American Sign Language Detection

Team #8

Team members:

Ahmed Essam El-deen Sadek	1	3
Ahmed Mohamed Zakaria	1	4
Abdullah Ezzat	1	34
Omar Ahmed Desoky	1	37

Used Algorithms:

Hand Detection:

- Smoothing using Gaussian filter to remove noise.
- Segmentation by skin tune Using Log-Chromaticity.
- Background Subtraction.
- Face Removal using Haar Cascade.
- Closing to fill hand.

Feature Extraction:

- Find Contours to get the largest Contour (Hand).
- Convex Hull of the hand.
- Convexity defects to get defects in the convex hull.

Classification:

- Classify hand based on number of fingers in each direction, orientation and bounding box aspect ratio.

Trials, Results and Analysis:

Hand Detection and Segmentation:

First, we used color model which is a combination of HSV, YcbCr and RGBA Color spaces.

Results:

- there are a lot of false negatives as Shown in Figure(1)

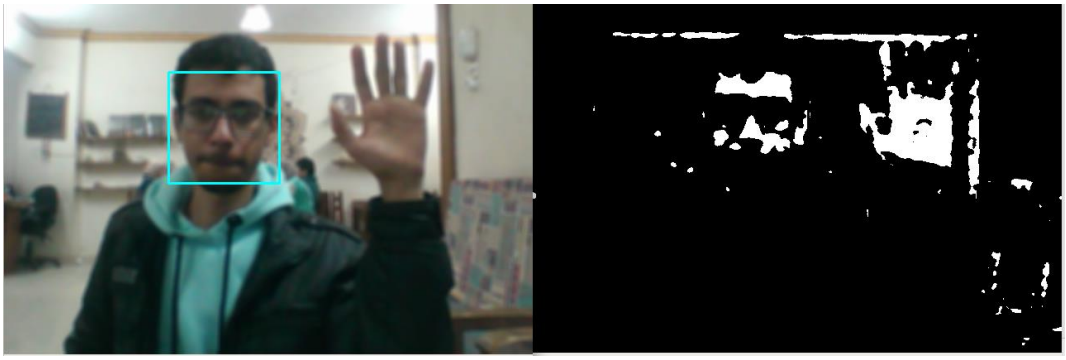

Figure 1

- So we finally used log chromaticity as shown in Figure(2)

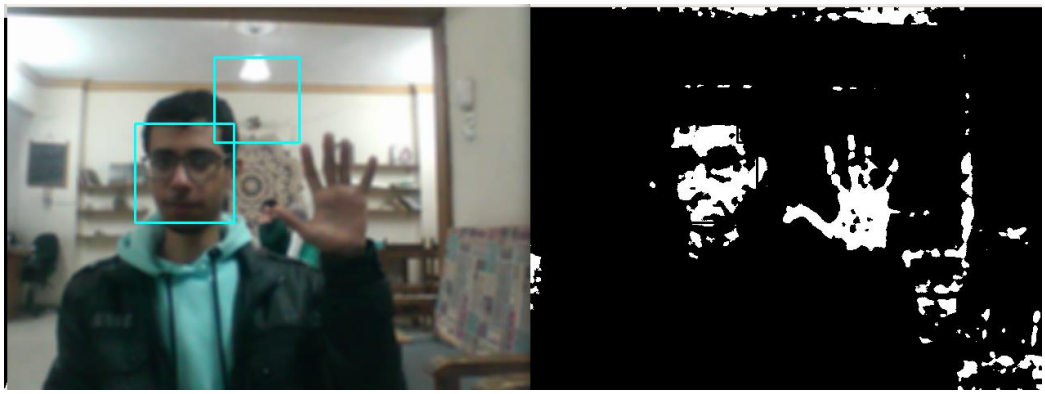

Figure 2

- regarding face removal, we tried first to use the aspect ratio of the contours of the binary image to filter hand from face, but it gives nearly 65% accuracy, that happens due to different orientations and dimensions, also if hand is closed it's almost the same as face.
- So we used Haar cascade Classifier to Remove Faces from image as Shown in Figure(3).

