ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

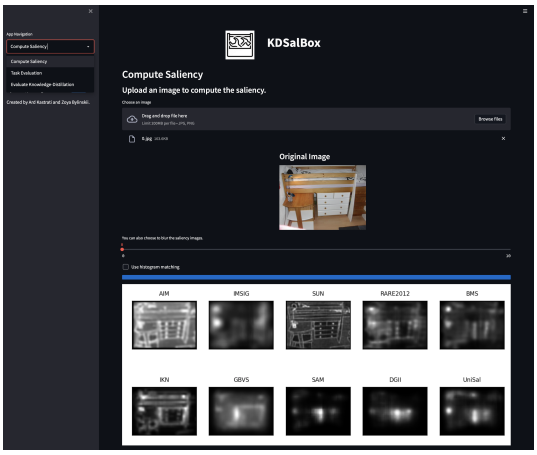
Distributed Computing

Prof. R. Wattenhofer

On the Generalization of Saliency Models

Dozens of saliency models have been designed over the last few decades, targeted at diverse applications ranging from image compression and retargeting to robot navigation, surveillance, and distractor detection. Barriers to their use include the different and often incompatible software environments that they rely on, as well as the computational inefficiency of older implementations. For application-purposes models are then frequently chosen based on convenience and efficiency, at the expense of optimizing for task performance.

To facilitate the evaluation and selection of saliency models for different applications, we created KDSalBox - a toolbox of 10 knowledge-distilled saliency models. Using the original model implementations available in their native environments, we produce saliency training data for efficient MobileNet-based architectures, that are identical in their architecture but differ in how they learn to distribute saliency over an image.



In this thesis we want to improve the existing framework such that the models generalize better for new applications. For example, depending on the application and setting (e.g. on-device deployment) a further reduction in the size or inference time might be needed in practice. Depending on the model, one can also design separate decoders that are more suitable for each model. Another goal that we are interested in this thesis is that our models are trained only on the natural images from the SALICON dataset and as our preliminary generalization experiments show, some models (particularly students of deep neural net based teacher models) struggle more with generalizing to new image types than others. Extensions of this work call for training on other domains and with larger datasets. Finally, since all models share the same architecture and reside in the same framework, a natural extension of this work would be to combine these models like building blocks to increase performance on saliency benchmarks and on particular tasks where different models can offer complementary features (combined expertise).

Requirements: Prior experience or a strong interest in programming. Creativity skills are advantageous.

Interested? Please contact us for more details!

Contact

- Ard Kastrati: kard@ethz.ch, ETZ G61.3