SMILER: Consistent and Usable Saliency Model Implementations

Toni Kunić, Calden Wloka, John K. Tsotsos

Computational modeling of saliency is a field with decades of development, and during that time an enormous array of different models have been proposed and implemented [1]. The diversity of model design choices and philosophies can be quite challenging to navigate. Insofar as the field has made efforts at standardization, these have largely focused on performance benchmarking [2]. However, the practical challenges of sharing code between different research groups have largely been ignored, which causes potential issues of consistency in the literature, as well as a high barrier to entry for new research.

To help address this situation, we present the Saliency Model Implementation Library for Experimental Research (SMILER), a curated bundle of saliency models with a standardized API for model execution and parameter handling in an easy-to-install package with both a MATLAB and a command line interface. SMILER currently supports twenty-three saliency models (see Figure 1), a number which will grow in the future through community contributions. Establishing a common standard facilitates interactions among different groups of saliency researchers, and helps move the field forward by reducing the effort involved in exploring new research directions.

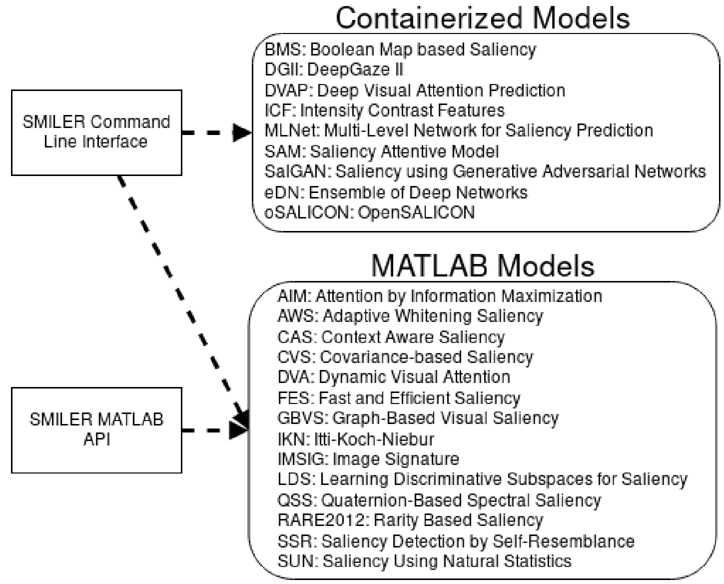

Figure 1: Diagram showing model implementations and interfaces provided by the SMILER system.

The standardized handling of model execution and parameter setting is important not only for reducing the burden of use, but also for ensuring consistent and correct results. In working with any implementation of a saliency model, it is not always immediately clear which parameters have been exposed by the authors, and the field largely lacks consistency with regard to default settings. Some models may be implemented to automatically execute using a predefined set of recommended defaults, whereas others may expect the user to familiarize themselves with a set of recommended settings which differ from the normal default parameters and manually set them. Lastly, there may be parameter settings which are never reported in the literature, leaving subsequent researchers to guess at the appropriate value. Without a common standard, the field is prone to comparisons which are not truly equivalent, which can result in inconsistent performance evaluations [3].

SMILER can help address these issues, and by doing so promote previously laborious avenues of research, for example a stronger interaction between psychophysical studies of primate attention and computational saliency modeling. A number of recent studies which have explored the correlation of saliency with human behaviour [4] or monkey neurological recordings [5] have done so with comparison to only a limited representation of saliency models. By enabling reliable and simplified access to a wide array of models, SMILER can increase the robustness of these types of study, while also identifying potential holes in model behaviour and performance through a more principled analysis of the behaviour which the models are seeking to encompass [6].

The source code for SMILER is available at https://github.com/tsotsoslab/smiler. The project is open source and encourages community contributions.

References

[1] Z. Bylinskii, E. M. DeGennaro, R. Rajalingham, H. Ruda, J. Zhang, and J. K. Tsotsos, "Towards the quantitative evaluation of visual attention models," *Vision research*, vol. 116, pp. 258–268, 2015.

[2] Z. Bylinskii, T. Judd, A. Borji, L. Itti, F. Durand, A. Oliva, and A. Torralba, "Mit saliency benchmark," 2015.

[3] C. Wloka, T. Kunić, I. Kotseruba, R. Fahimi, N. Frosst, N. D. Bruce, and J. K. Tsotsos, "Smiler: Saliency model implementation library for experimental research," *arXiv preprint arXiv:1812.08848*, 2018.

[4] A. Nuthmann and J. M. Henderson, "Object-based attentional selection in scene viewing," *Journal of vision*, vol. 10, no. 8, pp. 20–20, 2010.

[5] B. J. White, J. Y. Kan, R. Levy, L. Itti, and D. P. Munoz, "Superior colliculus encodes visual saliency before the primary visual cortex," *Proceedings of the National Academy of Sciences*, vol. 114, no. 35, pp. 9451–9456, 2017.

[6] N. D. Bruce, C. Wloka, N. Frosst, S. Rahman, and J. K. Tsotsos, "On computational modeling of visual saliency: Examining what's right, and what's left," *Vision research*, vol. 116, pp. 95–112, 2015.