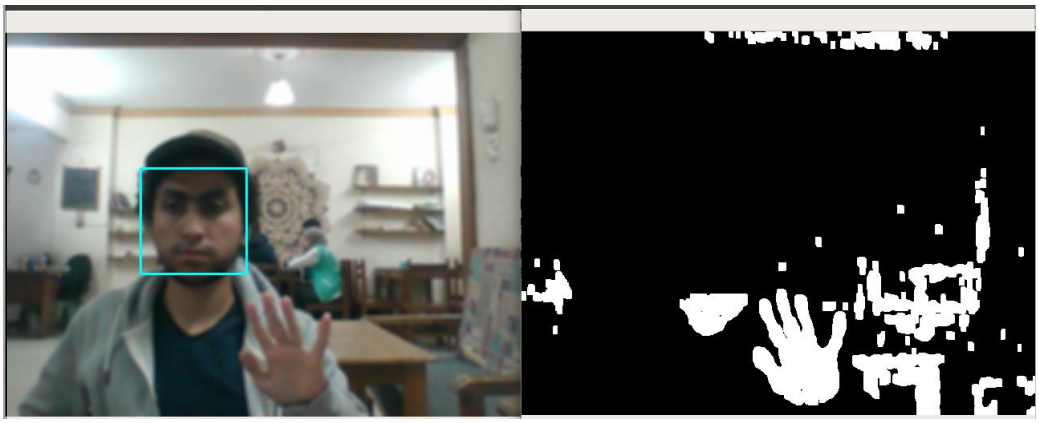

Figure 3

- regarding background subtractions, first we extracted region of interest (moving region) depending on difference among three frames.
- as shown in Figures (4,5,6,7,8)
  1. as long as hand moves, it works well.
  2. If nothing moves, outputs black images.
  3. We tried to handle this problem by keep tracking of the moving previous area, but it doesn't give good results as it always show the Palm of the hand.

