Supplementary material of the CVPR'17
Viraliency: Pooling Local Virality

Xavier Alameda-Pineda[1,2], Andrea Pilzer[1], Dan Xu[1], Nicu Sebe[1], Elisa Ricci[3,4]
[1] University of Trento, [2] Perception Team, INRIA [3] University of Perugia, [4] Fondazione Bruno Kessler
xavier.alameda-pineda@inria.fr, {andrea.pilzer,dan.xu,niculae.sebe}@unitn.it, eliricci@fbk.eu

1. Training details

We implemented our LENA pooling layer within the Caffe framework and ran all our experiments using a Tesla K40 GPU. All the networks were fine-tuned from the convolutional filters obtained when training these networks for the $1{,}000$ image classification task on the ImageNet dataset. We iterated the stochastic gradient descent algorithm for $10{,}000$ iterations with a momentum of $\mu = 0.9$ and a weight decay of $\lambda = 0.05$. The learning rate followed a `step` policy with factor 0.1 every $5{,}000$ iterations, with base learning rate set at 0.0001.

2. Virality Score

While the virality scores for the UIV dataset were provided with the images, this information was not available for the IVGP dataset. In order to have similar measures of virality and therefore be able to compare results obtained on the two datasets, we adapted the annotation procedure of [1] to the metadata provided in [2]. The original virality score was defined in [1] as:

$$V_i = \frac{L_i}{\bar{L}} \log\left(\frac{M_i}{\bar{M}}\right), \tag{1}$$

where V_i is the virality score of the i-th image, L_i and $\bar{L}$ are the number of likes associated to the i-th image and the average number of likes over the dataset, respectively, and M_i and $\bar{M}$ are the number of resubmissions of image i and the average number of resubmissions over the dataset, respectively. The terms "likes" and "resubmissions" are valid for images extracted from Reddit, but need to be adapted for images downloaded from GooglePlus. We intuitively choose to replace the "resubmissions" by "reshares" and "likes" by the difference of "upvotes" minus "downvotes", and use the exact same formulation (previous equation) to retrieve the virality score of each image in IVGP.

3. Extra viraliency maps

In the next pages we present more examples of the viraliency maps obtained for different images with the three pooling strategies (GAP, GMP and GLENAP) with and without objectness. We show results on the IVGP dataset from Figure 1 to Figure 5 and on the UIV dataset from Figure 6 to Figure 9.

Generally speaking we can observe the following trends. First, GMP is generating spread viraliency maps, where many image locations are partly highlighted. Second, we observe that GLENAP is able to produce viraliency maps composed of several strong active areas. Compared to GAP, this has the advantage of pointing to many virally salient locations in the images, since GAP is mostly producing one strong compact area. When objectness is added things may change in two directions depending on the image (and this holds for the three pooling strategies that tend to keep their relative differences). If the image contains clearly defined objects (*i.e.* in the foreground), the viraliency maps tend to focus more on these objects. Otherwise, the viraliency maps are modified in a way that may be perceived as "independent of the image content." This behavior is quite expected since objectness cannot provide strong localization cues if the objects in the image are not strongly localized.

We believe this large set of examples (chosen to show both the successes and failures of the proposed GLENAP layer) provides rich insights on the capabilities of the novel layer for virality localization and on how does it function.

References

[1] A. Deza and D. Parikh. Understanding image virality. In *IEEE CVPR*, 2015. 1

[2] M. Guerini, J. Staiano, and D. Albanese. Exploring image virality in GooglePlus. In *Int. Conf. on Social Comp.*, 2013. 1

[3] M. Oquab, L. Bottou, I. Laptev, and J. Sivic. Is object localization for free?-weakly-supervised learning with convolutional neural networks. In *IEEE CVPR*, 2015. 2, 3, 4, 5, 6, 7, 8, 9, 10

[4] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba. Learning deep features for discriminative localization. In *IEEE CVPR*, 2016. 2, 3, 4, 5, 6, 7, 8, 9, 10

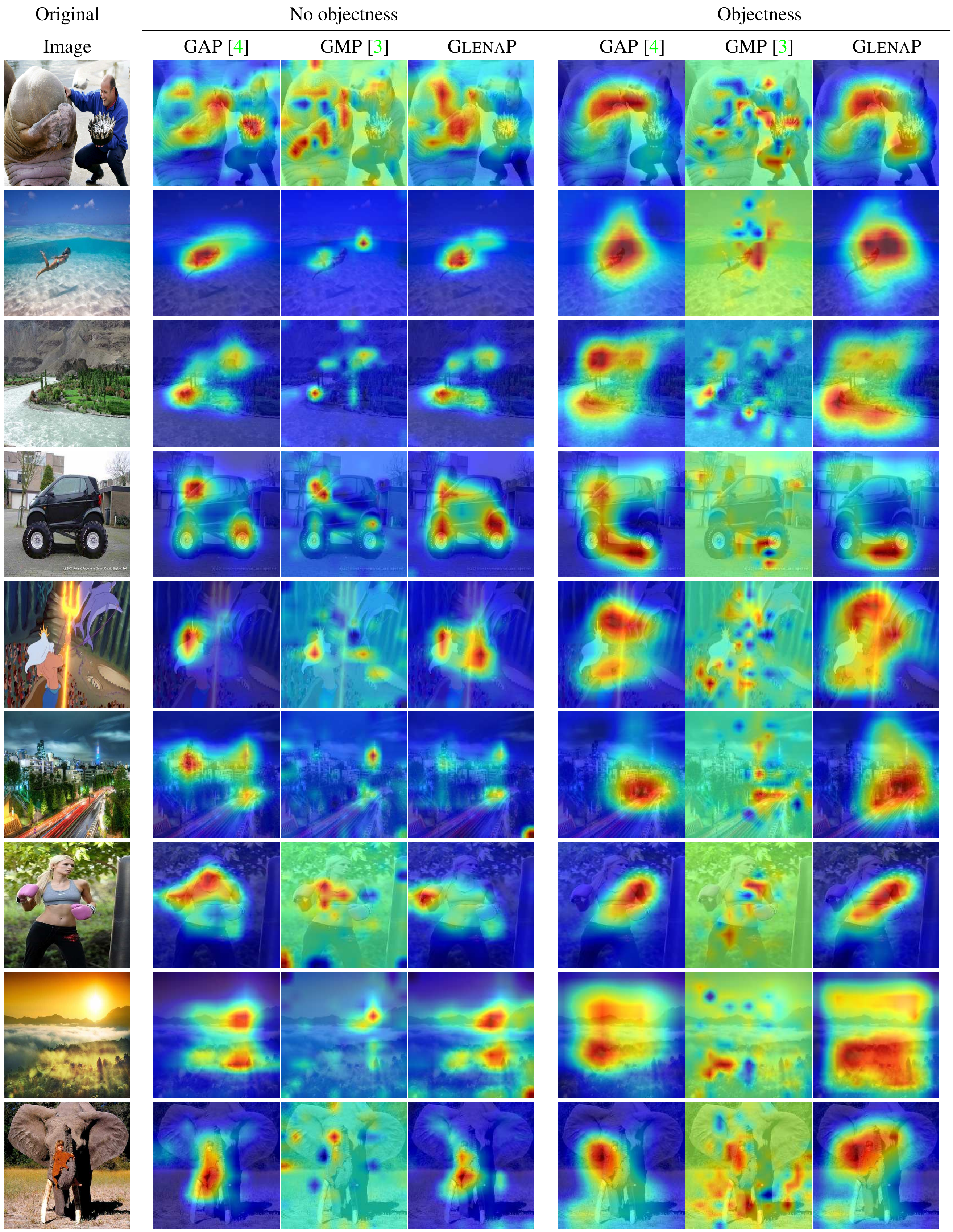

Figure 1. Sample viraliency maps for the IVGP dataset.

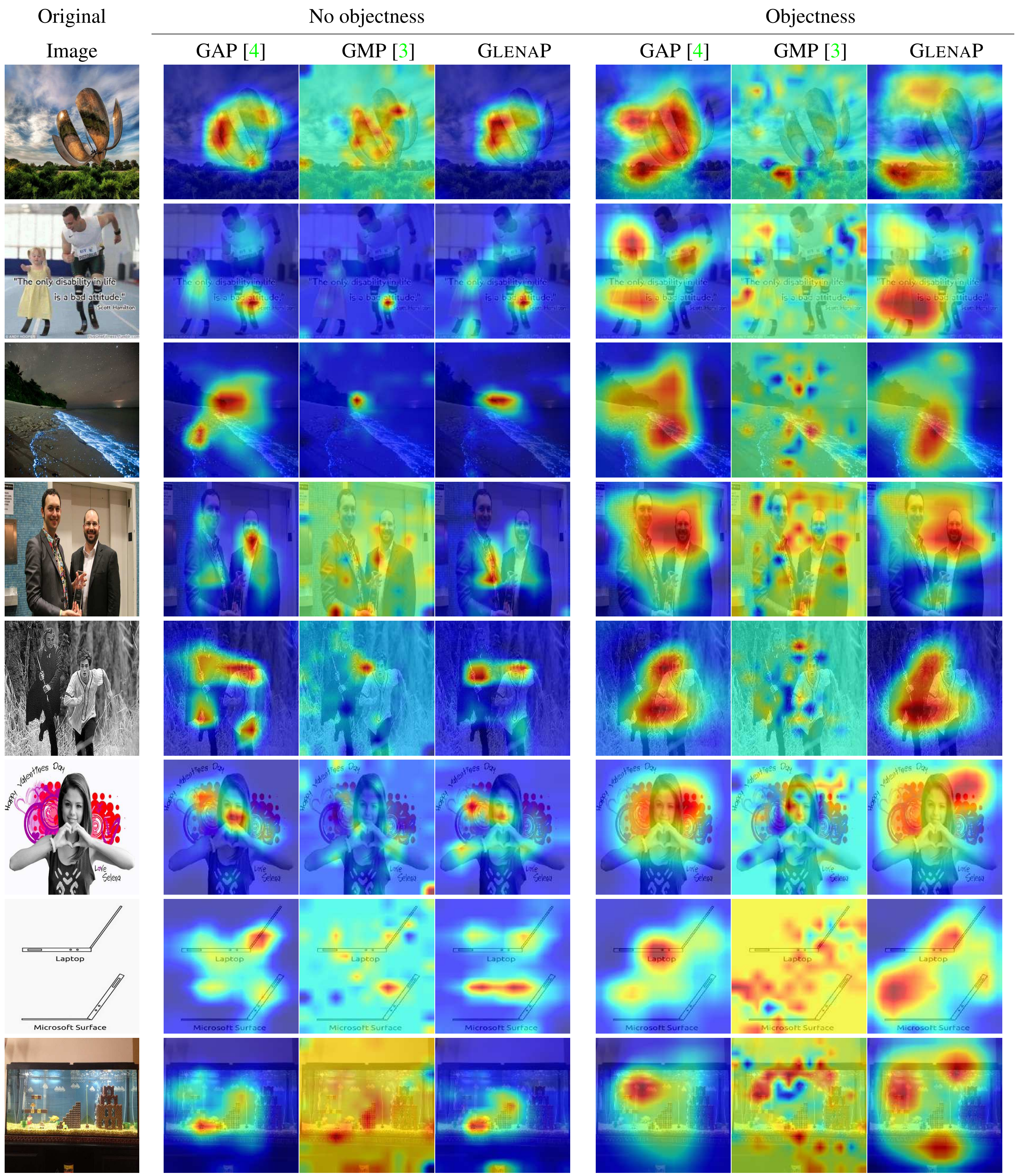

Figure 2. Sample viraliency maps for the IVGP dataset.

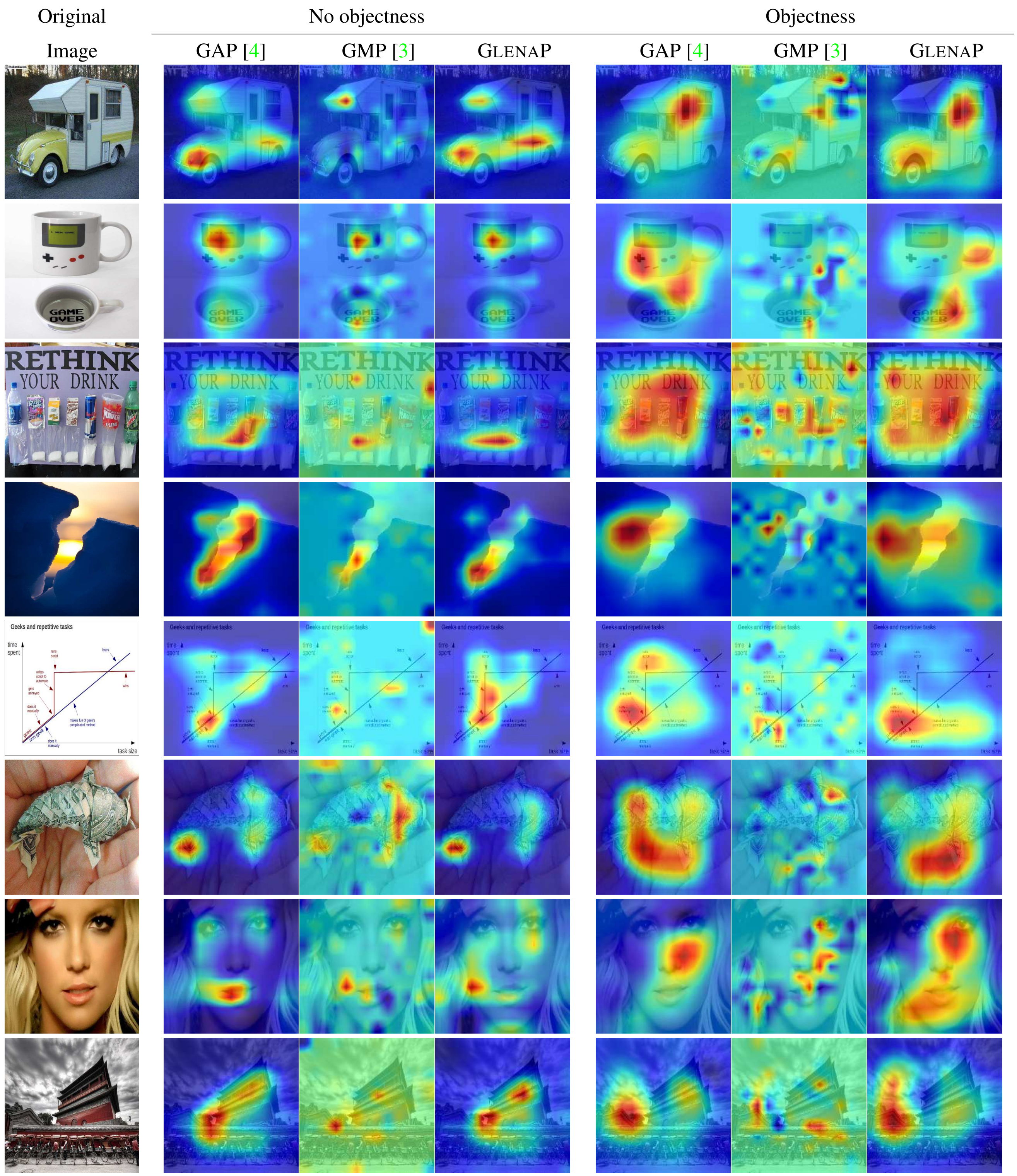

Figure 3. Sample viraliency maps for the IVGP dataset.

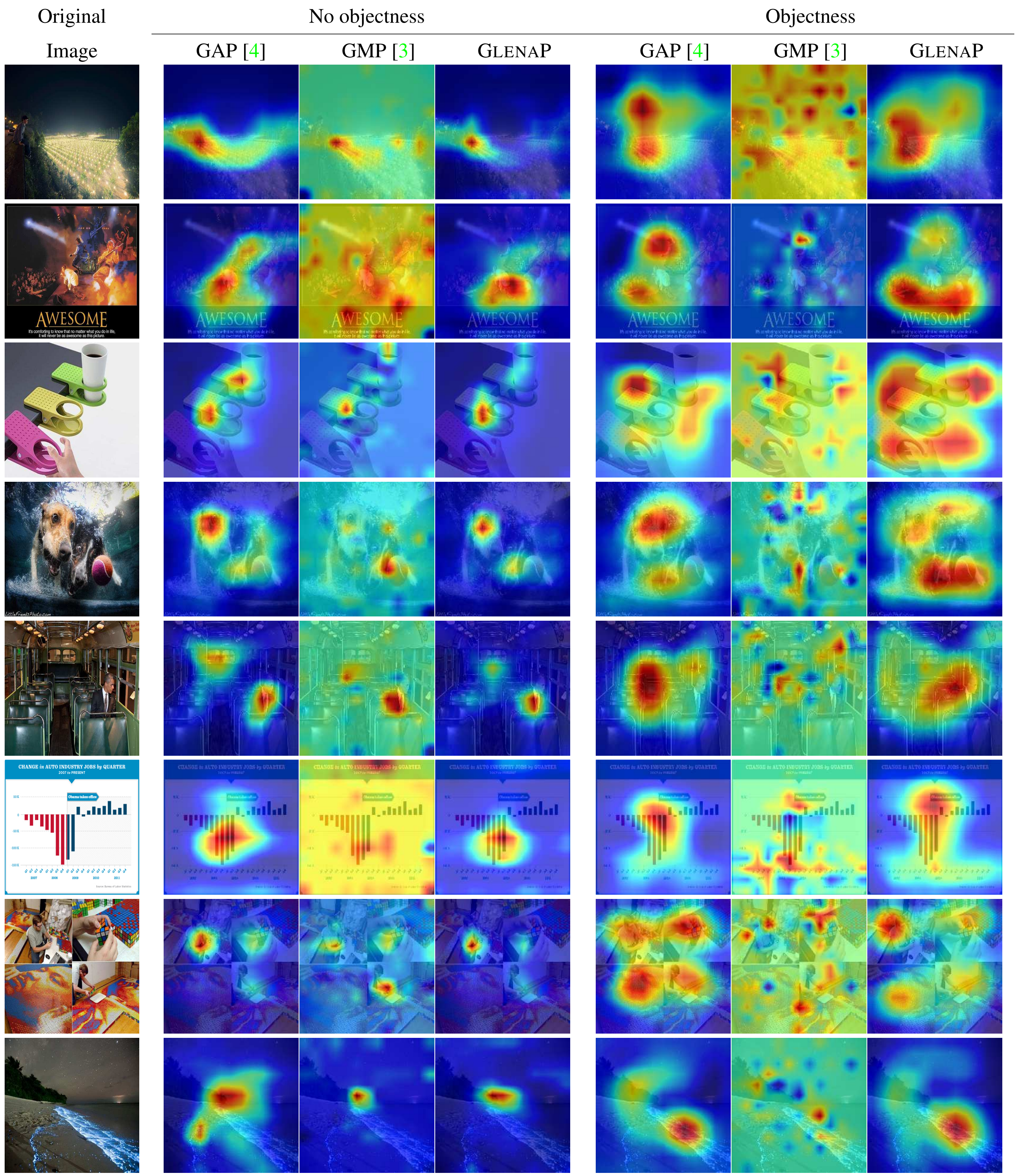

Figure 4. Sample viraliency maps for the IVGP dataset.

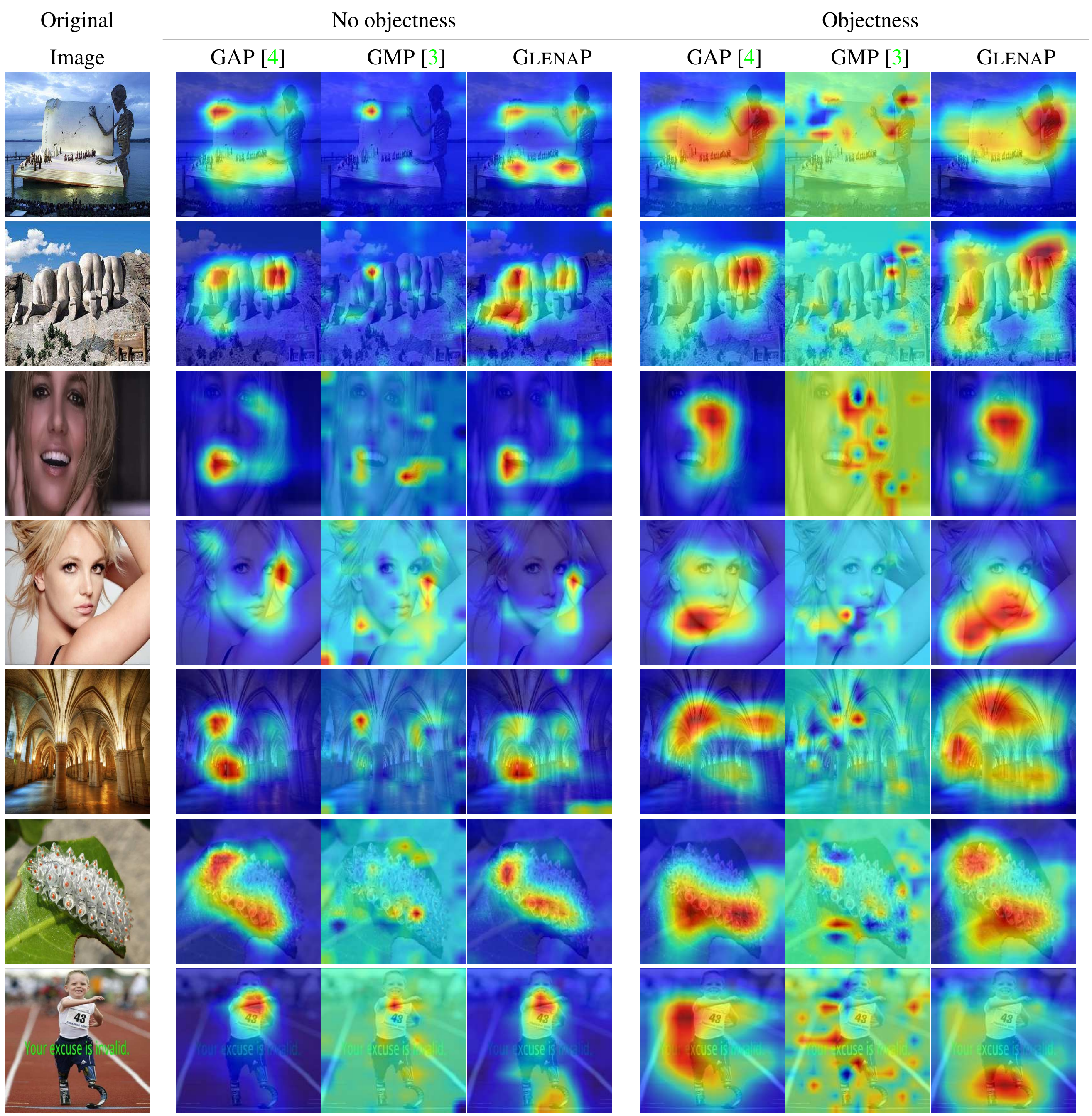

Figure 5. Sample viraliency maps for the IVGP dataset.

Original Image	No objectness			Objectness		
	GAP [4]	GMP [3]	GLENAP	GAP [4]	GMP [3]	GLENAP

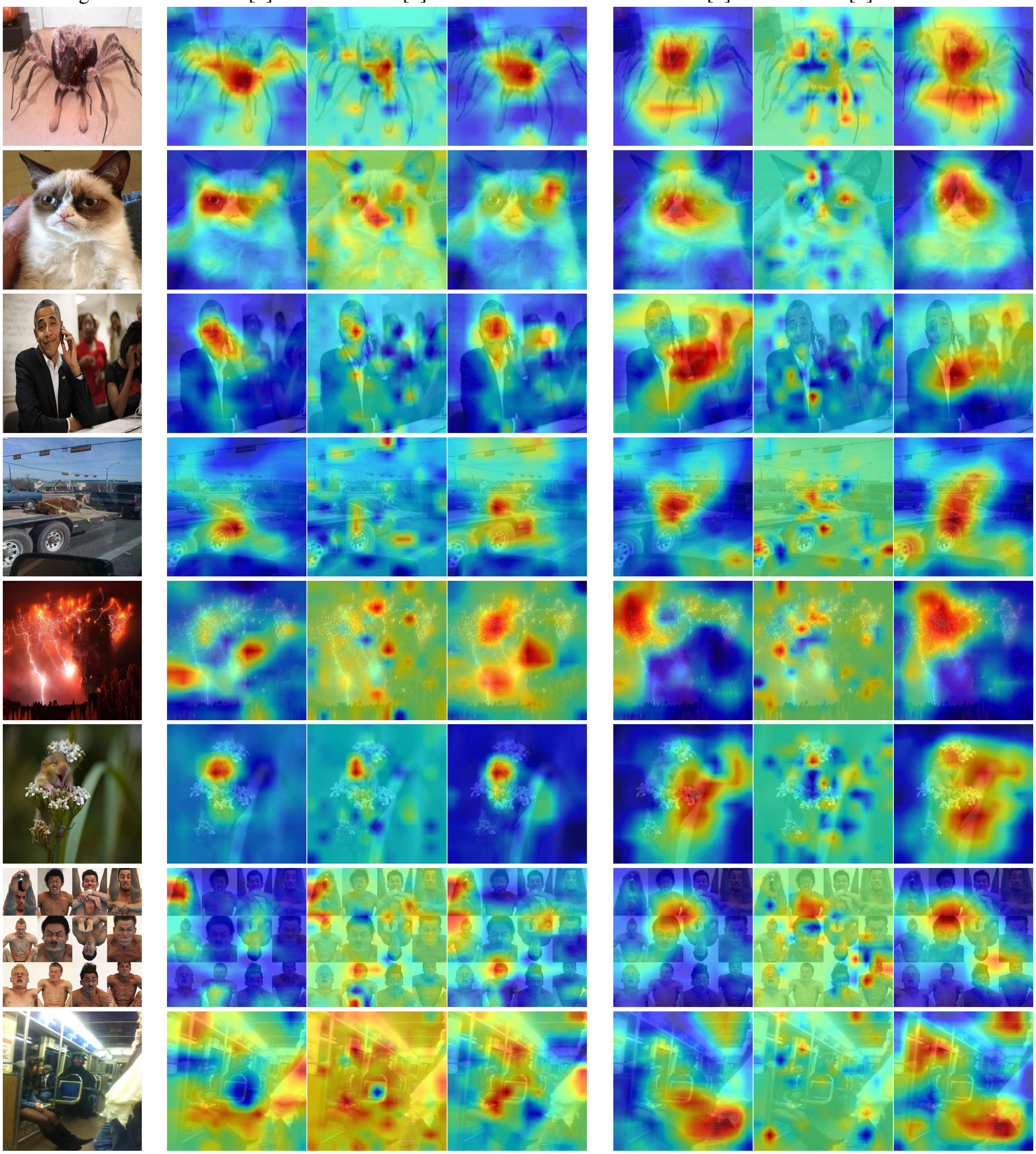

Figure 6. Sample viraliency maps for the UIV dataset.

Original Image	No objectness			Objectness		
	GAP [4]	GMP [3]	GLENAP	GAP [4]	GMP [3]	GLENAP

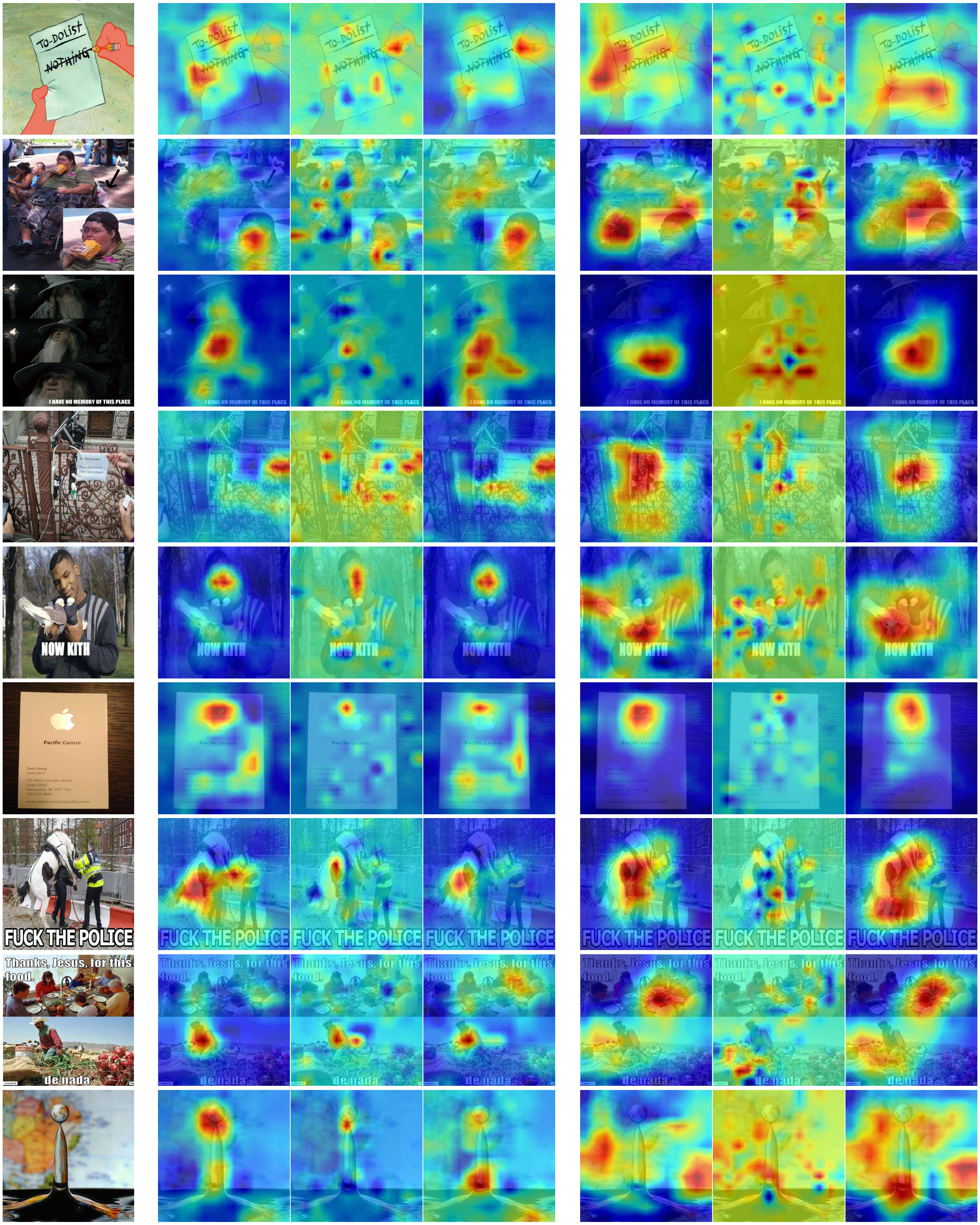

Figure 7. Sample viraliency maps for the UIV dataset.

