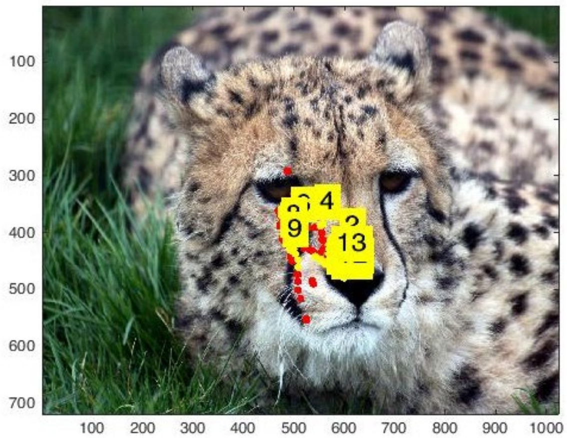
In The Name of God

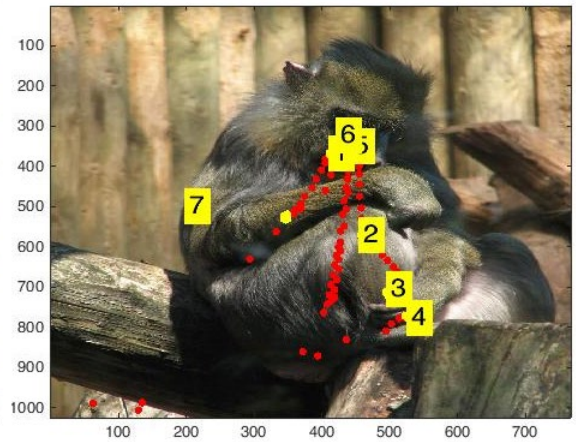
Negin Esmaeilzade 97104034

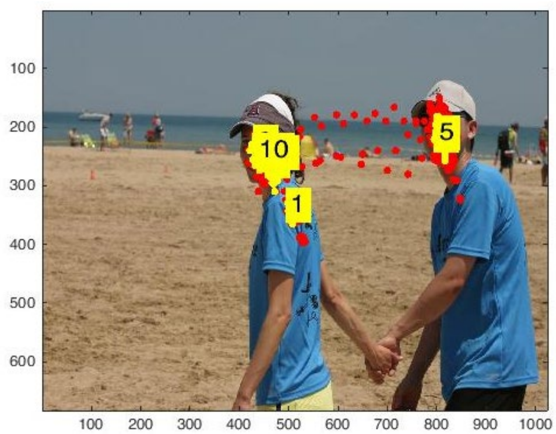
Advanced Neuroscience HW8

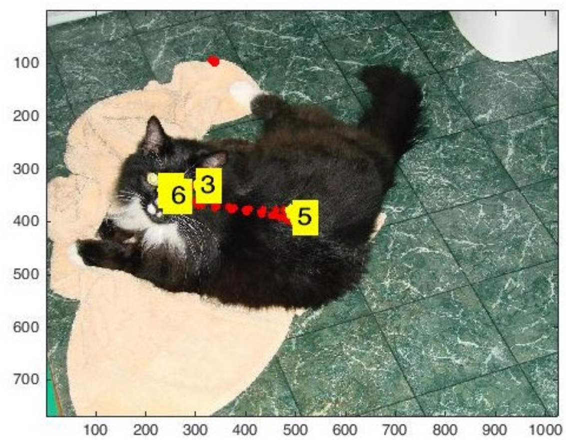
Dr. Ali Ghazizade

Part 1) In this part, we examinate the dataset and toolbox functions. Below are the examples of eye data for a single subject across 4 random selected images. The yellow dots represent the fixation, and the red dots represent the eye data and due to the fixations, we can say that some features like faces, texts and numbers, human and animal body and some parts of cars takes the attentions the most:

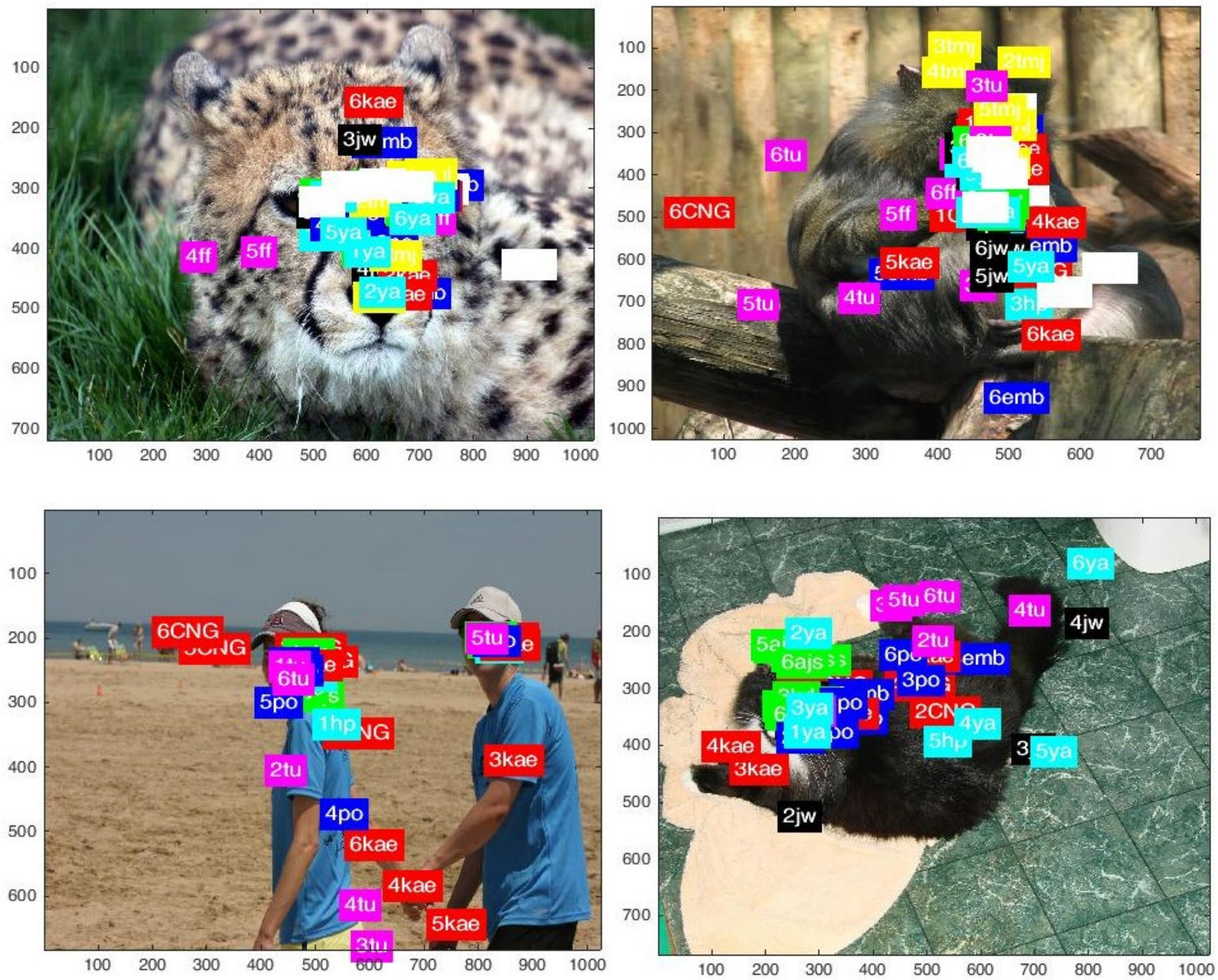




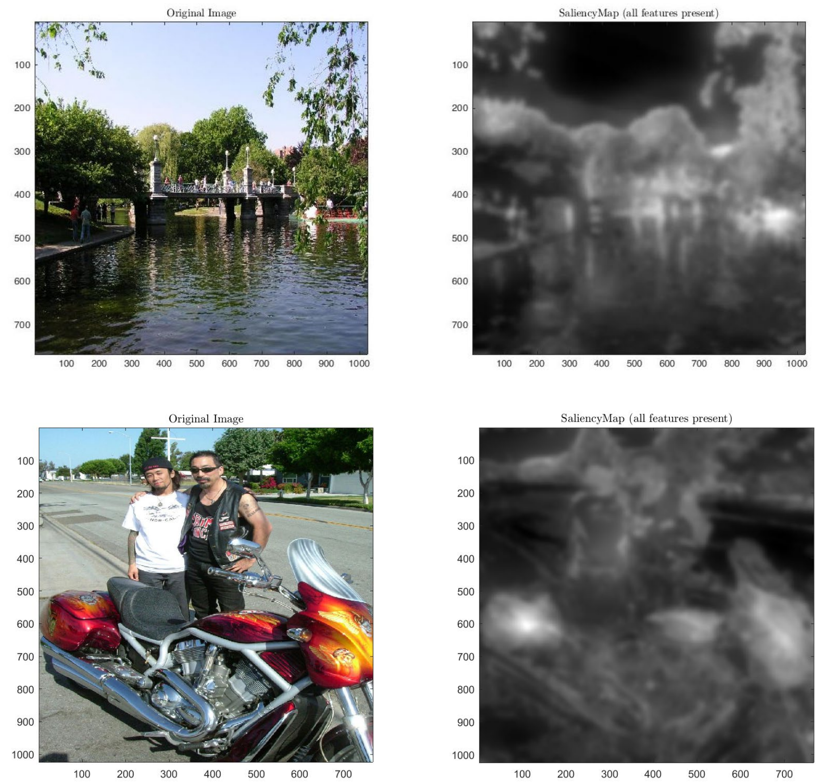




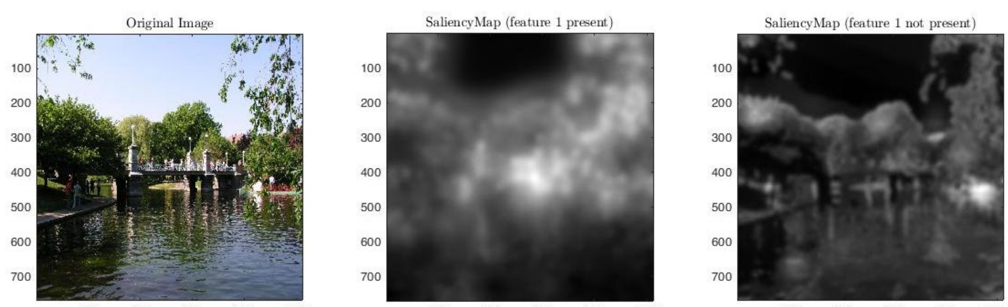
We also test the data across all subjects using the function “showEyeDataAcrossUsers” and here are the examples for the same previous random selected images (the labels on the image are the subject names):

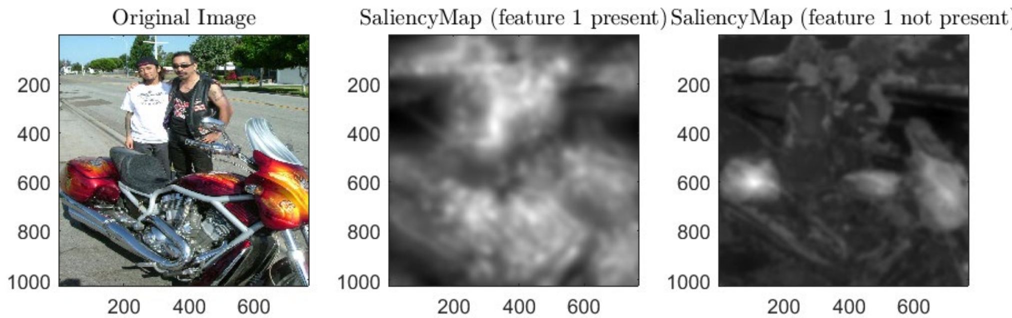


Part 2) In this part, we examine the different saliency models performances. Below is the resulted saliency map using all features:

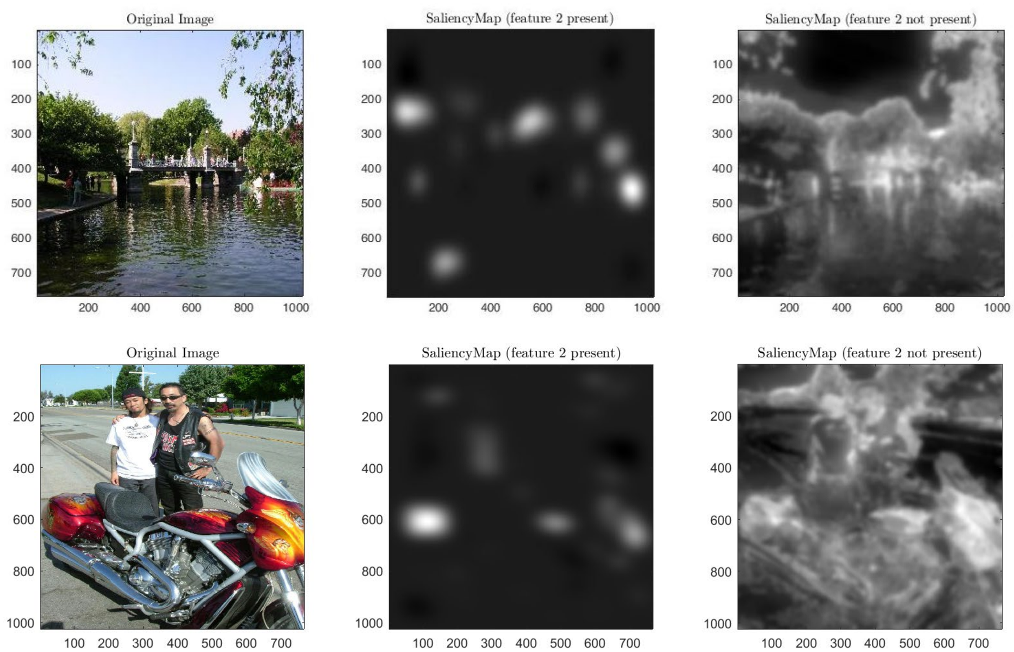


1) **Sub-band features:** These features are sub-band coded features in time-frequency domain. If there exist patterns in a specific band in an image, it will appear into a feature in their wavelet sub-band. Here we can see the result of using this feature and omitting only this feature in the saliency model on a single image:

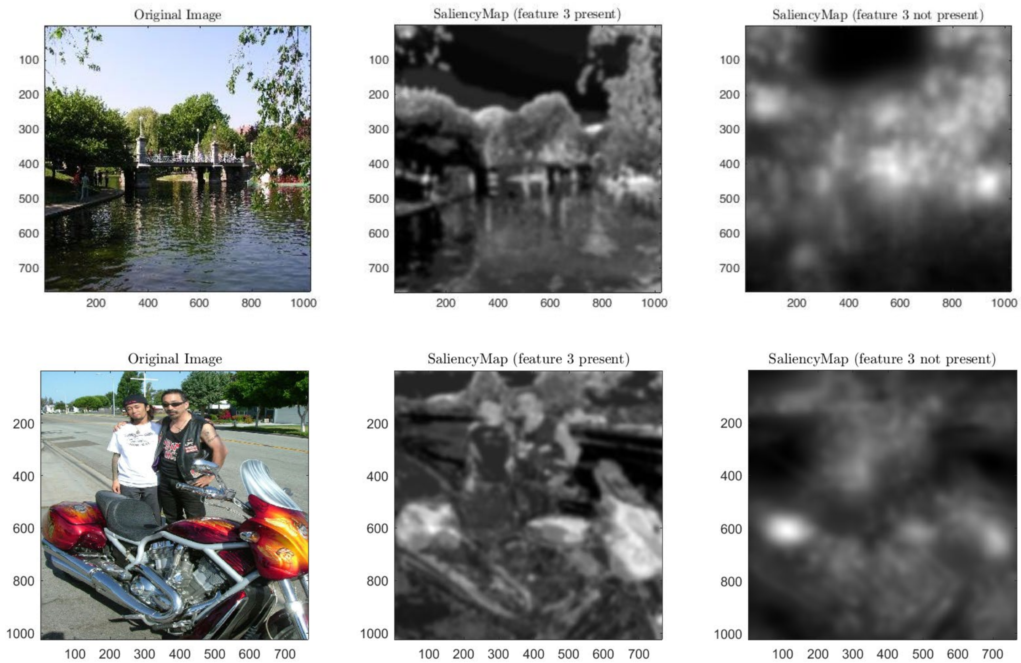




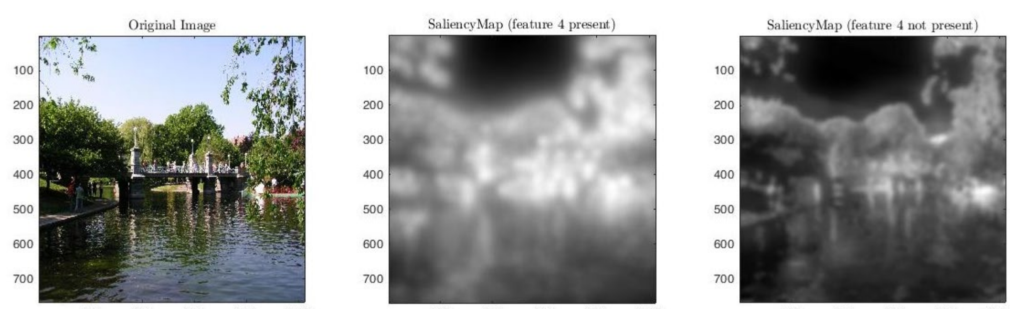
2) **Itti & Koch channels:** The saliency model introduced by Itti and Koch decomposes the image into orientation/color/intensity and then based on winner-takes-all neural model, competes on the next salient candidate. Here we can see the result of using this feature and omitting only this feature in the saliency model on a single image:

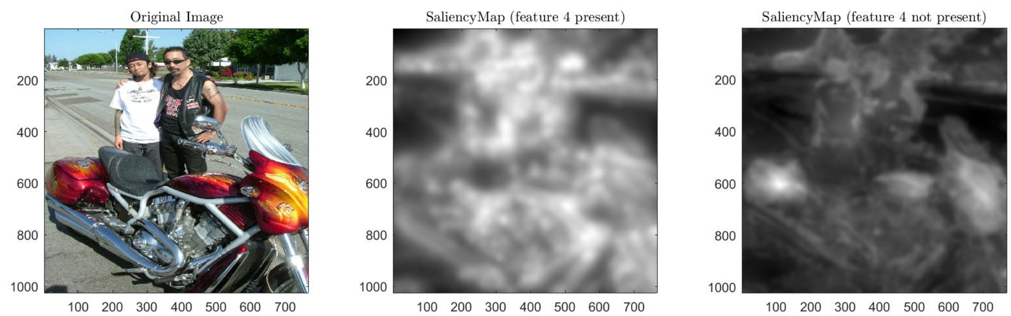


3) **Color features:** In these features the significant factor is color and actually it returns the 3D color histograms. For instance, a red object is more likely to be weighted than a gray object in the saliency map. Here we can see the result of using this feature and omitting only this feature in the saliency model on a single image:

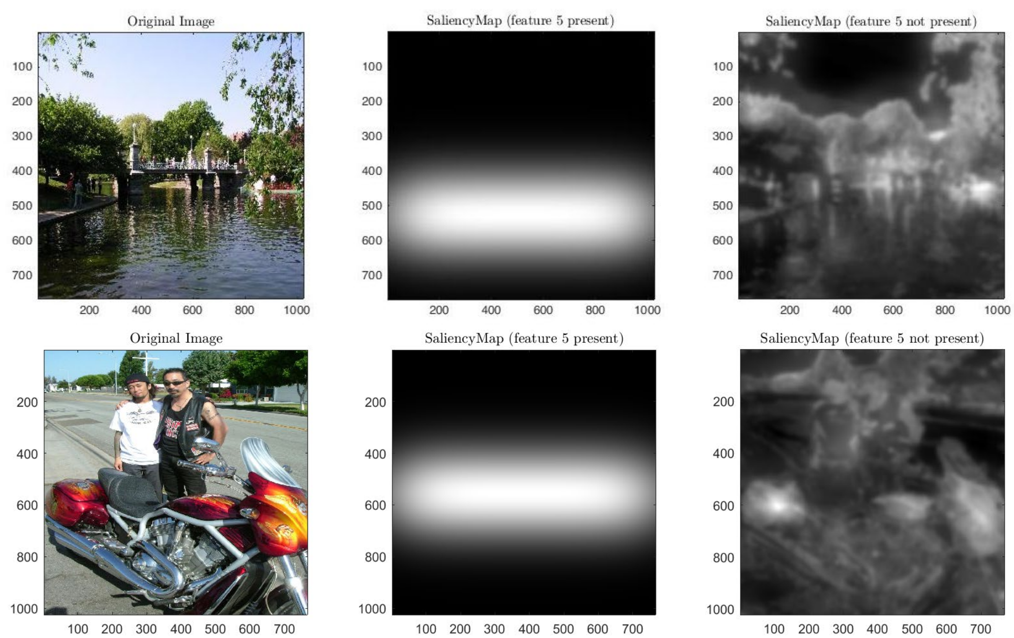


4) **Torralba features:** In fact, Torralba feature extraction method uses some basic feature kernel such as lines or color kernels and with applying multiple kernels gives a map in output that indicates the candidate for each class, so if we define some features as attractive features for visual system, we can define a saliency model based on voting among attractive features, using Torralba. In this example, we can see that Torralba feature extraction is almost same with overall feature extraction, However, the major difference are the details or edges details. Note that Torralba insist on lower computational costs such as memory and time and hard details are not in their method's interests.