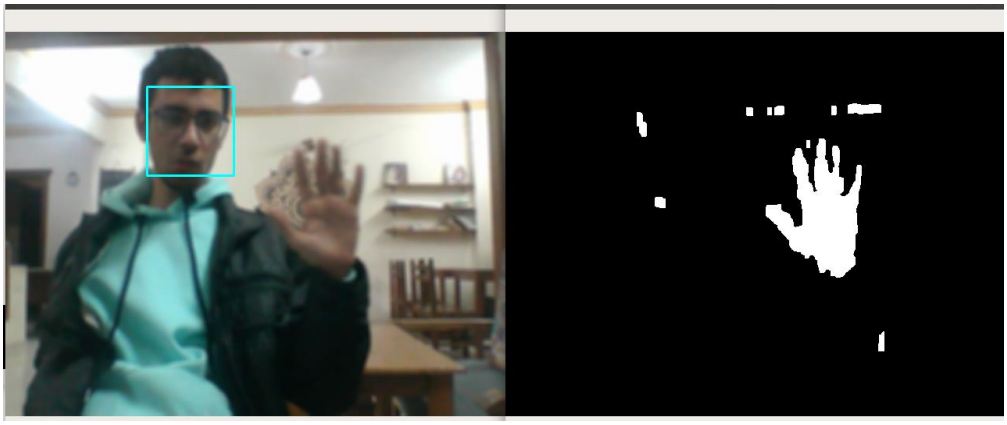

Figure 4

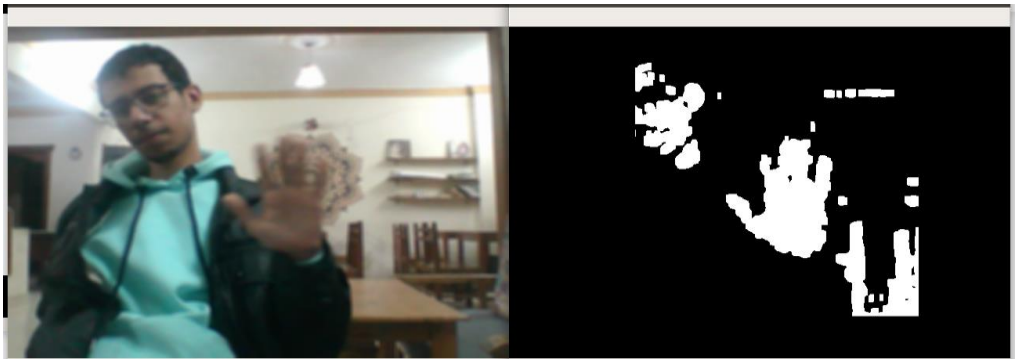

Figure 5

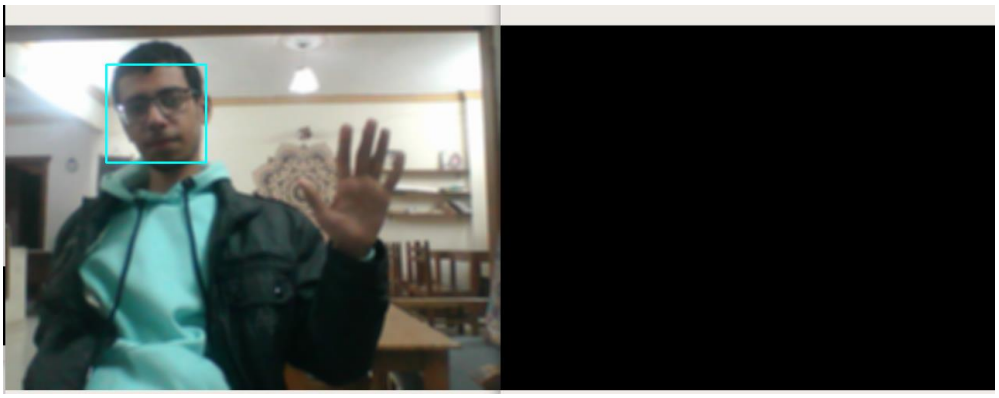

Figure 6

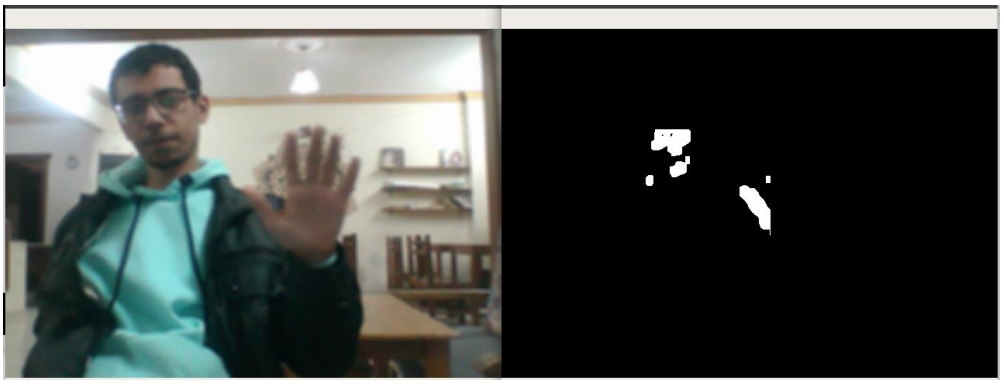

Figure 7

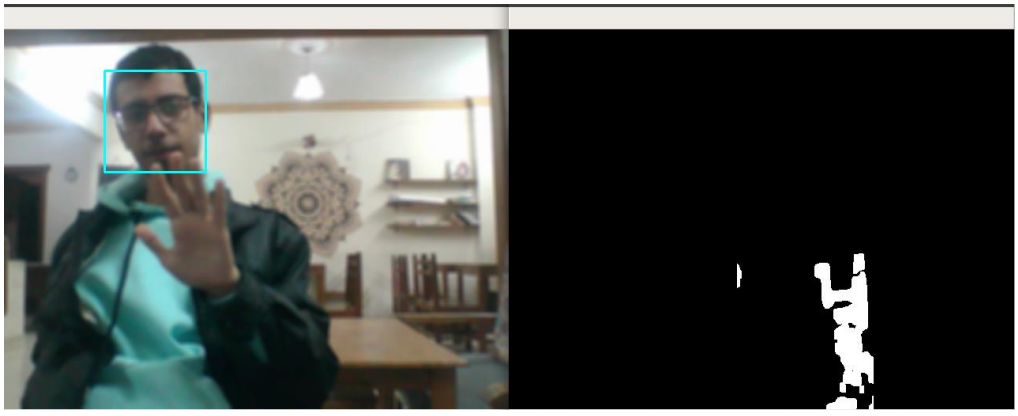

Figure 8

- So we finally used, a background subtraction method which has a dynamic background model, updated regularly from previous frames to calibrate itself.
