Sharif University of Technology
Electrical Engineering School

Advanced Neuroscience HW8
Based on Judd & Ehinger & Durand & Torralba 2009

LEARNING TO PREDICT WHERE HUMANS LOOK

Armin Panjehpour
arminpp1379@gmail.com

Supervisor(s): Dr. Ghazizadeh
Sharif University, Tehran, Iran

05/06/2022

Part.1 - Eye Tracking Dataset

Part.1.1 - Samples of Eye Tracking Dataset

Below, you can find some eye tracking data of one the subjects over 9 images:

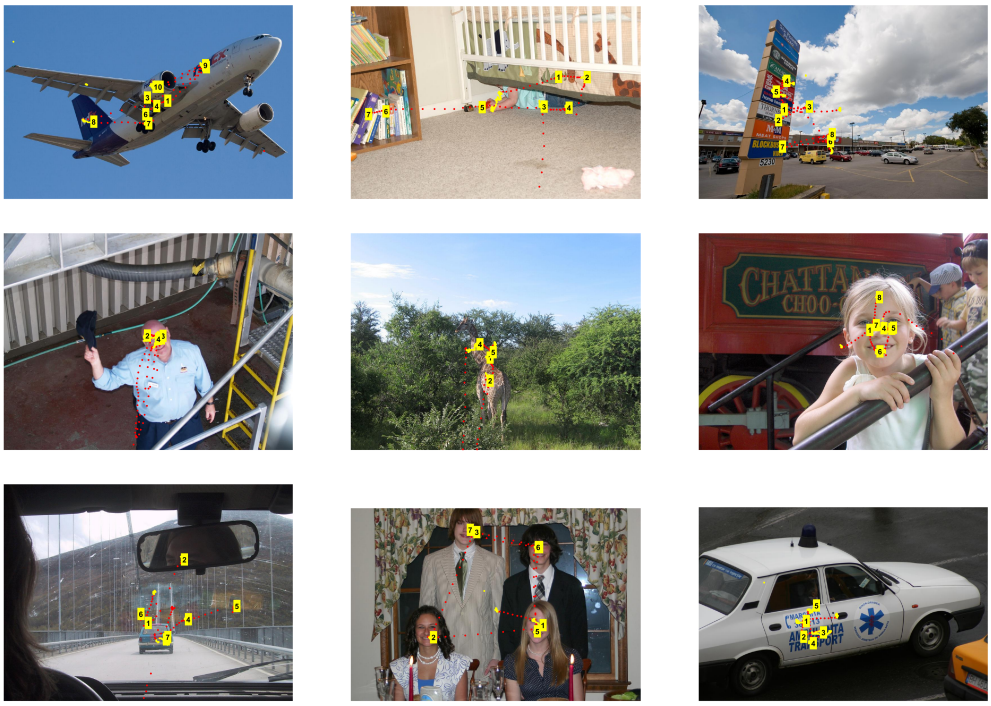

Figure 1: Fixation Maps of Subject "hp"

According to the fixation maps, It seems that features like faces, people, texts, body parts and cars are in interest of the subjects mostly.

Part.2 - Saliency Maps

There is a total of 32 features (in 7 groups) in this assignment which are calculated in "saliency.m" and then linear weighted using the model that the paper has trained. At the end, a saliency map for the input image will be returned. Below you can find 8 saliency maps created for a selected image, each time using one of the features and the 8th one is created using all features:

Figure 2: Selected Image

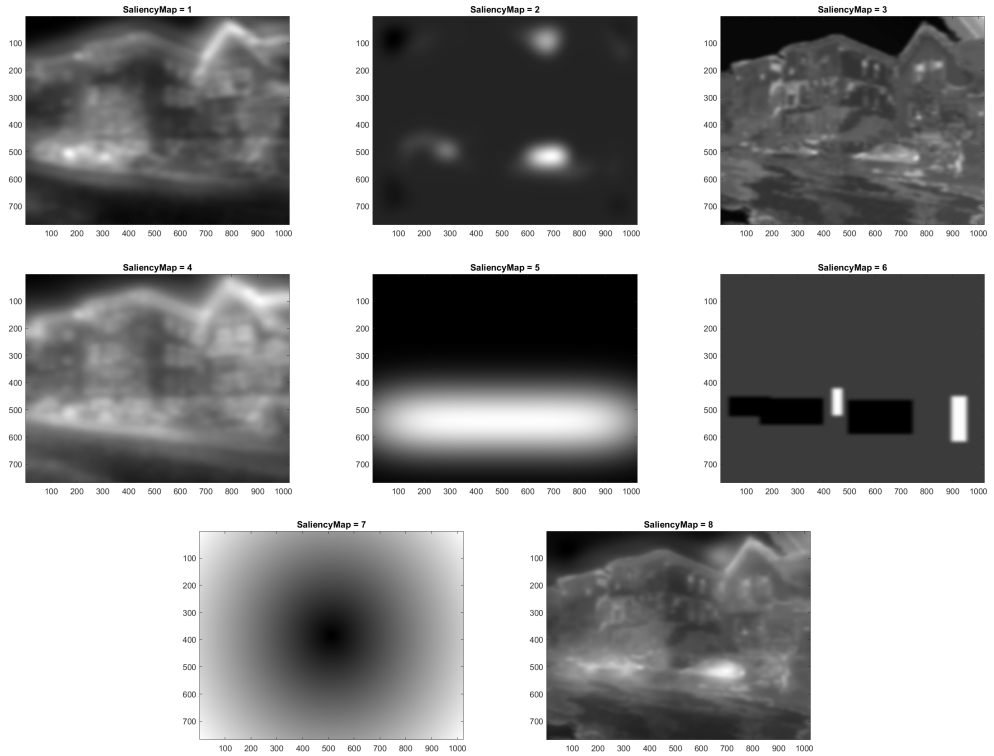

Figure 3: 8 Saliency Maps using each Groups of Features

1. The first group of features is the "Subband Features" which finds coefficients of the subbands of the steerable pyramid.

2. The second group of features is the "Itty Features" which finds three channels of Itti and Koch's saliency model (Color, Intensities & Orientations)

3. The third group of features is the "Color Featuers" which finds the colors, colors' probabilities & 3D histogram of colors.

4. The forth group of features is the "Torallba Featuers" which creates the Torralba saliency.

5. The fifth group of features is the "Horizon Featuers" which finds Horizons in an image.

6. The sixth group of features is the "Object Featuers" which finds faces, cars & people in images.

7. the seventh group of features is "Dist To Center Features" which finds distance of each pixel to the center of image. The reason of presence of this feature is that people usually look at the center of an image more!

Here are the saliency maps for the same image but this one all features expect one is used in creating the saliency map:

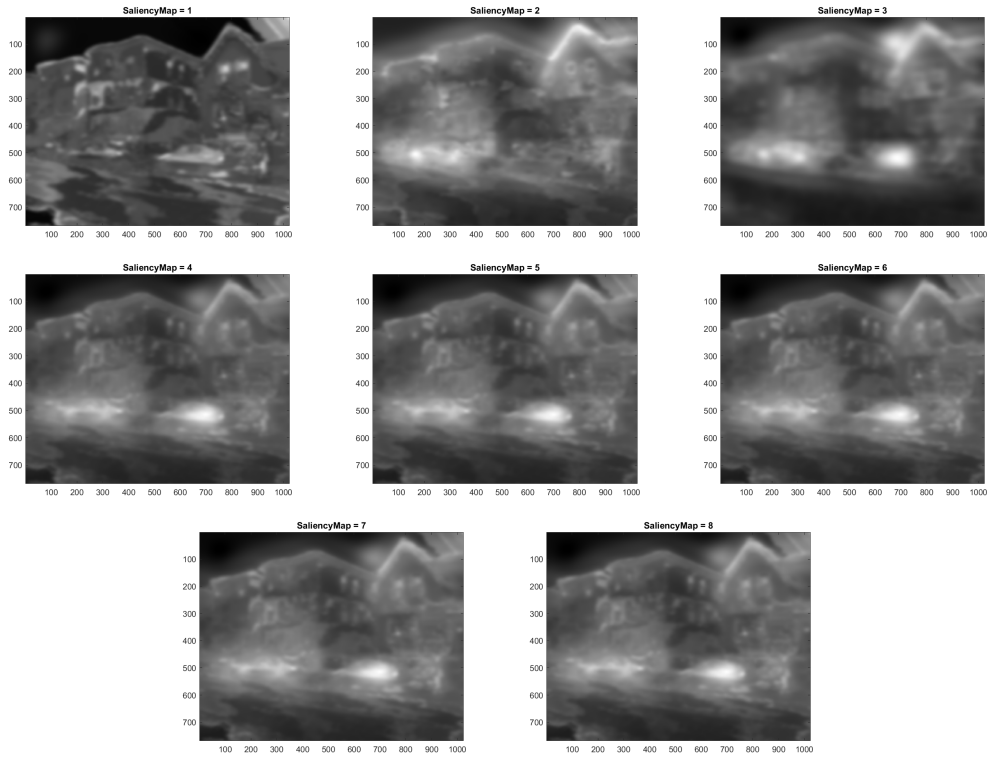

Figure 4: 8 Saliency Maps Excluding just One Group of Features

Part.1.3 - Compare Saliency Maps with Fixation Maps

Here, we create saliency map for each image each time by just keeping one feature (and finally with all features) and calculate the ROC between the saliency maps and the fixation points. In order to compare the bottom-up and top-down effect, the fixation points is divided to two 1.5s periods and ROCs are calculated for both. So finally, we have a ROCs matrix with a size of (1003,15,8,2) which is for all 1003 images, all 15 subjects and 8 different groups of features for two periods. Here we plot the histogram of ROCs for just the first 1.5s period since the behavior of these plots doesn't change a lot for the two periods:

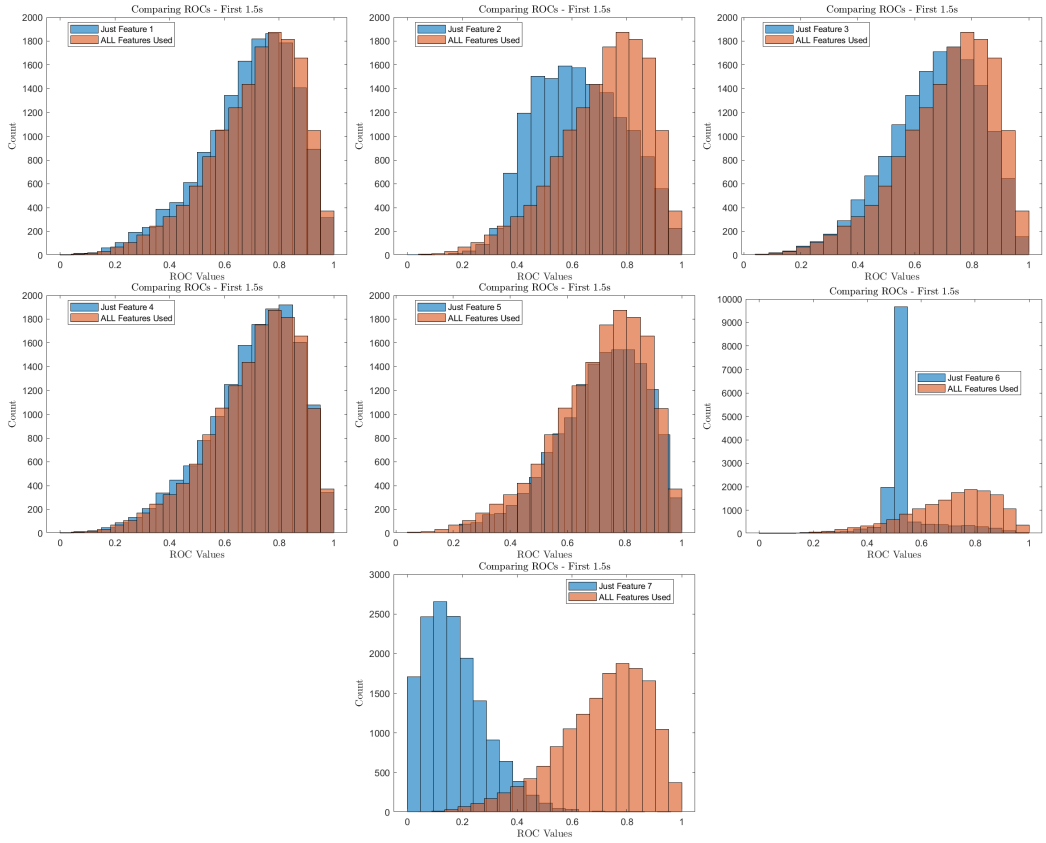

Figure 5: ROC Histogram for different Features Comparing to All Features Available (For the First 1.5s) Over All Images & Subjects

As you can see, the first groups of features have high ROC values which means this group has a big effect. In the opposite, the last feature has low ROC values which indicates that this group of features doesn't explain a lot of ROCs in difference to the others.

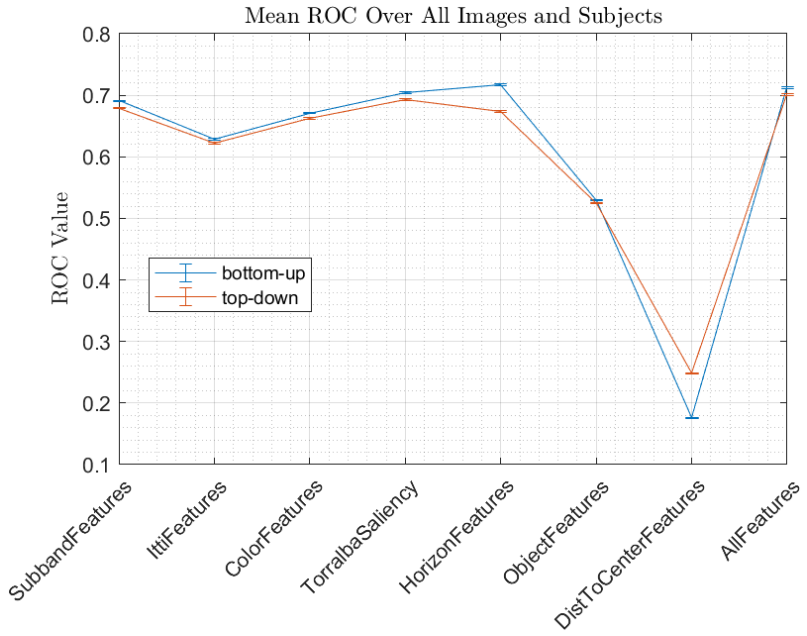

Figure 6: ROC Mean Over All Images & Subjects for First 1.5s and Second 1.5s - Just One Feature Included

As you can see, All features have a high mean ROC expect the last feature.

Now, we calculate ROCs by excluding one feature and again calculate the ROCs to see what is the effect of removing one groups of features, we expect the results to match with the previous ROC plots:

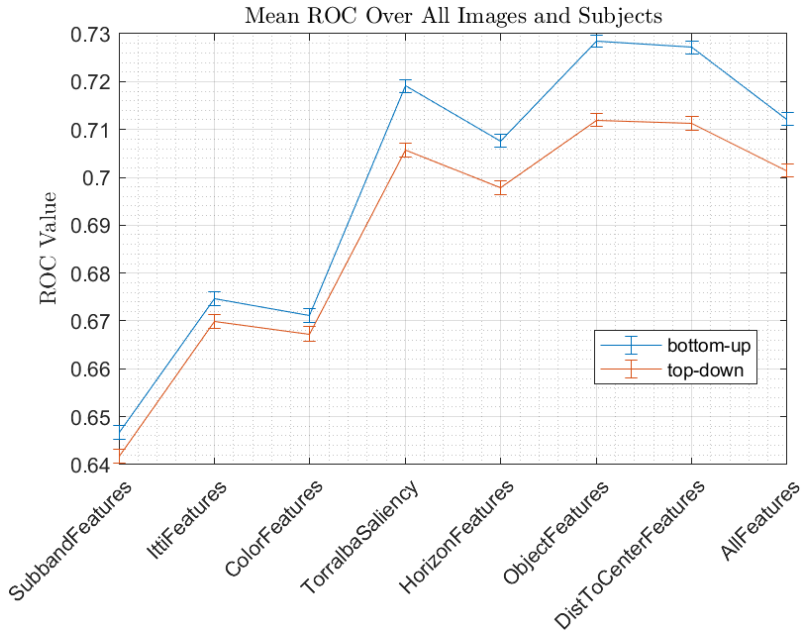

Figure 7: ROC Mean Over All Images & Subjects for First 1.5s and Second 1.5s - Just One Feature Excluded

As you can see, by removing the Subband Features, the mean ROC will decrease significantly and this shows us that the first group of features is very effective. As you can see in Fig.6, ROC values for all bottom-up s are more than the top-down computations but in the two last features. So

we can say all of the features are effect of bottom-up computation but the two last ones. Another thing that we can see, is that the last feature has low effect on ROCs as we saw in the previous results, too. From Fig.7 we can sort the features from their most significant: 1. Subband Features, 2. Color Features, 3. Itti Features, 4. Horizan Features, 5. Torralba Featuers, 6. DistToCenter Features, 7. Object Features. I was expecting to have a high decrease when I remove the object features since they're high level but this doesn't happen.