



HAL
open science

Saliency aggregation: Does unity make strength?

Olivier Le Meur, Zhi Liu

► **To cite this version:**

Olivier Le Meur, Zhi Liu. Saliency aggregation: Does unity make strength?. 12th Asian Conference on Computer Vision (ACCV 2014), Nov 2014, Singapor, Singapore. hal-01085898

HAL Id: hal-01085898

<https://inria.hal.science/hal-01085898>

Submitted on 21 Nov 2014

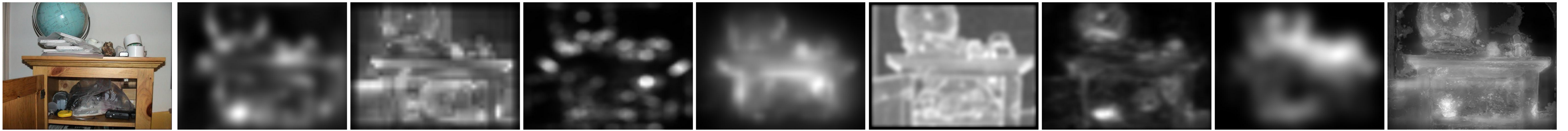
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Saliency aggregation: Does unity make strength?

O. Le Meur (olemeur@irisa.fr) & Z. Liu

Motivation: we investigate whether we could improve on the saliency quality by aggregating a set of predicted saliency maps.



Does the linear combination improve the prediction?

$$p(x|M_1, \dots, M_K) = \sum_{k=1}^K w_k \times p(x|M_k)$$

$p(x|M_k)$ is the probability of an image pixel x from the saliency map M_k to be salient,
 $\sum_{k=1}^K w_k = 1$

Unsupervised methods to compute the weights

- **Uniform:** weights w are uniform and spatially invariant, $w_k = 1/K$
- **Median**
- **M-estimators:**

$$g_{Welsh}(e(x|M_k)) = \exp\left(-\frac{e(x|M_k)^2}{\sigma^2}\right)$$

$$g_{L1L2}(e(x|M_k)) = \frac{1}{\sqrt{1 + e(x|M_k)^2}}$$

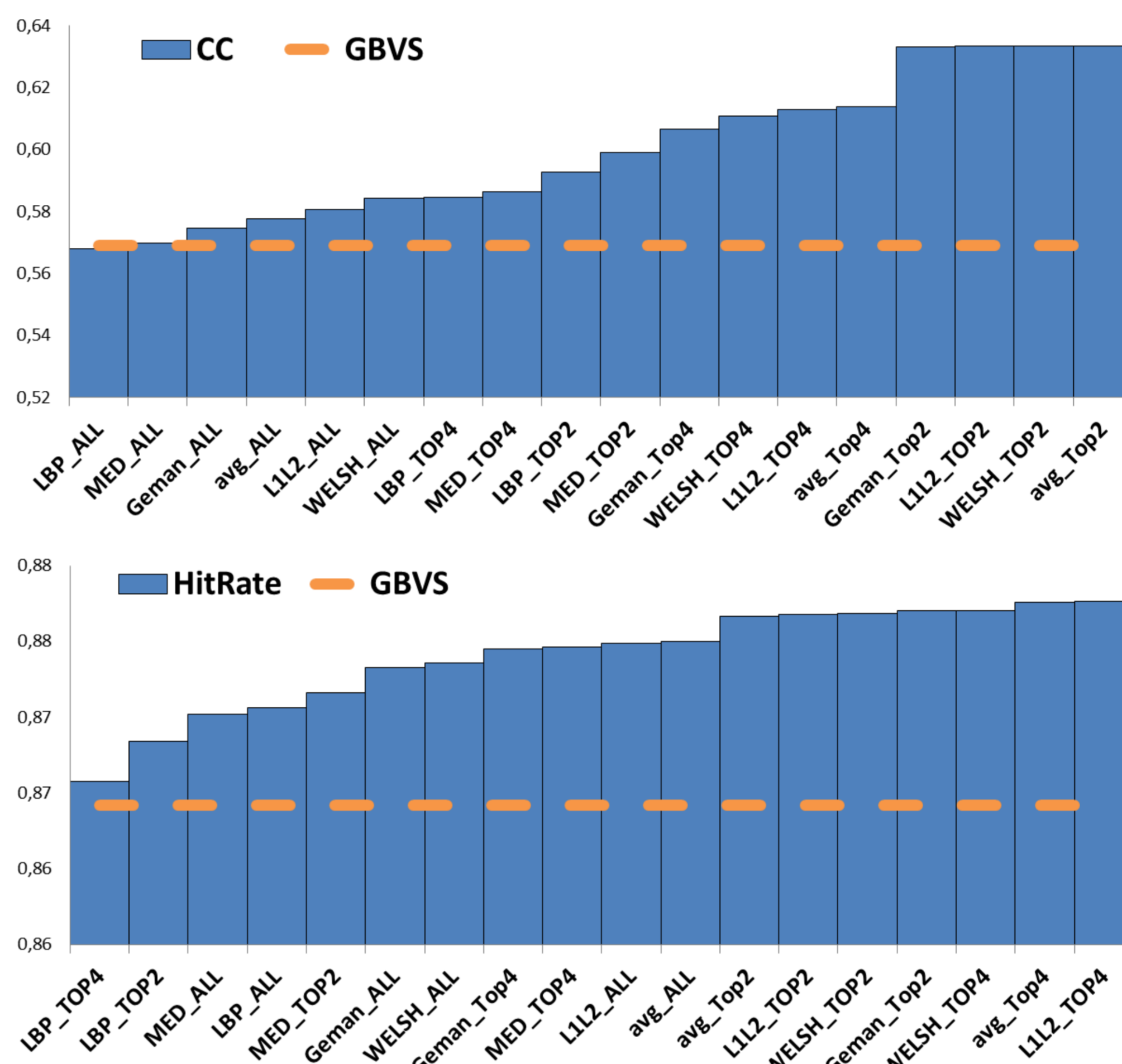
$$g_{Geman}(e(x|M_k)) = \frac{1}{1 + e(x|M_k)^2}$$

- **Global minimization of an energy function (Loopy Belief Propagation)**

TOP 2 : GBVS and RARE2012

TOP 4 : GBVS, RARE2012, Judd and AWS

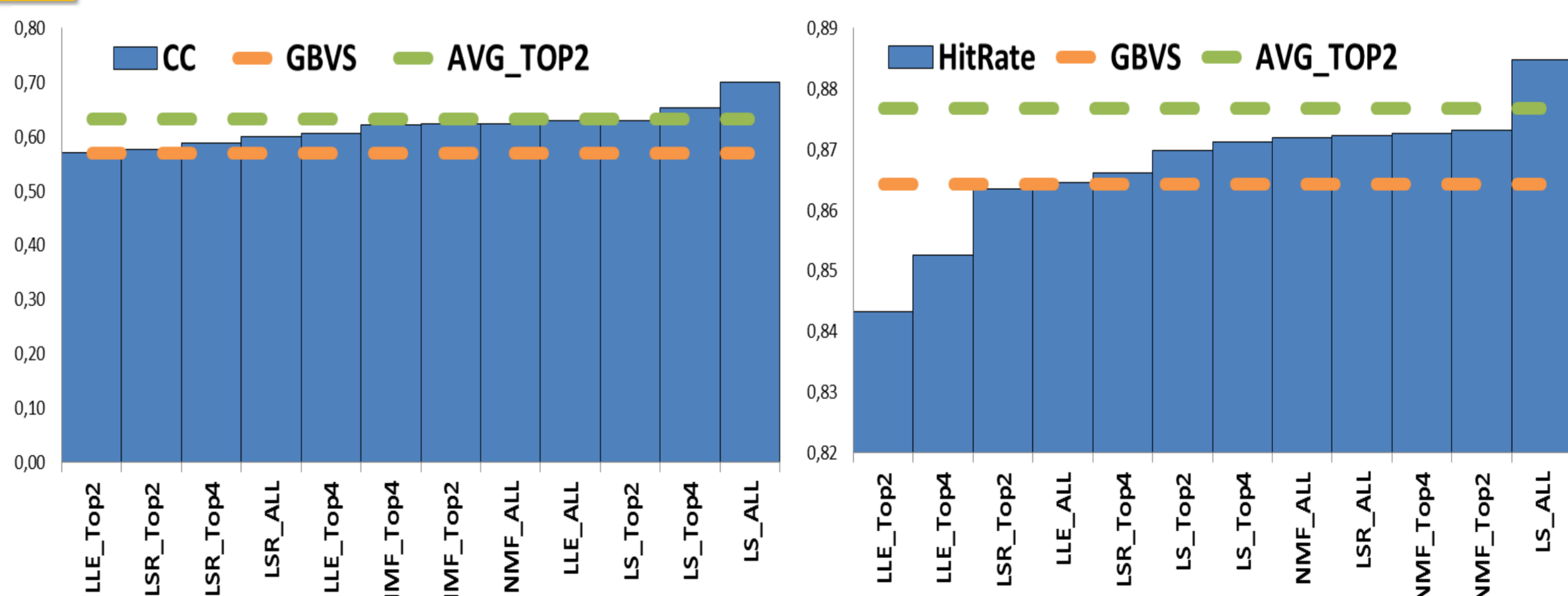
ALL: GBVS, RARE2012, Judd, AWS, Le Meur, Hou, Itti, Bruce.



