Accurate Image Segmentation using Gaussian Mixture Model with Saliency Map

Navya and Saksham Rathi

April 2024

1 Introduction

GMM is a flexible tool for image segmentation and image classification. But a major drawback in it is that it doesn't consider spatial information present in the image. The research paper on the basis of which this report has been generated has tried to come up with a method so that we can incorporate spatial information as well.

2 Saliency Map

The saliency map reflects the regions of an image, which can present an interest in the sense of visual per-ception. It highlights the pixels, which can potentially contain information to be used in a more complex image classification scheme.

This is how Saliency Map is calculated: L(f) is log spectrum of image

$$h_n(f) = \boxed{\begin{matrix} \frac{1}{n^2} & \frac{1}{n^2} & \cdots & \frac{1}{n^2} \\ \frac{1}{n^2} & \frac{1}{n^2} & \cdots & \frac{1}{n^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n^2} & \frac{1}{n^2} & \cdots & \frac{1}{n^2} \end{matrix}}$$

A(f) - Average Spectrum, S(x) - Saliency Map, G(x) - Gaussian Blur, F - Fourier Transform

$$A(f) = h_n(f) * L(f)$$

$$R(f) = L(f) - A(f)$$

$$S(x) = G(x) * F^{-1} exp[R(f) + P(f)]^2$$

The following images will show that why saliency map indeed capture salient features.

As we can see that those locations are captured where there is gradient in the image. The pixel intensities at the edge location are very high as compare to other pixels. Informally saying, the saliency map has captured the outline of different objects in the original image.

We also did some experiments on the method of calculating saliency maps. By varying the parameters we were able to exaggerate the boundaries. All this helps at times when we want minute details to be reflected in segmented image. The images below are to illustrate a few of such experiments.

Figure 1: Original Saliency Map

Figure 2: After Exponentiation of Features

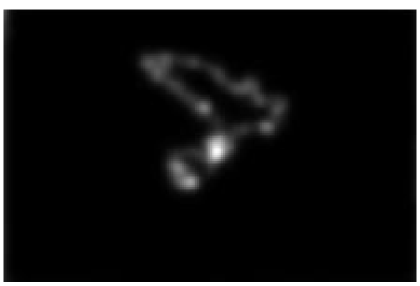

Figure 3: Original Saliency Map

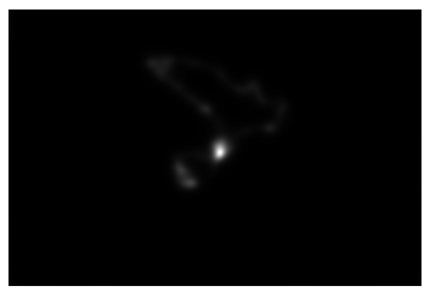

Figure 4: After Exponentiation of Features

3 Improvised Likelihood Function

The new likelihood for the data point x_i to belong to cluster j is given by:

$$\phi(x_i|\theta_j) = \frac{1}{2\pi|\sum_j|} exp[-\frac{1}{2}(x_i - \mu_j)\sum_j^{-1}(x_i - \mu_j)]$$

In a conventional GMM, the pixel value distribution can be discribed by the following equation

$$f(x_i|\Pi, \Theta) = \sum_{j=1}^{L} \pi_j \phi(x_i|\theta_j)$$

Π are parameters constituting prior probabilites and Θ consitutes the parameters of the gaussian distributions.

In our GMM we will incorporate the spatial information by using a weighted neighborhood template for calculating conditional probability by using neighborhood probabilities.

$$f(y_i|\omega) = \sum_{j=1}^{L} \pi_{ij}[\sum_{m \in N_i} \frac{S(x_m)}{R_i} p(y_m|\theta_j)]$$

This is the likelihood for our Saliency Weighted GMM.

One can see that in structure the formula is similiar to conventional GMM in the sense that this likelihood also constitutes summation over different exponentials. That said, we can actually use EM method for finding optimal/approximate values for parameters for which the likelihood for the given image is maximized.

4 Flow Chart Of GMM-SMSI

Step 1. Salience Map Extraction.
1. The image is converted in the spectral domain using FFT. This gives the amplitude spectrum F(f) and the phase spectrum P(f) of the image.
2. The log spectrum representation L(f) is given by the logarithm of F(f).
3. The estimation of the average spectrum A(f).
4. The calculation of the residual value R(f).
5. The generation of the saliency map S(x).

Step 2. GMM incorporating the saliency map as spatial information.
1. The k-means algorithm is first used to initialize the parameters set $\Psi^{(0)}$.
2. Using the saliency map S(x) the EM algorithm is applied for the parameters estimation until convergence. At the end, we get the parameters set $\Psi(c)$.
3. The image pixels are then classified (labeled) based on the highest posterior probability.

Figure 5: Eagle

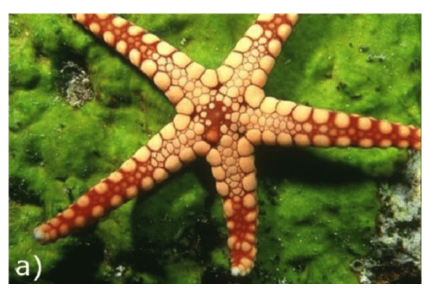

Figure 6: Starfish

5 Dataset

We have performed experiments from on a set of real images from the Berkeley Image Dataset. Some of the images from the dataset are shown in Figure 5 and 6.

6 Experiments

In this section we will showcase the results of the experiments mentioned in the research paper. Out of four experiments, we have found good results in two of them, in the later two experiments we have found results not much better than conventional GMM. But the reason for the same to whatever extent we could think of, has been mentioned in the respective sections.

6.1 Eagle - Tiny Object In Large Background

The original image is shown in Figure 5. The image below is of segmented image outputted by conventional GMM (Image obtained from the research paper).

Figure 7: Conventional GMM

The images below represent our results obatained using GMM-SMSI. On the left is the saliency map and on the right is the segmented result of the GMM-SMSI.

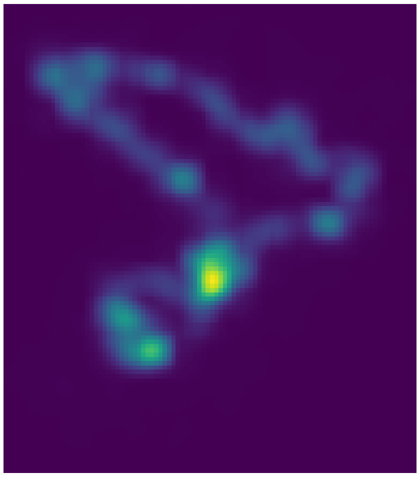

Figure 8: Saliency Map

Figure 9: Segmented Image

Here the result of GMM-SMSI is clearly better than conventional GMM.
1. The feathers are properly separated.
2. The outline of the birds resemble to that in the original image.

One can take a close look at the fourier transform to see that how fantastically it has captured the outline of the birds. Moreover, observe the feather section in the fourier transform.

But what result was research paper claimed? You can see below.

Figure 10: GMM-SMSI

Where are we lacking ? The middle thing that is properly segmented in there GMM-SMSI but not in ours. To get this we tried various initialization and changed the method to generate the fourier transform. The result of one such experiment is shown below. Although the result is not clean but it does to some extent preserve the middle part.

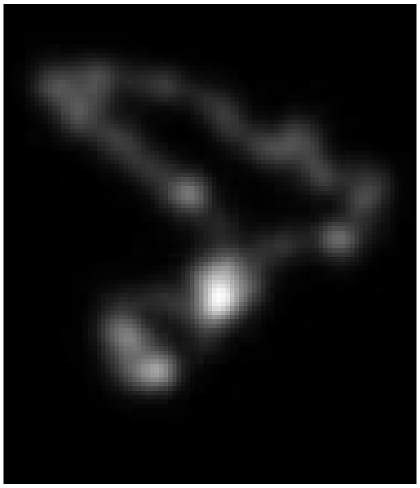

Figure 11: Saliency Map

Figure 12: Segmented Image

6.2 Minute Details in Images

Figure 13: Original

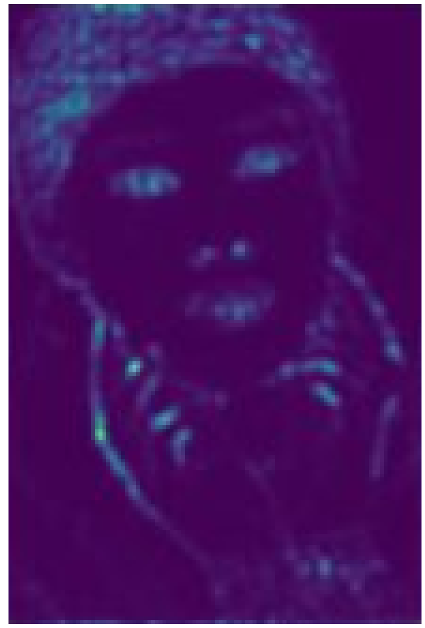

Figure 14: Saliency Map

This experiment was to test that how effectively we capture the minute details present in the image.

Here we would like to point out the details that have been captured by saliency map:
1. The part of cap which overlaps with forehead. The details of the triangles is captured.
2. The eyebrows(look closely at the saliency map!) are captured.
3. The details of the sleeves

Let's see the result obtained by using the conventional GMM.

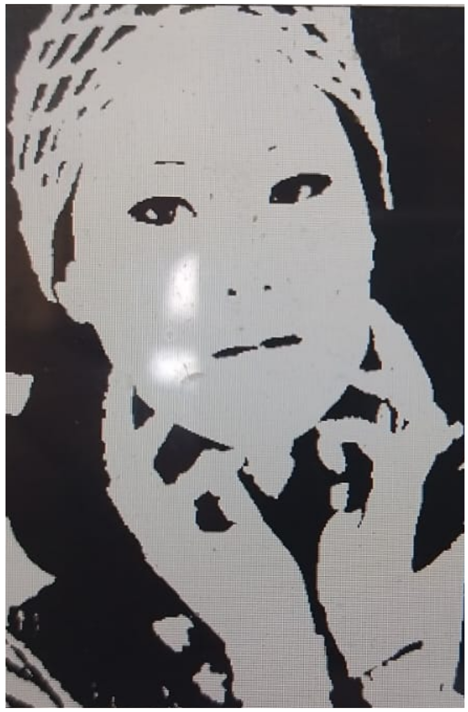

Figure 15: GMM

There are a few problems with this segmentation.

1. The cap and the forehead are not separated anymore.
2. Eyebrows are not clear anymore.
3. The linings in the sleeves are not preserved.

Let's have a look at the result obtained using gmm-smsi.

Figure 16: GMM-SMSI

We can see that the minute details of images have been handled now in relatively better.
1. Although minutely but some of the cap which got merged with forehead is now visible separately with white blogs in between.
2. Left eyebrow is very clear now, although not much improvement has happened for the right eyebrow.
3. Lining in the sleeves are fantastically preserved.

We can see that how with the help of saliency map, we are able to preserve minute details comparatively much better.

6.3 Starfish

The Original Image and the Saliency Map obtained are as below.

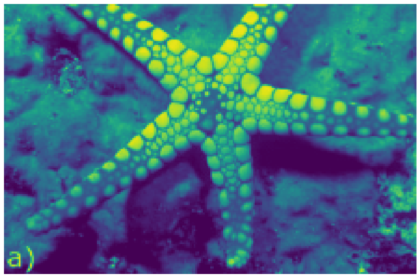

Figure 17: Original

Figure 18: Saliency Map

Here we would like to point out the detail that have been captured by saliency map:
1. The circles of the startfish has been captured properly.

The obtained result are as follow.
Let's see the result obtained by using the conventional GMM.

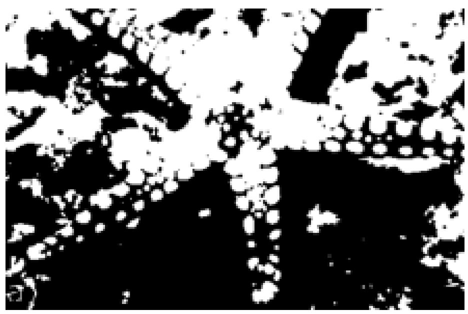

Figure 19: GMM

1. We can see that the segmented image has lost the information of the circular blogs on the surface of starfish.
2. Much of the background pixels has been clustered together with the pixels of circular discs.

Let's see the result obtained by gmm-smsi.

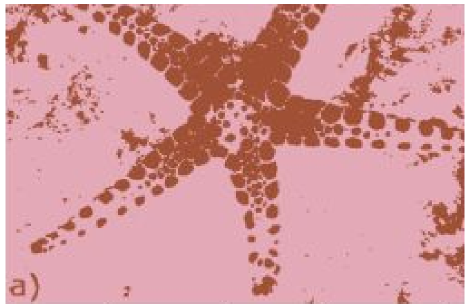

Figure 20: GMM-SMSI

We can see now that
1. The circlular blogs on the starfish surface are now properly captured.
2. The background pixels are mostly in a different cluster that the cirucluar discs pixels.

6.4 Building

Note: the result of this experiment are not quite good and are not consistent with that of the research paper. We have anyways tried to think of the possible reasons for it and has mentioned them.

The Original Image looks like as below.

Figure 21: Original

Corresponding Saliency Map.

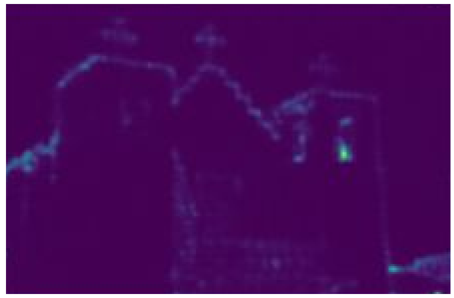

Figure 22: Saliency Map

Features Captured:
1. Left window(very minutely)
2. Right window(clearly)
3. Door details

Result obtained by conventional GMM:

Figure 23: GMM

Result obtained by GMM-SMSI:

Figure 24: GMM-SMSI

First of all, we are not getting the door part by any means. Why the brightly white part will get clubbed with dark part, while the less white-greyish church will be in another cluster. It does not seem to make any sense. But if we see the formula, it indeed can be obtained if variance of the black cluster is high as compare to white cluster.

Secondly, the left window which has been not preserved, it is perhaps because although there is an edge but the window itself has same pixel intensites as the usual church along with the slightly darker part.

The research paper claims that with gmm-smsi the dark pixel internsites will get clubbed with the white background which seems to be the reason for both the above differences. But we are not claiming anything against research paper, perhaps it is because of some issues with our implementation. One reason could be that since our method of saliency map calculation is slightly different that might be causing this issue.

6.5 Result Comparison

For the four experiments, we have shown above, we also ran using GMM with MRF. The resukls are shown below (clearly, worse than GMM-SMSI):

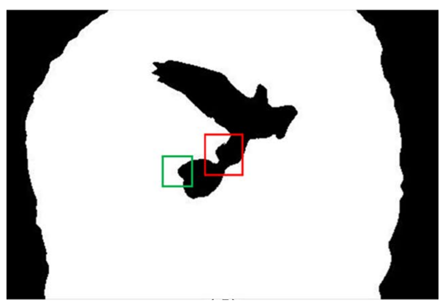

Figure 25: Eagle

Figure 26: Starfish

Figure 27: Church

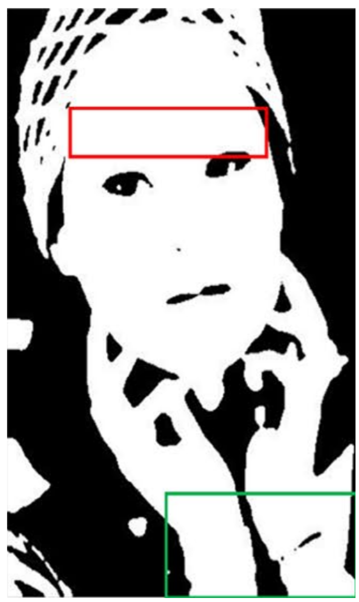

Figure 28: Girl

6.6 Medical Images

We chose two medical images. The results of GMM-SMSI are as follows on them: We have also increased the number of clusters from 2 to 3. (The brain patches and other portions are clearly segmented. Therefore our method can be used on medical images too.)

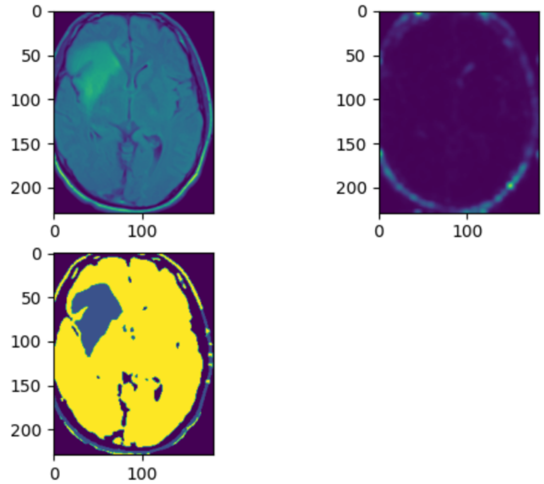

Figure 29: GMM-SMSI

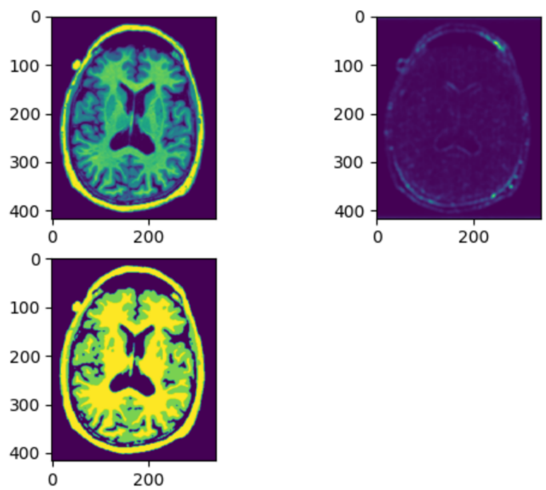

Figure 30: GMM-SMSI

7 Conclusion

The project as already mentioned was a attempt to regenerate the results of the research paper as already mentioned. After overall experimentation we have realized that the gmm-smsi is indeed a better approach. Althouth the research paper has also done PR value comparison with many other fields but since our results were not that good as claimed by the paper, we have just done simple visual anaylysis with conventional gmm. But it is sure that saliency map loaded GMM is much better than conventional GMM.

8 References

1. Research paper has been taken from this link - https://link.springer.com/article/10.1007/s10044-017-0672-1

2. A part of the code was taken from - https://github.com/SrikanthAmudala/GaussainDistribution