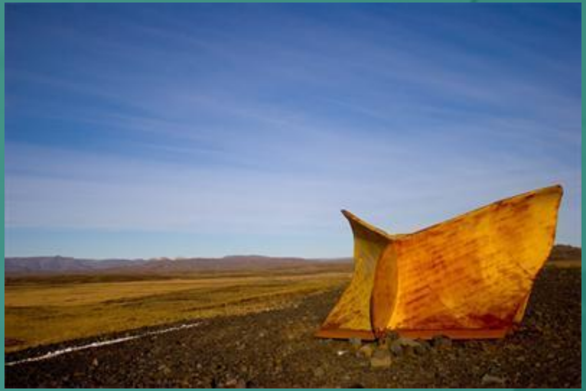
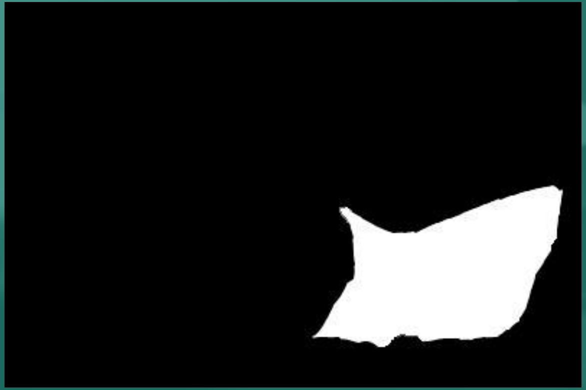
Saliency Filters: Contrast Based Filtering for Salient Region Detection

Avinash Vadlamudi
Sravan Chandra
Kamineni Vikas





Saliency

- ➢ Saliency is what stands out to you and how you are able to quickly focus on the most relevant parts of what you see.
- ➢ Saliency detection, essentially, can be used in any area in which you’re trying to automate the process of understanding what stands out in an image.

Approach-1:

We have implemented the paper:

Saliency Filters: Contrast Based Filtering for Salient Region Detection.

CVPR 2012

Abstraction:

- ➔ For the image abstraction, we use an adaptation of SLIC superpixel to abstract the image into perceptually uniform regions.

Uniqueness:

$$U_i = \sum_{j=1}^{N} \|\mathbf{c}_i - \mathbf{c}_j\|^2 \cdot \underbrace{w(\mathbf{p}_i, \mathbf{p}_j)}_{w_{ij}^{(p)}} .$$

Distribution:

$$D_i = \sum_{j=1}^{N} \|\mathbf{p}_j - \mu_i\|^2 \underbrace{w(\mathbf{c}_i, \mathbf{c}_j)}_{w_{ij}^{(c)}},$$

Saliency Assignment

- We start by normalizing both uniqueness U_i and distribution D_i to the range [0..1].
- We compute Saliency value S_i for each element using,