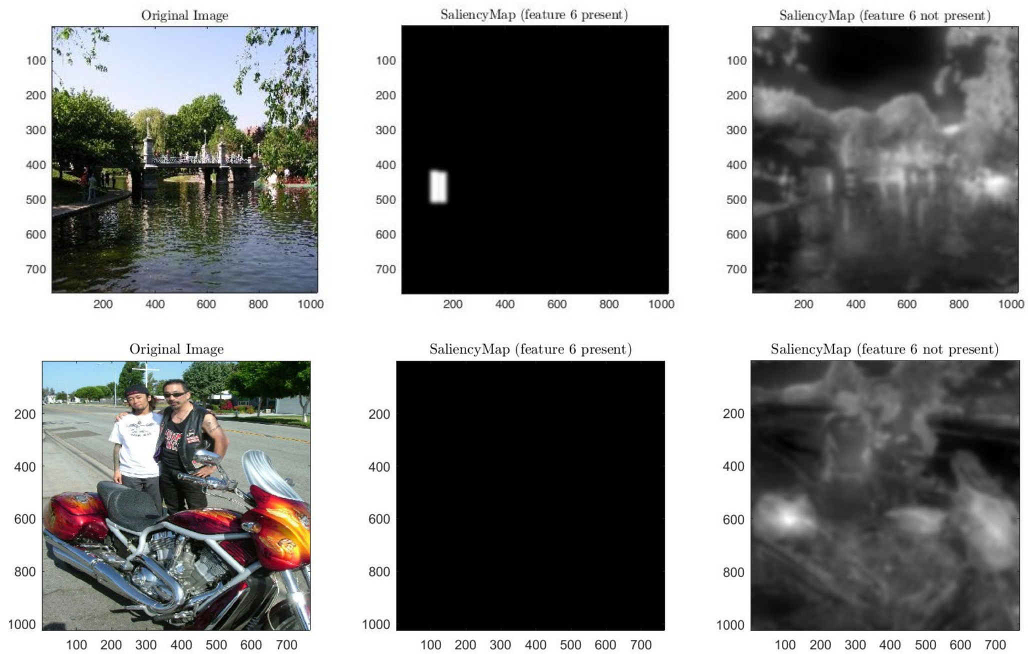




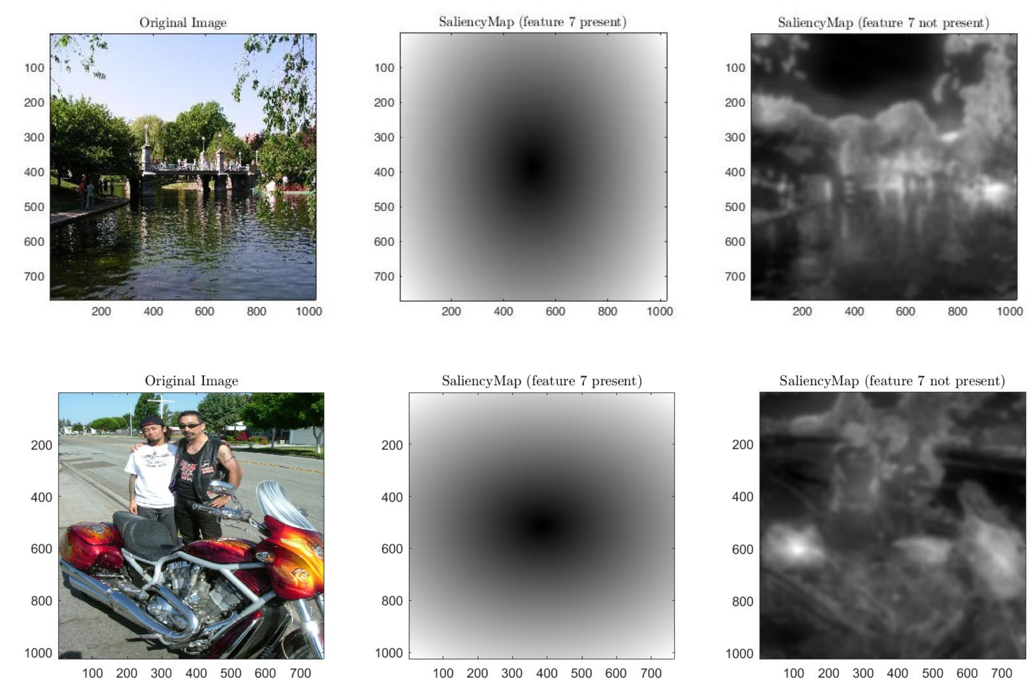
5) **Horizons:** A horizon is a wide line in image that visual systems usually look at that rectangle/line. Horizon detectors do not consume much memory and time and could be useful for saliency map detections. Here we can see the result of using this feature and omitting only this feature in the saliency model on a single image:



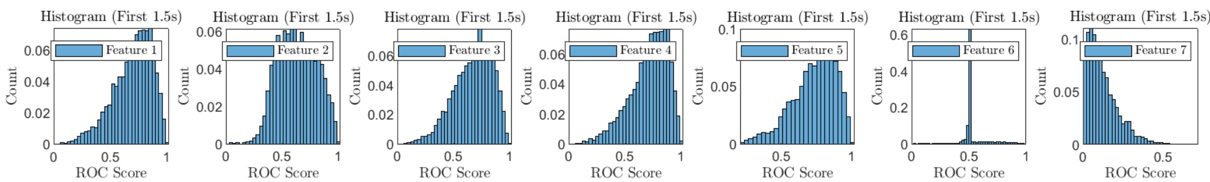
6) **High level objects:** High level objects are complex defined objects such as cars, signs, trees, faces and buildings. Detection of these objects usually require high amount of memory and time (CNNs) and even it is not guaranteed to be in attention. Here we can see the result of using this feature and omitting only this feature in the saliency model on a single image:

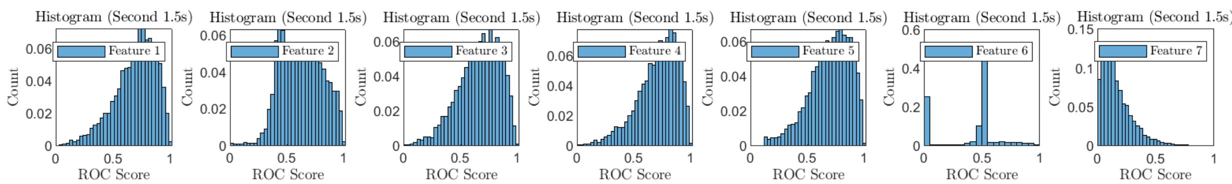


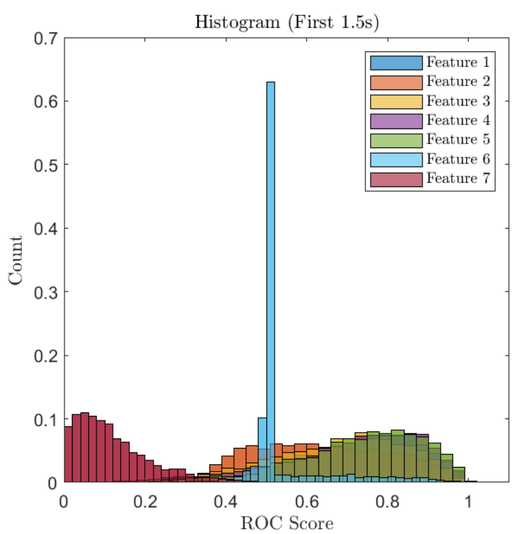
7) **Distance to center:** Simply measures the distance of different image points to the center which gives a simple 2D gaussian. This is important because people usually tend to look the center of their field of view (the image) the most. Here we can see the result of using this feature and omitting only this feature in the saliency model on a single image:

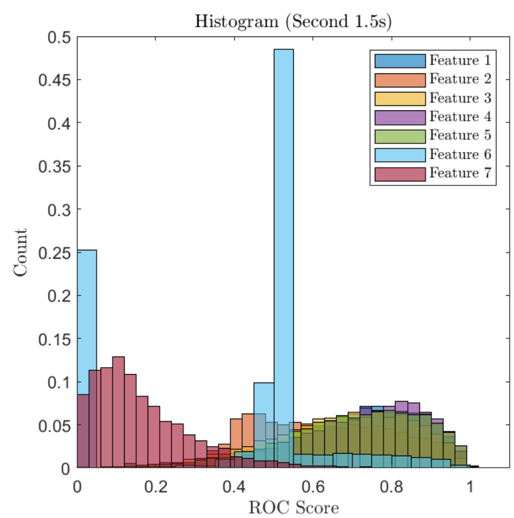


Part 3) To compare the ROC between saliency maps and fixations at both first 1.5s interval and the last 1.5s interval we first divided the eye data into two 1.5s intervals' data. Then for each group we calculated the ROC between the saliency maps using each single feature and the fixations data respectively. These steps are done across all subject and images. Below we can see the resulted histograms of ROC scores, in which the first row represents the first 1.5 data and the next row represent the second 1.5s data:









It can be seen as an example that the first feature has bigger ROC scores on average which means that this feature explains the real fixation the most while its almost the opposite for the last feature. And also, as we can see, for detail-concerned features the ROC scores between the saliency maps and fixations in the first 1.5s data is lower than the next and vise versa. This points to the fact that at the early visual process the basic features are mostly attractive for people and as the time passes people pay attention to the details more. Here is a comparison between the bottom-up (second 1.5s interval) and the top-down (second 1.5s interval) average ROC scores which represents the explanations well: