How I understand metrics

We will try to understand metrics that use both ground truth and the saliency map.

Figure 1: Image with its ground truth (top right) and a computed saliency map (bottom right)

1) Similarity (SIM)

It’s a simple comparison between two saliency maps where the result is the computation of the minimum on each pixel.

$$\sum_i P_i = \sum_i Q_i = 1$$

$$SIM(P, Q) = \sum_i \min(P_i, Q_i)$$

Visually we can observe on the computed images where both networks show positive value. It means that only true positive will appear. False positive and lake of saliency will not be reported on the visual result. We can see it as a selection of the overlapping area.

The value of the score is included between 0 and 1. The best value is 1 if the images are the same so we expect the higher possible score.

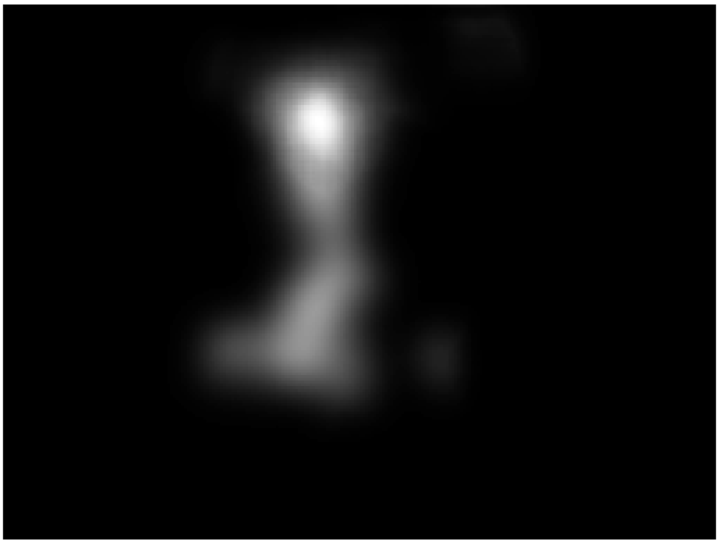

Figure 2: Evaluation (SIM) of the saliency map with comparison to ground truth; Score=0.7277

2) Pearson's Correlation Coefficient (CC)

This norm is used to evaluate how correlated/independent two variables are. I don't understand why this norm can be use on saliency map and how pixels are seen. Are they seen as a sample on a unique random variable or each pixel is a different random variable? In the first case, there is no position information and the order of pixels as no importance. In the second case, the sampling's size on each value is 1 and we have not any statistical view. How can we find a correlation?

$$CC(P,Q) = \frac{\sigma(P,Q)}{\sigma(P) * \sigma(Q)}$$

The problem here is the way we compute σ the covariance and its meaning on the saliency maps. The perfect score is 1 and the value are included in [0,1].

For the visualization of this norm, we use this formula on each pixel:

$$V_i = \frac{P_i * Q_i}{\sqrt{\sum_j (P_j{}^2, Q_j{}^2)}}$$

Because it is a symmetric computation between P and Q, we cannot distinguish a false positive or a false negative.

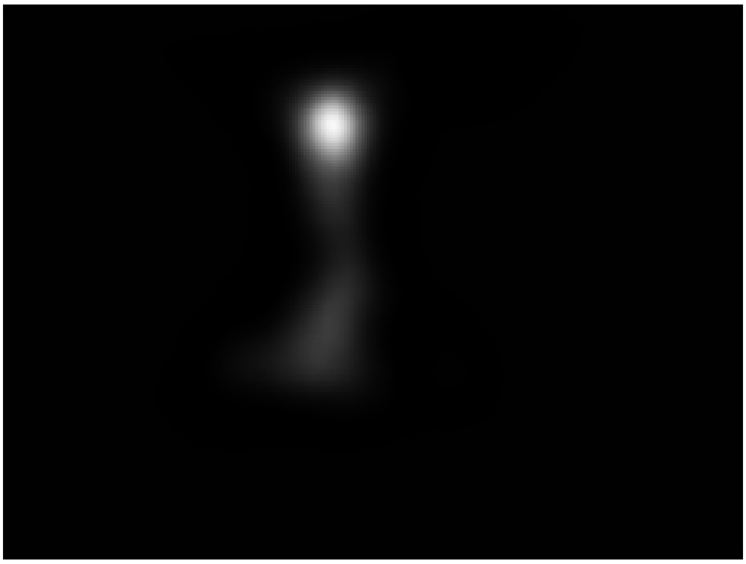

Figure 3: Evaluation (CC) of the saliency map with comparison to ground truth; Score=0.8606

3) Kullback-Leibler divergence (KL)

The KL norm measure the difference between two probability distribution. There are many ways to compute it but in the one we use is:

$$\mathrm{KL}(\mathrm{P},\mathrm{Q}) = \sum\nolimits_i Q_i * \log(\varepsilon + \frac{Q_i}{\varepsilon + P_i})$$

where ε is a very small number. The way it is computed make it receptive to false negative. The only value that matter is when the ground truth is non-zero. Then, because of the log, close values between Q and P does not appear in the norm. Only the false negative impact the norm.