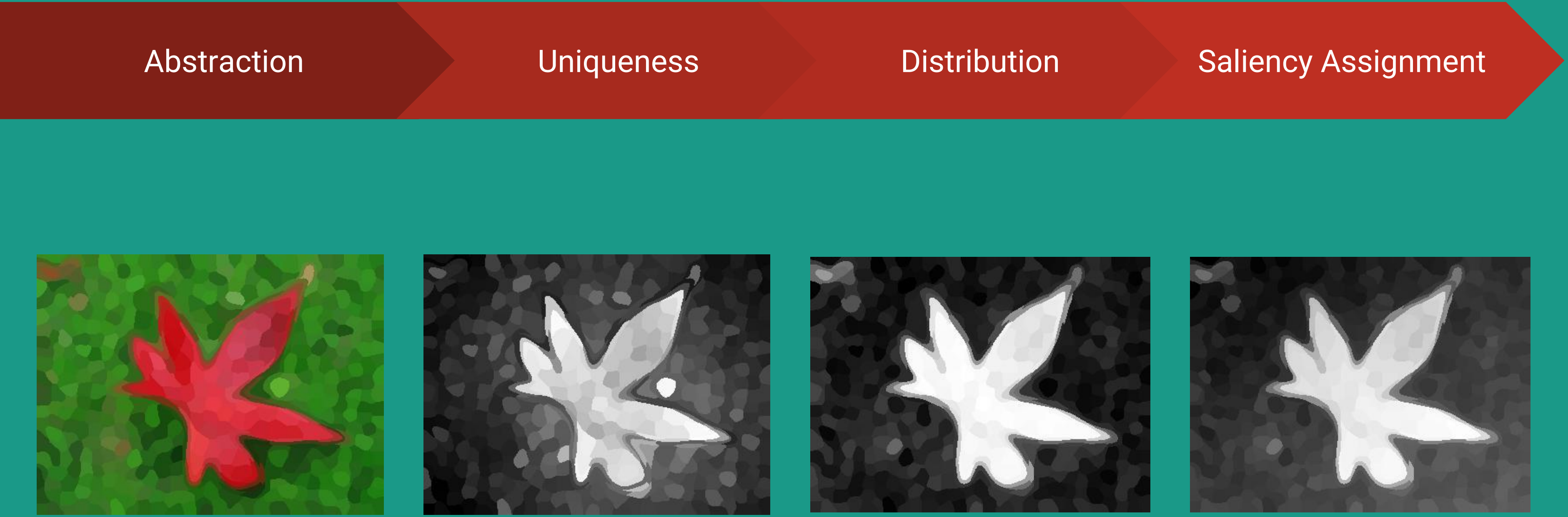
$$S_i = U_i \cdot \exp(-k \cdot D_i)$$

- We used $k = 6$ as the scaling factor for the exponential.

- As the final step, we need to assign a final saliency value to each image pixel, which can be interpreted as an upsampling of the per-element saliency S_i .
- We define the saliency $\tilde{S}_i$ of a pixel as a weighted linear combination of the saliency S_j of its surrounding image elements.

$$\tilde{S}_i = \sum_{j=1}^{N} w_{ij} S_j .$$

- By choosing a Gaussian weight as **$w_{ij} = 1/Z_i \times \exp(- 1/2 \times (\alpha||c_i - c_j||^2 + \beta||p_i - p_j||^2))$** we can ensure the up-sampling process is both local and color sensitive.
- Here α and β are parameters controlling the sensitivity to color and position.



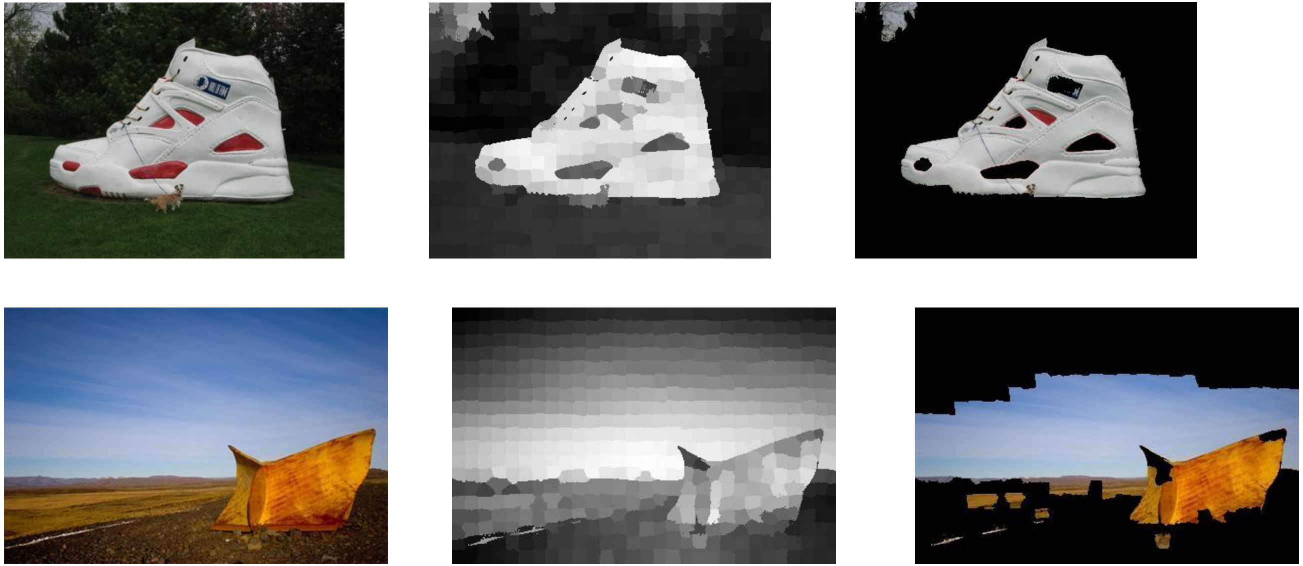
Abstraction

Uniqueness

Distribution

Saliency Assignment

Results:

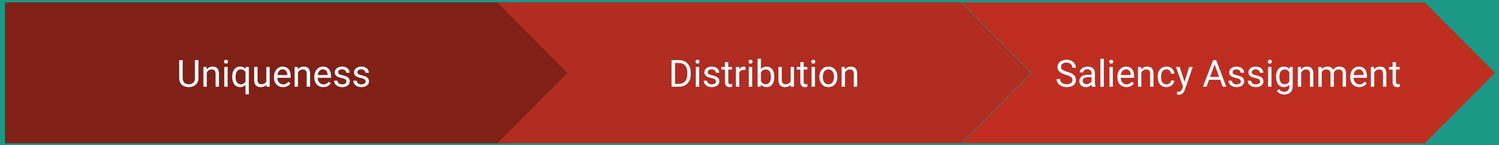


Extension:

- As, In the above Approach, we noticed that some of the results are not good, So we tried to perform on pixel level instead of superpixels.

$$
\begin{aligned}
U_i &= \sum_{j=1}^{N} \|\mathbf{c}_i - \mathbf{c}_j\|^2 w_{ij}^{(p)} \\
&= \mathbf{c}_i^2 \underbrace{\sum_{j=1}^{N} w_{ij}^{(p)}}_{1} - 2\mathbf{c}_i \underbrace{\sum_{j=1}^{N} \mathbf{c}_j w_{ij}^{(p)}}_{\text{blur } c_j} + \underbrace{\sum_{j=1}^{N} \mathbf{c}_j^2 w_{ij}^{(p)}}_{\text{blur } c_j^2}
\end{aligned}
$$

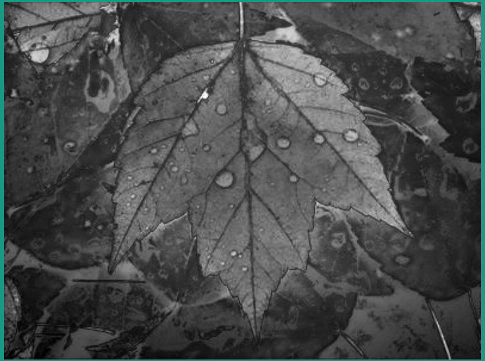
$$
\begin{aligned}
D_i &= \sum_{j=1}^{N} \|\mathbf{p}_j - \mu_i\|^2 w_{ij}^{(c)} \\
&= \sum_{j=1}^{N} \mathbf{p}_j^2 w_{ij}^{(c)} - 2\mu_i \underbrace{\sum_{j=1}^{N} \mathbf{p}_j w_{ij}^{(c)}}_{\mu_i} + {\mu_i}^2 \underbrace{\sum_{j=1}^{N} w_{ij}^{(c)}}_{1} \\
&= \underbrace{\sum_{j=1}^{N} \mathbf{p}_j^2 w_{ij}^{(c)}}_{\text{blur } \mathbf{p}_j^2} - \underbrace{\mu_i}_{\text{blur } \mathbf{p}_j}{}^2 .
\end{aligned}
$$