- We subtract it from the current frame to get the foreground image as shown in Figure(9).

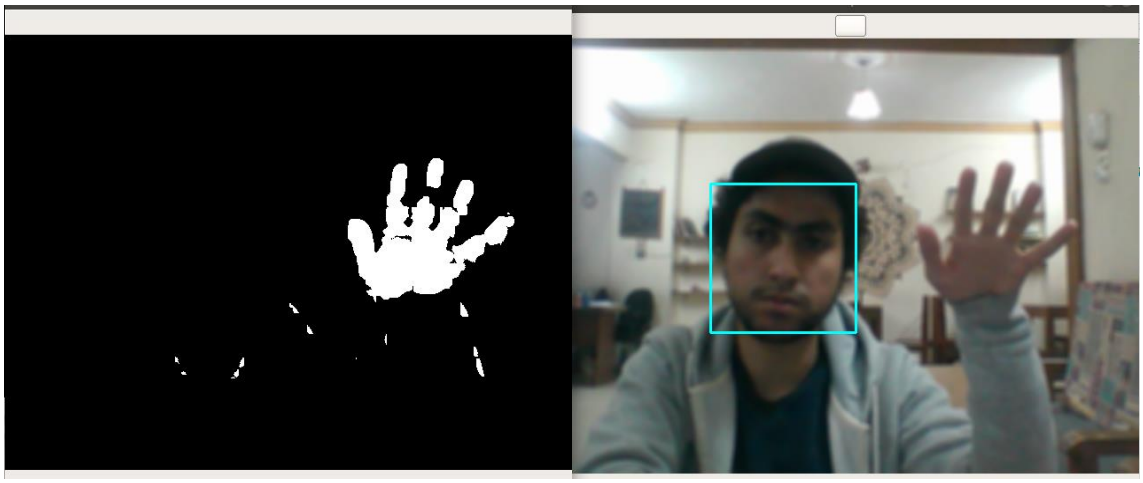

Figure 9

Feature Extraction:

First, we got the contours of the mask. And started working on the maximum contour, to avoid noise, to get the convex hull of the hand, which we implemented ourselves (shown in green lines in figure 10).

Then, we got the defects of the convex hull to get how many fingers are there (shown in dots in figure 10).

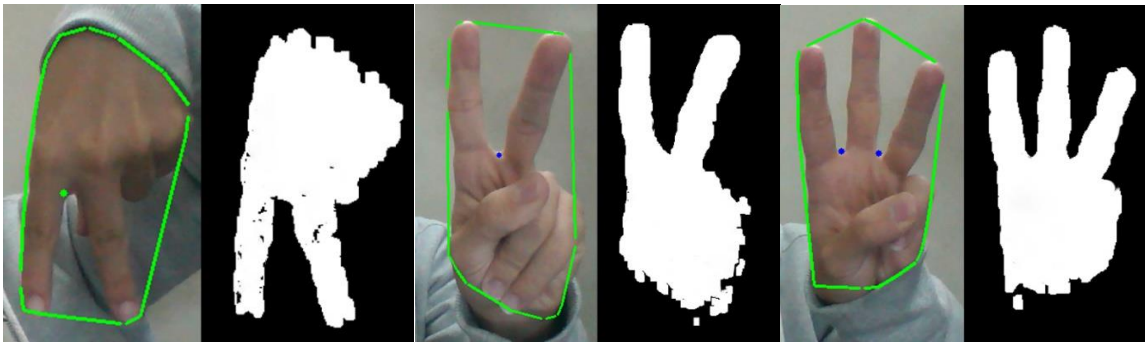

Figure 10

Then, we counted the defects in each direction (shown in the different coloring of the dot) to help us detect different signs that has different directions.

Also, we used the maximum contour to get the bounding box of the hand which we use to determine the orientation of the sign.

We also used the ratio between the area of the left half and right half of the contour to determine some other signs.

We also tried to get number of fingers using the local maximums in the convex hull but it was too difficult to set the right threshold as shown in figure 11 (vertical fingers shown in green and horizontal fingers shown in red).

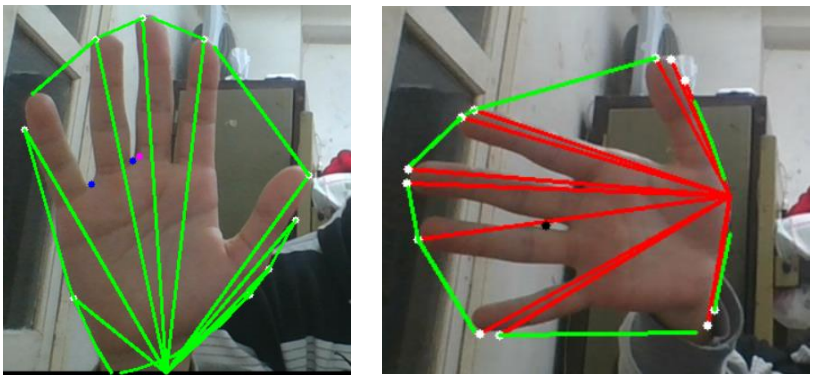

Figure 11

Classification:

We used the features mentioned above to classify the signs and recognize 17 different letter/number listed below in figure 12.