Visually, we compute pixel per pixel this norm so only false negative appeared.

The values are positive and the best result is 0.

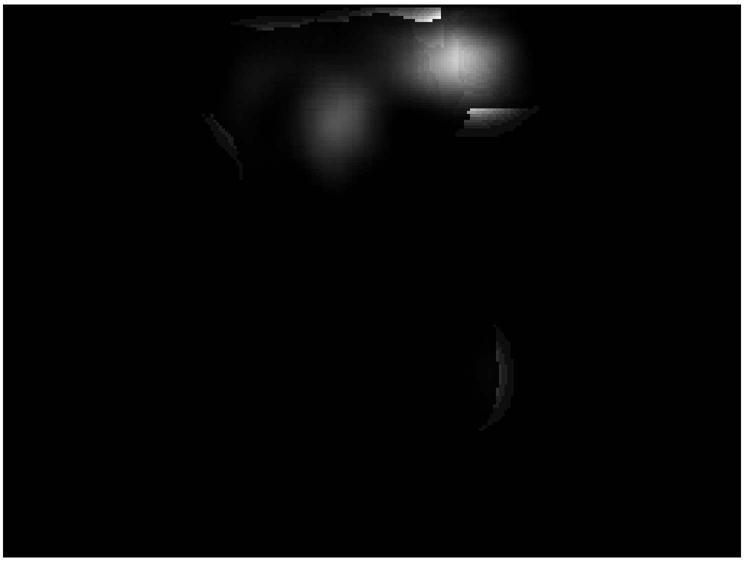

Figure 4:Evaluation (KL) of the saliency map with comparison to ground truth; Score=0.4825

4) Information Gain (IG)

This norm computes the gain between two maps knowing the ground truth. It can be use with only one saliency map and a center prior baseline to check any biases in the prediction.

In our results, it is useful to evaluate results in comparison between them.

The formula is:

$$\mathrm{IG}(\mathrm{P1}, \mathrm{P2}, \mathrm{Q}) = \frac{1}{N}\sum_{i} Q_i * (\log(\varepsilon + P1_i) - \log(\varepsilon + P2_i))$$

where ε is a very small number and N the total number of pixel.

Like for the KL metric, only the values that corresponds to a non-zero ground truth value matter. It means that false negative (or less false negative) are much more significant that true (or better) positive.

A positive score indicates a better prediction of the map P1 and a negative one the opposite.

For the visualization, we compute separately both logarithm. Then we can plot improvement in blue (P1 better than P2) and worst prediction in red (P2 better than P1).

Figure 6: Ground truth, first computed saliency map and second one (left to right)

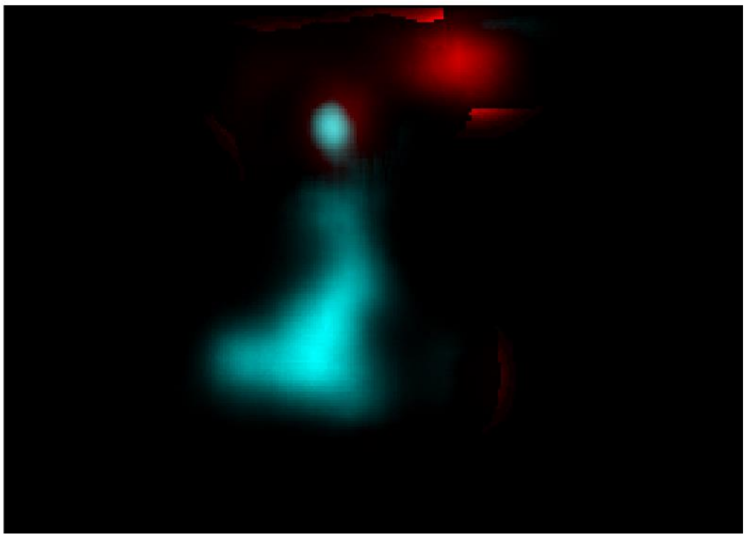

Figure 5: Information Gained from the first saliency map to the second according to ground truth; Score = -0.10307