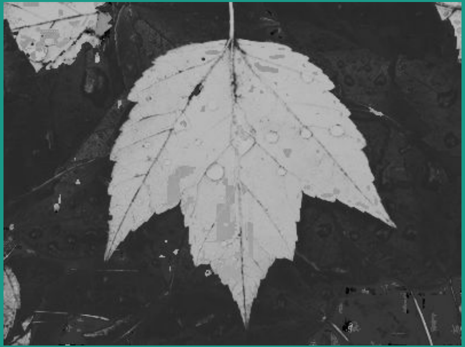


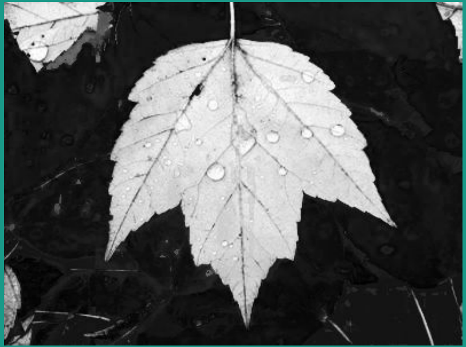
Uniqueness

Distribution

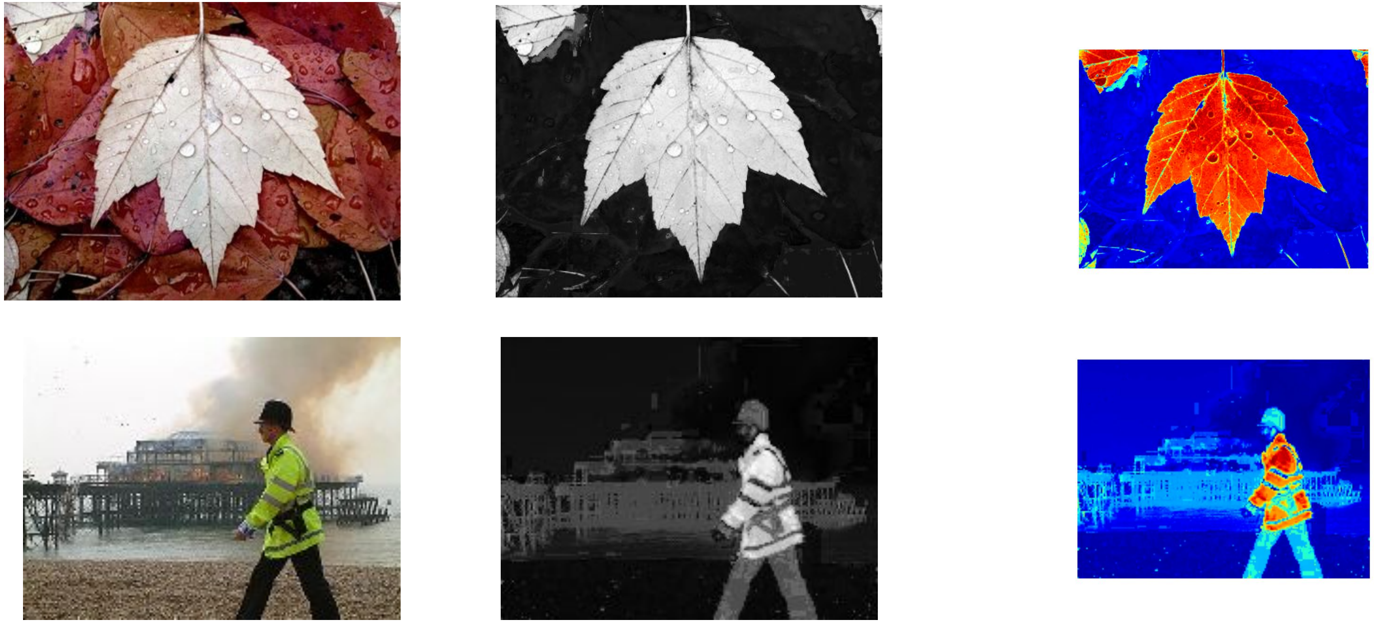
Saliency Assignment



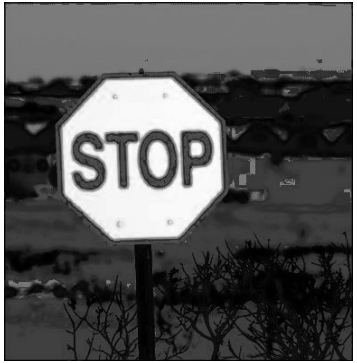


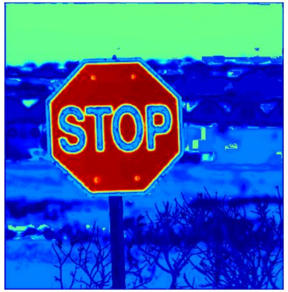


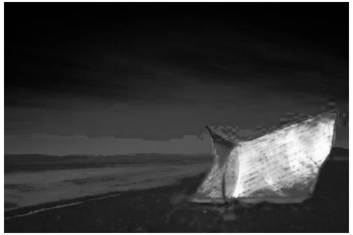
Results:

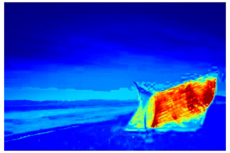


STOP


STOP


STOP





Approach-2: Energy Minimization

- Saliency Optimization from Robust Background Detection (CVPR 2014)
- Here, We solve the problem from the fact that background objects are more linked to the boundary where as salient objects are not much linked to the boundary of the image.
- The image is first abstracted as a set of nearly regular superpixels using the SLIC method.

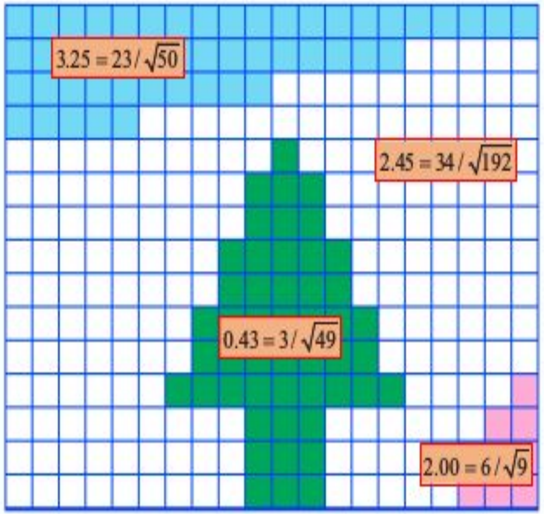
- We then construct an undirected weighted graph by connecting all adjacent superpixels (p, q) and assigning their weight $d_{app}(p, q)$ as the Euclidean distance between their average colors in the CIELab color space.
- We further add edges between any two boundary superpixels. This is useful when a physically connected background region is separated due to occlusion of foreground objects.
- The geodesic distance between any two superpixels $d_{geo}(p, q)$ is defined as the accumulated edge weights along their shortest path on the graph.

- So, we give the measure to the saliency through following idea:
  - Here, we use the below ratio to measure saliency,

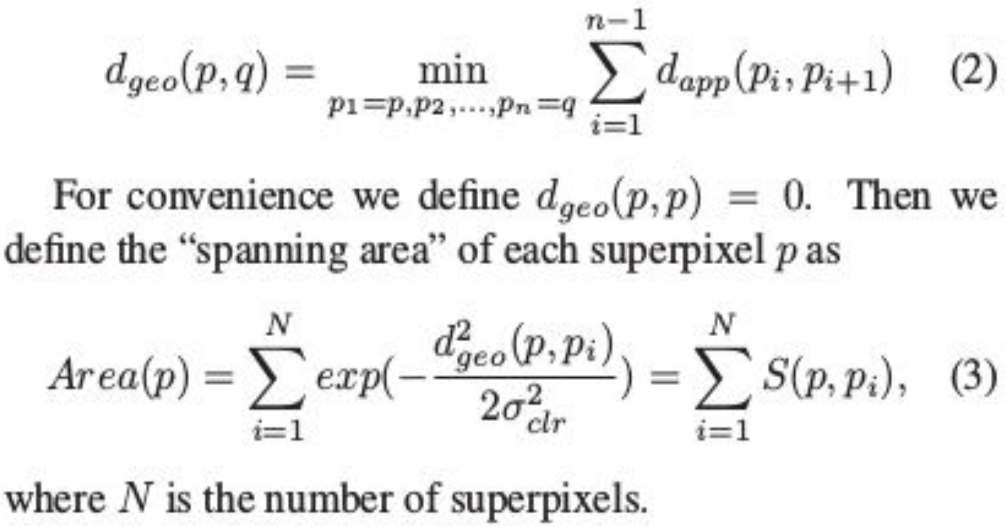
$$BndCon(p) = \frac{Len_{bnd}(p)}{\sqrt{Area(p)}}$$

Where, length of a superpixel p is the length along the boundary.

$$Len_{bnd}(p) = \sum_{i=1}^{N} S(p, p_i) \cdot \delta(p_i \in Bnd)$$



And Area(p) is the area of the superpixel p:



$$d_{geo}(p,q) = \min_{p_1=p,p_2,\ldots,p_n=q} \sum_{i=1}^{n-1} d_{app}(p_i,p_{i+1}) \quad (2)$$

For convenience we define $d_{geo}(p,p) = 0$. Then we define the "spanning area" of each superpixel p as

$$Area(p) = \sum_{i=1}^{N} exp(-\frac{d^2_{geo}(p,p_i)}{2\sigma^2_{clr}}) = \sum_{i=1}^{N} S(p,p_i), \quad (3)$$

where N is the number of superpixels.

We minimize the below non-linear function,

$$\underbrace{\sum_{i=1}^{N} w_i^{bg} s_i^2}_{\text{background}} + \underbrace{\sum_{i=1}^{N} w_i^{fg} (s_i - 1)^2}_{\text{foreground}} + \underbrace{\sum_{i,j} w_{ij} (s_i - s_j)^2}_{\text{smoothness}}$$

Where,

$$w_i^{bg} = 1 - exp(-\frac{BndCon^2(p_i)}{2\sigma^2_{bndCon}})$$

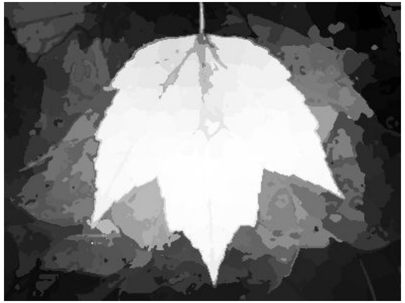
$$w_{ij} = exp(-\frac{d^2_{app}(p_i, p_j)}{2\sigma^2_{clr}}) + \mu$$

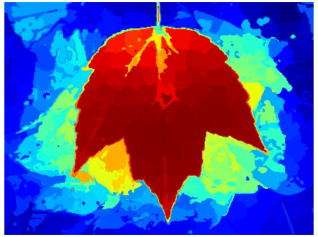
$$wCtr(p) = \sum_{i=1}^{N} d_{app}(p, p_i) w_{spa}(p, p_i) w_i^{bg}$$

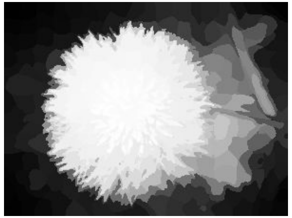
$$w_{spa}(p, p_i) = exp(-\frac{d^2_{spa}(p,p_i)}{2\sigma^2_{spa}}). \quad d_{spa}(p, p_i)$$

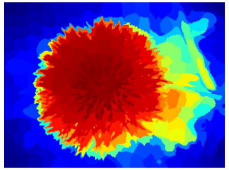
Here, wCtr is background weighted contrast, which is equivalent to w^{fg} .

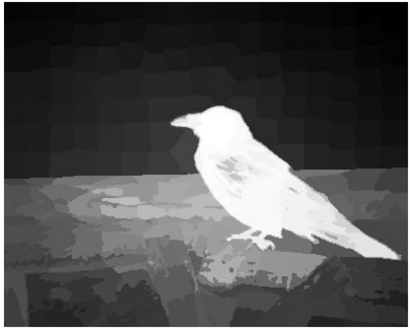
Results:

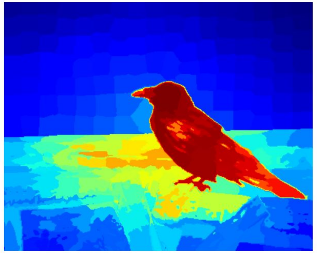


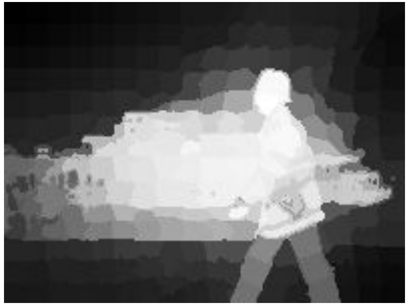


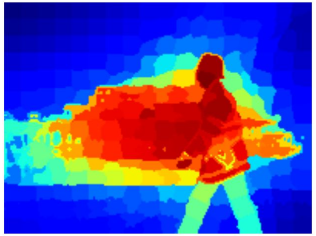








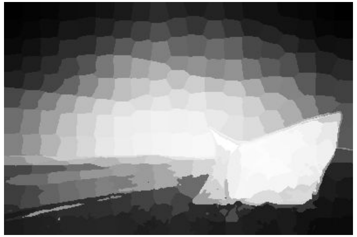


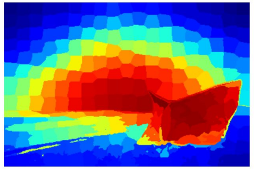


STOP

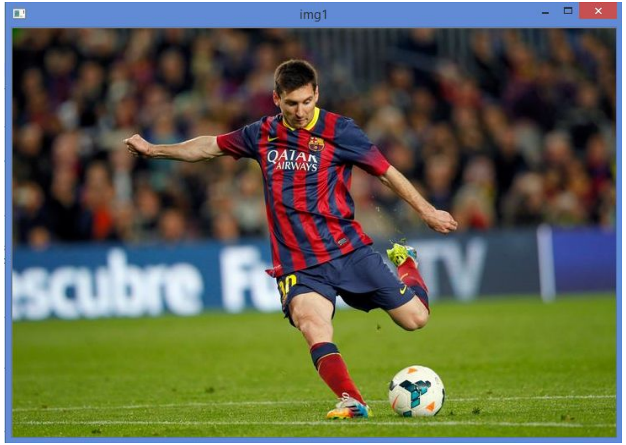
STOP

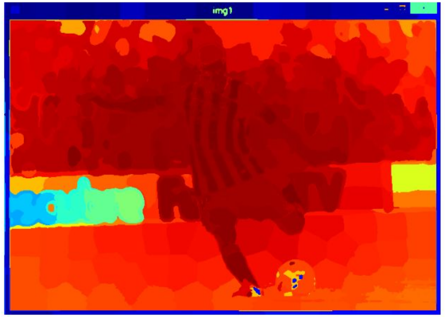
STOP

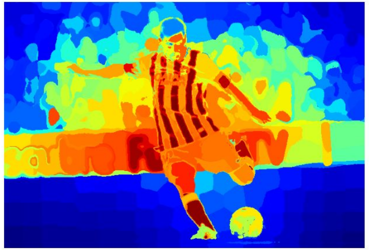




Failures of Approach-2:







- In the first image, it contains the blue border, so using this approach, it considers that entire image is distinct with border, so everything is considered as salient.
- As we can see, the second image doesn’t have blue border, so we are getting good detection of features.

Analysis:

	Approach-1	Extension
Abstraction	0.55301452 s	--------
Uniqueness	0.009122849 s	0.1385328 s
Distribution	0.01104879 s	77.074188 s
Saliency	0.01099658 s	6.2409470 s

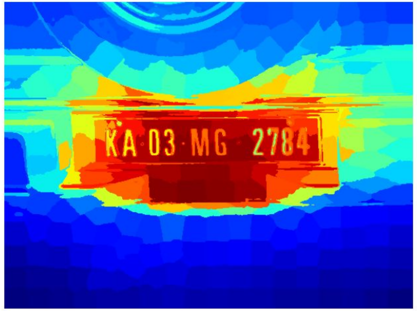
	Approach-1	Extension	Approach-2
Total Time	0.58 s	83.3974419 s	7.242528

SLIC	Using Inbuilt	Coded (For 5 iter’s)
Time Taken	0.2575693 s	21.516971 s

Applications

- ➔ Incorporate higher-level features such as face detectors, number plate detection.
- ➔ Automatic cropping of Images (Seam Carving).
- ➔ Thumbnail generation.

Number Plate Detection



Citations

- Main Paper:- http://www.philkr.net/papers/2012-06-01-cvpr/2012-06-01-cvpr.pdf
- Second Paper:- https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6909756
- Dataset: http://saliencydetection.net/dut-omron/