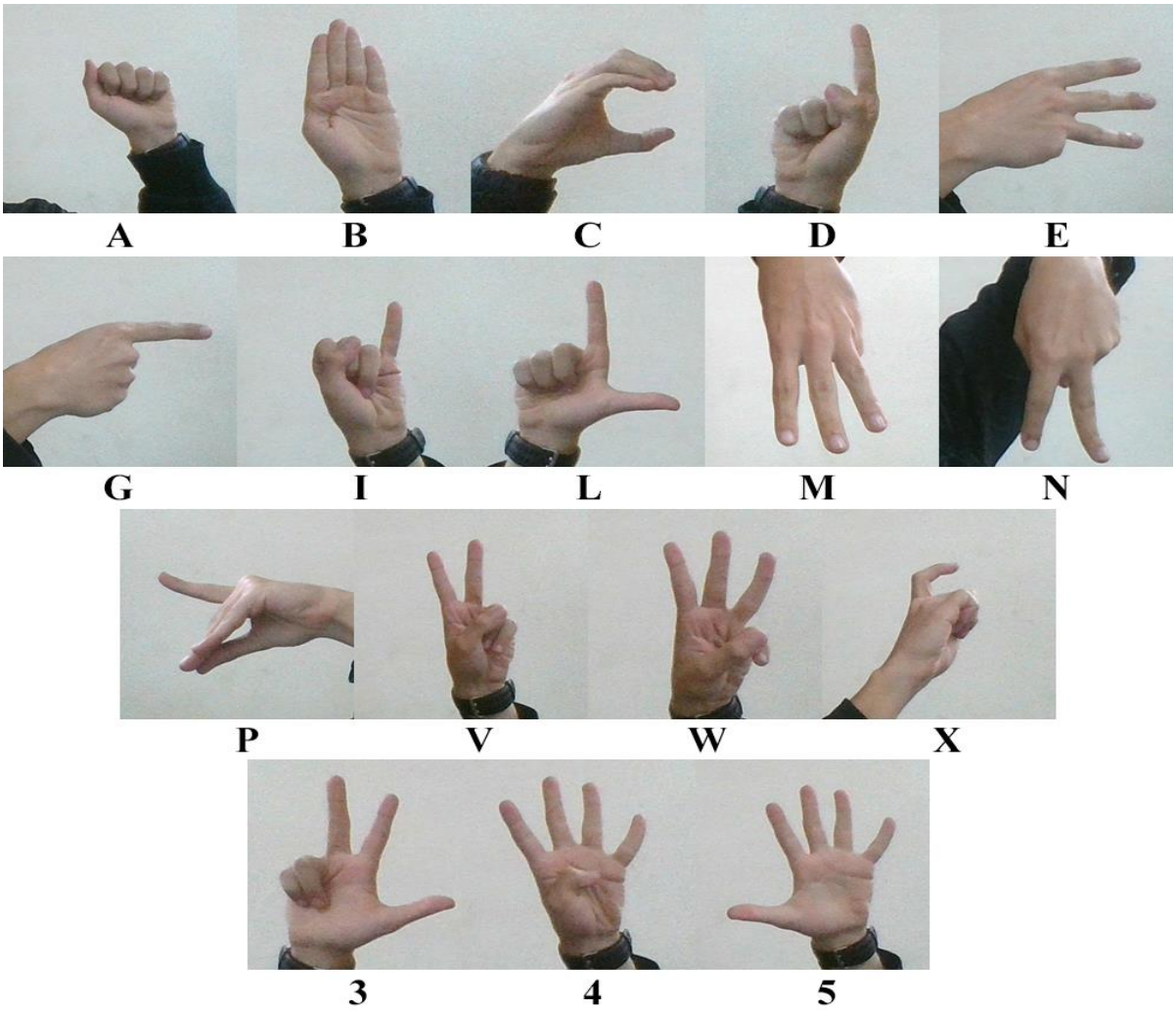

Figure 12

Weaknesses:

Prone to weak/strong lightening conditions (causes noise on hand segmentation).
Only works on one hand; if detected multiple skin color in the frame, it works on the biggest one of them.
Face removal is prone to quick movement.

Work Division:

We worked in pairs to implement project phases as follows:
Ahmed Essam and Omar Ahmed: hand detection and segmentation
Ahmed Zakaria and Abdalla Ezzat: feature extraction and classification

Conclusion:

This report proposes a real-time American Sign Language (ASL) recognition system that goes through 3 phases; hand segmentation which detects skin using Log-Chromaticity Color Space (LCCS) and then removes the face from the frame, feature extraction that uses many image processing algorithms, and classification that uses these features without any type of learning or training to recognize 17 different letters/numbers.

References:

Hardy Francke et all. *Real-Time Hand Gesture Detection and Recognition Using Boosted Classifiers and Active Learning.* 2007.
Nasser H. Dardas and Nicolas D. Georganas, *Fellow, IEEE.* *Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques.* 2011.
Mathias K¨olsch and Matthew Turk. *Robust Hand Detection.* 2004.
Jayshree R. Pansare, Kirti S. Rampurkar, et all. *Real-Time Static Devnagri Sign Language Translation using Histogram.* 2013.
Jayashree R. Pansare et all. *Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background.* 2012.