A simple aggregation function, such as the average function, operating on the top 4 or top 2 models is a good candidate to improve significantly the performance of saliency models.

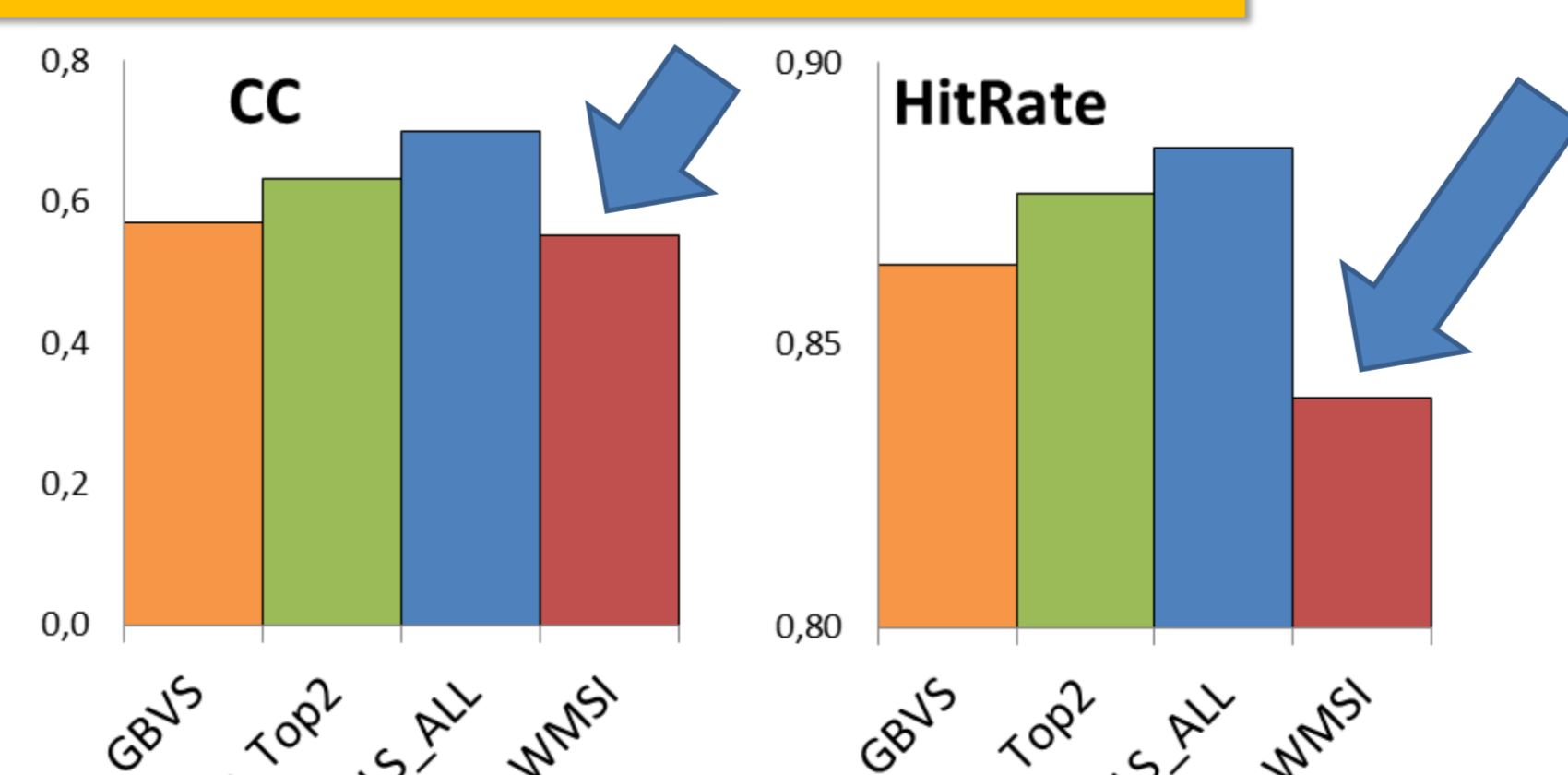
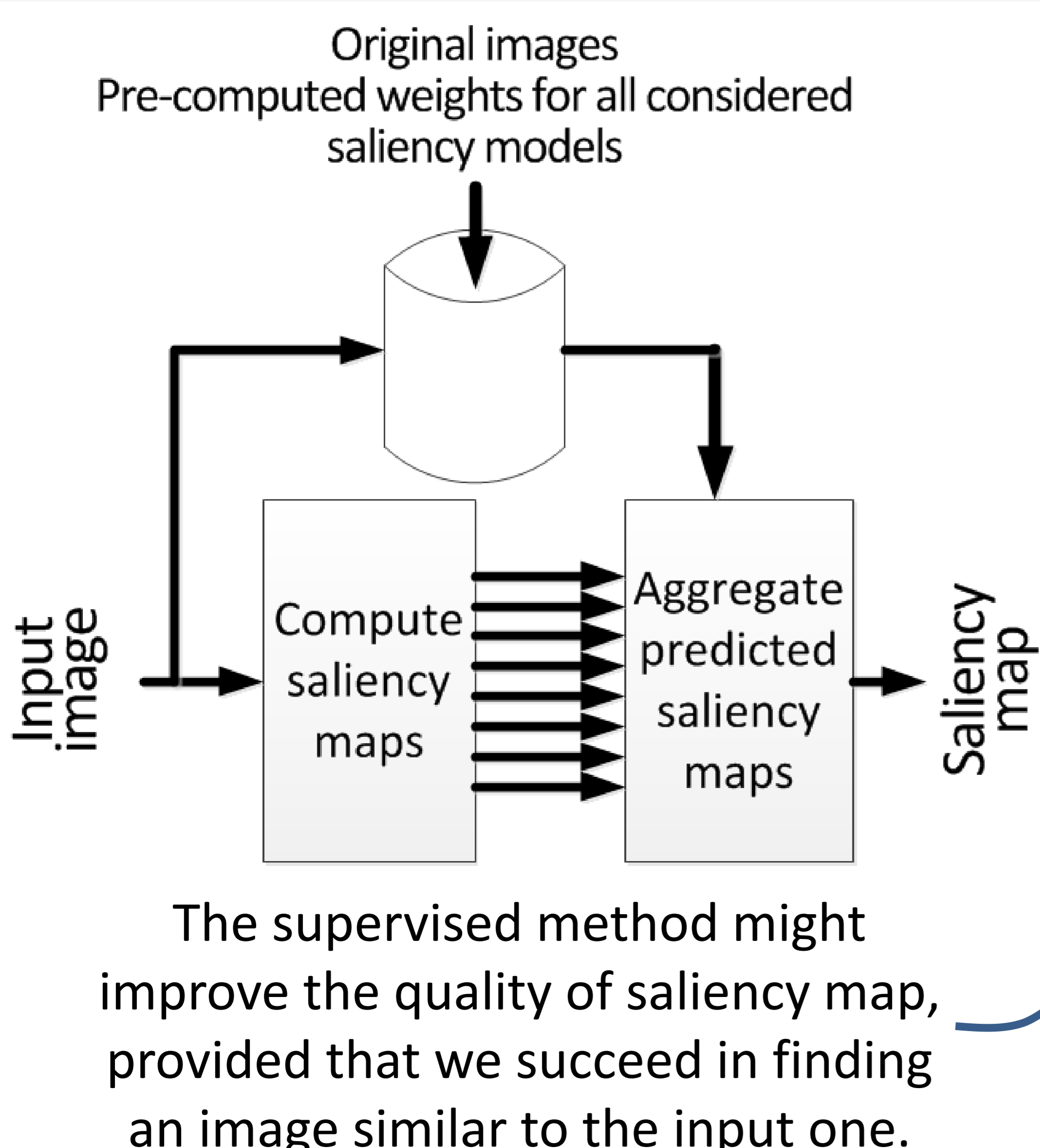
Least-squares estimation of the weights

- The classical least-squares method (LS)
- LS with the sum-to-one constraint of the weights (LLE)
- LS with positive weights (NMF)
- Robust least-squares (LSR)