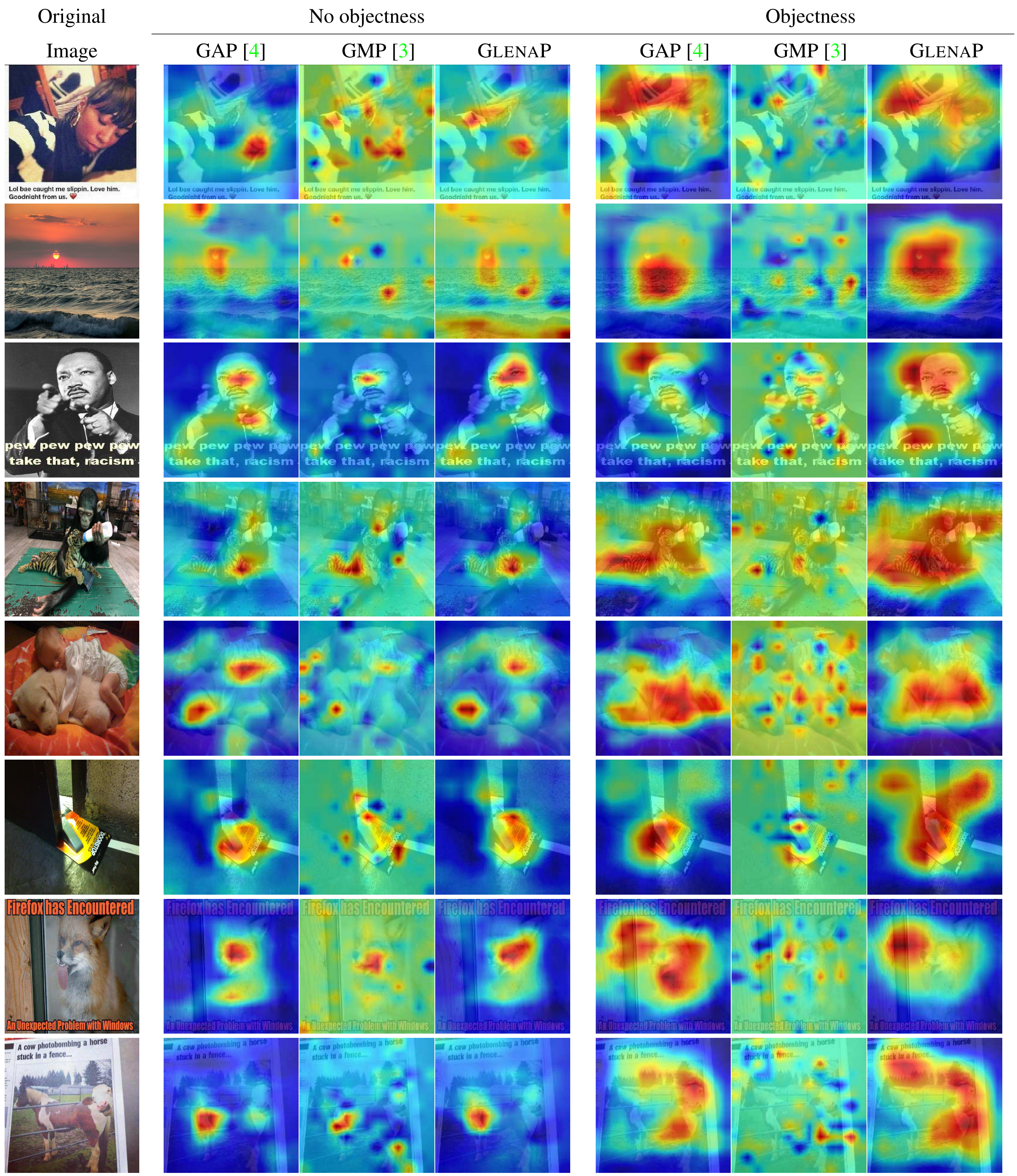

Figure 8. Sample viraliency maps for the UIV dataset.

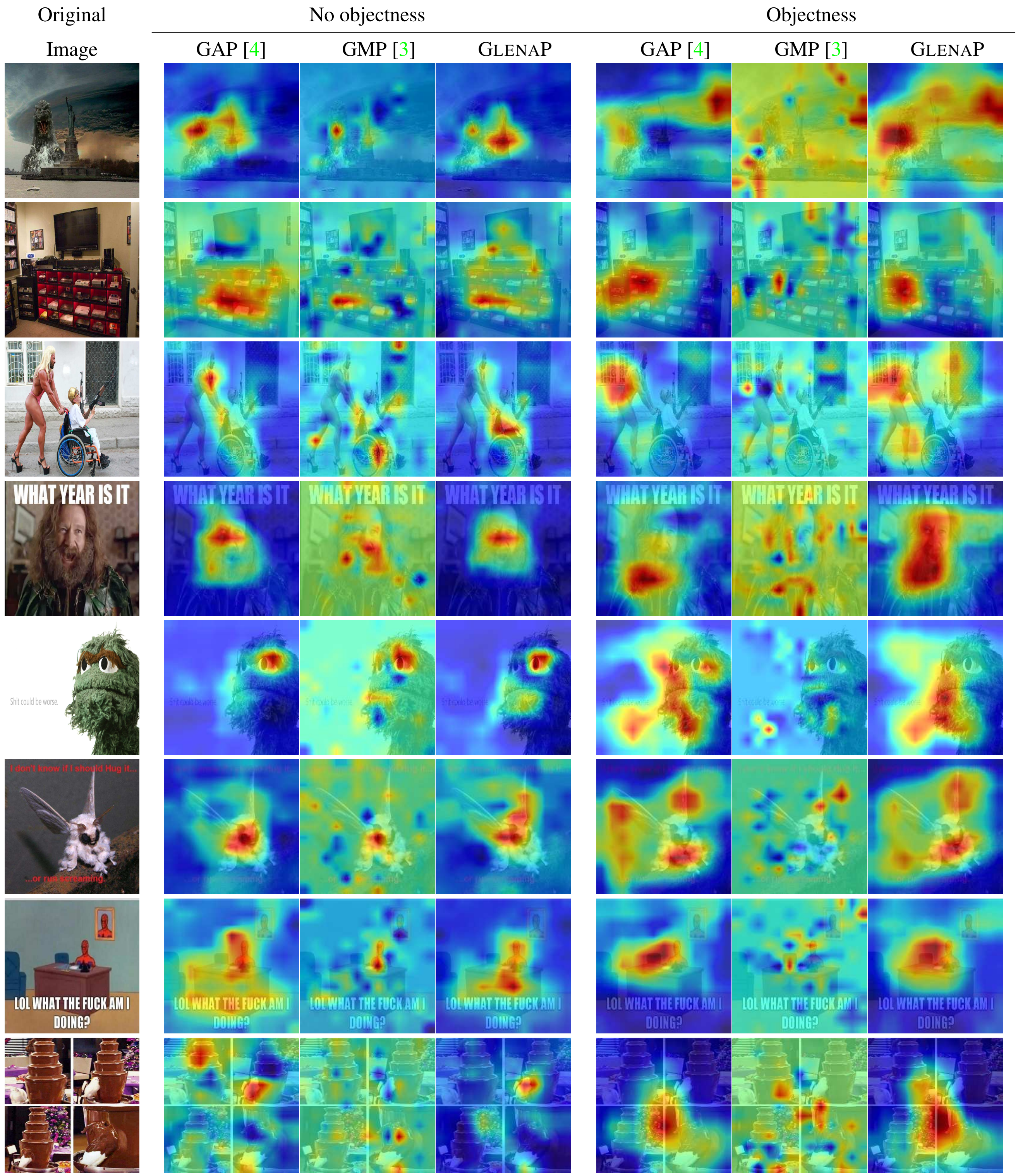

Figure 9. Sample viraliency maps for the UIV dataset.