessive Blurring

Originally, blurring was used many times with large kernel sizes. It was theorized that this blurring would minimize dark and light spots on the hand and therefore make thresholding and segmentation easier. However, this level of blurring made the edges harder to detect and did not work well when lighting conditions were changed.

Too Little Blurring

When little to no blurring was used, the segmentation of the hand was patchy. This can result in problems because a single finger could result in two or more segmentations. It is important to use enough blurring so that the black patches are limited to parts of the hand that are not counted for segmentation.

Binary Thresholding with a Constant

When a single constant was used for segmentation, there were a few problems. This constant needed to be modified for variable levels of brightness and skin colors. Segmentation worked well when there was a large amount of contrast in between the background and the hand. However, it did not work well in other conditions. The segmentations were much patchier, which causes problems with finger counting.

Excessive Contrasting

It was theorized that more contrast would help improve the accuracy of the contours. While this may have been accurate in some areas, the higher use of contrasting caused patchiness inside of the hand. For this reason, contrasting was used very sparingly.

Excessive Use of the Bilateral Filter

While the bilateral filter is used 3 times in this program, it was not used as the sole method of blurring in this program. This is because the bilateral filter can cause the program to run slowly. This program does run slowly but not excessively so. It was important to find a balance in using this filter so that the hand would segment well, and the program would run without crashing or being obnoxiously slow.

Radius Multipliers that were too High or too Low

A factor that was multiplied by the maximum Euclidian distance was tested at different values. A value that was too high resulted in segmentations that were not counted. A value that was too low resulted in counting segmentations that were not represented by fingers. These segmentations would come from other parts of the hand. Ultimately, a value of 0.85 was chosen.

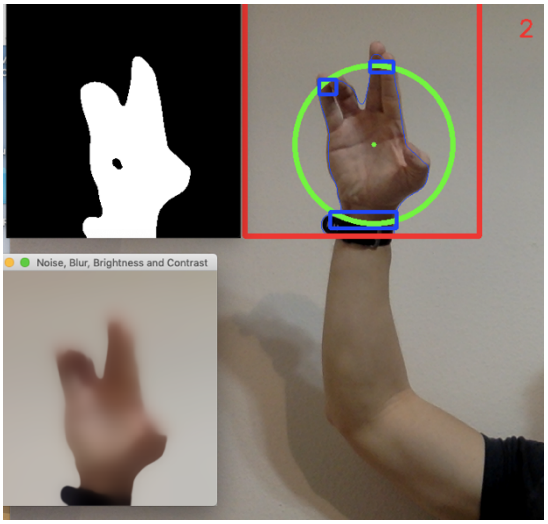

Figure 6 – Joining of Segmentations

Known Limitations and Possible Future Extensions

Joining of Segmentations

One problem with this program is that the fingers must be clearly separated to segment correctly. When the star trek symbol is held up, the program shows only two fingers. We are unsure how to solve this problem and could be rather difficult to solve without using more advanced libraries.

Watches or Other Wearables

Reid Glaze
Computer Vision
Dr Iona Fleming
4/25/22

Sometimes, segmentations can be inaccurate due to watches or other wearables. The segmentation of the watch changes the size of the circle and can result in problems.

Shadows

Shadows can alter the segmentations and cause the size of the circle to change. In this example, the circle is made to be too large and one of the fingers is not counted.

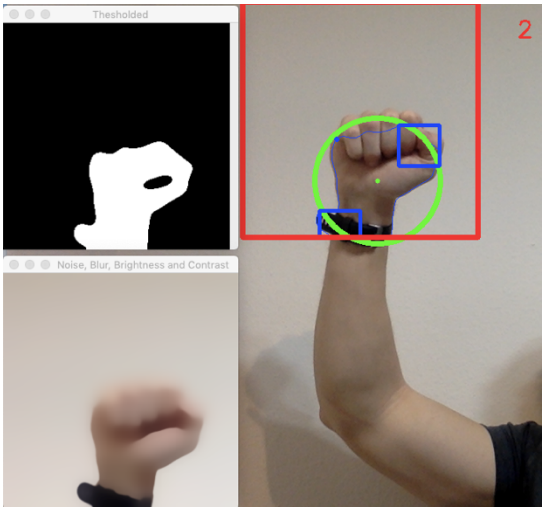

Figure 7 – Watches or Other Wearables

Inaccurate Segmentations

Inaccurate segmentations can be caused for a variety of reasons including lighting, angle and proximity of the hand, hand shape. This program is far from perfect and has lots of room for improvement. Visual ques can be used to help the user position the hand correctly, but this program may not be reliable without such feedback. However, it provides nice insight into how the program is implemented and strives to improve robustness and adaptability to variable conditions.

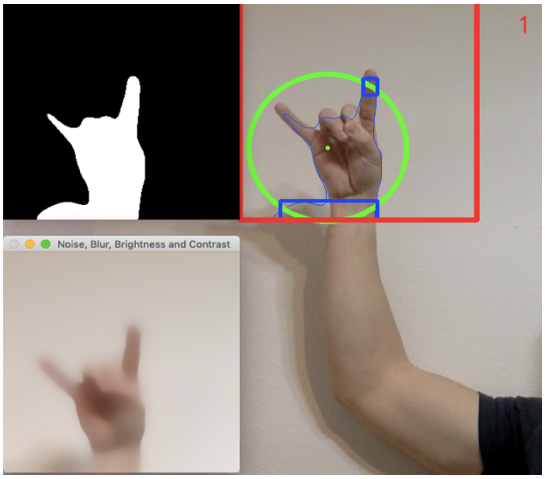

Figure 8 - Shadows

Contributions

Reid worked on implementation of different methods and visual demonstration of the segmentation and finger counting. He worked on the CVproject.py file.

Rohit worked on testing thresholding methods, blurring methods, morphological operations, and evaluation of distance measures. He worked on the Capstone_BlurSmooth.py, Capstone_DistMeas.py, Capstone_Moropho.py, and Capstone_Threshold.py files.

Things Learned

1) I learned the advantage of using OTSU binary thresholding over regular binary thresholding. I learned the significance of this tool by experimenting with variable lighting conditions.
2) I learned how a bilateral filter can be useful for segmentation of an object from a background. This was discovered by using the tool in place of other forms of blurring.
3) I learned how the use of a mask can be used to count segmentations. This was done by utilizing the circle.

Reid Glaze
Computer Vision
Dr Iona Fleming
4/25/22

Advice to Next Year's Computer Vision Students

I would recommend starting on the homework assignments as early as possible. The work is a lot more enjoyable when you do not have to stress last minute. I would also recommend working with a partner. The assignments take a lot of time, so it is good to be able to make a friend while working. It is also easier to be creative when different minds are working together.