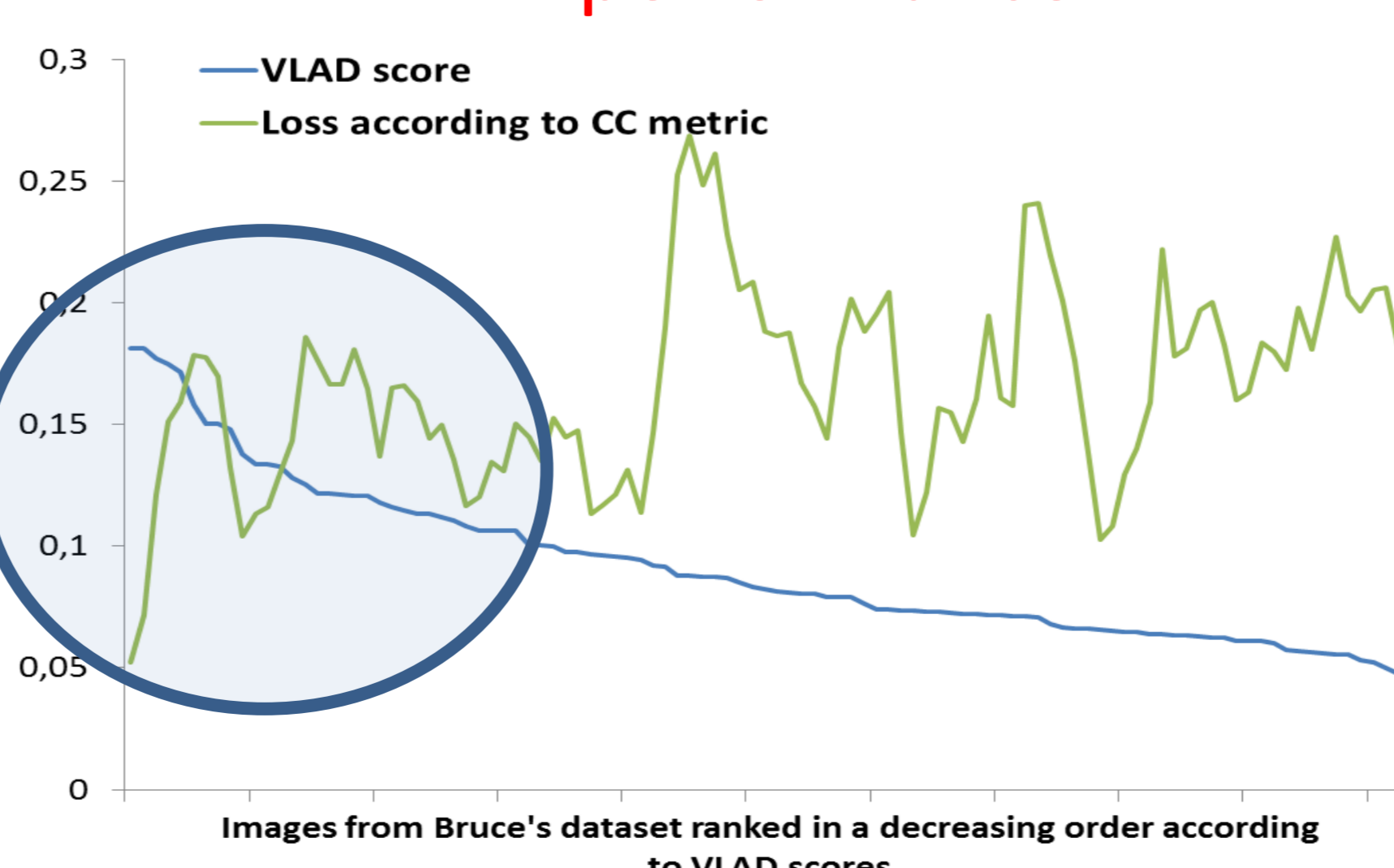
The best learning function is the simple least-squares method involving all saliency maps.

The use of the learned weights of the most similar image



WMSI = Weights of the Most similar Image

The WMSI method gets the lowest performance



Conclusion

The robustness of the prediction can be enhanced by:

- ✓ averaging the saliency maps of the top 2 models;
- ✓ considering a **dedicated training dataset**.

Acknowledgment

This work is supported in part by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme under Grant No. 299202 and No. 911202, and in part by the National Natural Science Foundation of China under Grant No. 61171144.