Saliency and the population receptive field model to identify images from brain activity

Alex Hernandez-Garcia[1], Wietske Zuiderbaan[2], Akhil Edadan[2], Serge O. Dumoulin[2,3,4], Peter König[1,5]

[1]Institute of Cognitive Science, University of Osnabrück, Germany • [2]Spinoza Centre for Neuroimaging, Amsterdam, Netherlands • [3]Department of Experimental and Applied Psychology, VU University Amsterdam, Netherlands • [4]Department of Experimental Psychology, Helmholtz Institute, Utrecht University, Netherlands • [5]Institute of Neurophysiology und Pathophysiology, University Medical Center Hamburg-Eppendorf, Germany

ABSTRACT

One of the goals of visual neuroscience is to develop predictive models of brain activity to better understand the underlying mechanisms of the visual cortex. Here, we use the population receptive field (pRF) model to identify the presented stimulus from a set of natural images (Zuiderbaan et al., 2017) combined with saliency-related information from the images. Ultimately, we seek answers to these questions: What fraction of the fMRI responses is driven by saliency? Where in the visual cortex is saliency most represented? We calculate the prediction response profile of every image as the summed overlap of its saliency map with the pRF at each cortical location. Then, we compute the correlation between the fMRI recordings and the prediction profiles of all images to assess the predictive power of saliency compared to contrast at different cortical areas.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 641805

ahernandez@uos.de

alexhernandezgarcia.github.io

METHODS

1. Data acquisition

• 45 greyscale natural images from Berkeley Segmentation Dataset and Benchmark data set (Martin et al., 2001)
• As **ground truth**, we measured the **fMRI response** profiles of 2 subjects to each stimulus presented at 11x11° diameter of visual angle (Zuiderbaan et al., 2017)
• Cortical areas considered: V1, V2, V3, hV4, LO12 & V3AB
• We estimated the **pRF**-properties of each cortical location (Dumoulin & Wandell, 2008)

2. Response prediction

The task is to identify the presented stimulus from the pRF model and information from the images. Our goal is to compare three models or feature maps:

• **Contrast**: RMS contrast as in Zuiderbaan et al. (2017)
• **Deep Gaze II**: saliency map based high-level features learnt by a deep neural network (Kümmerer et al., 2017)
• **ICF**: saliency map optimized by a neural network, but restricted to low-level (intensity and contrast) features.

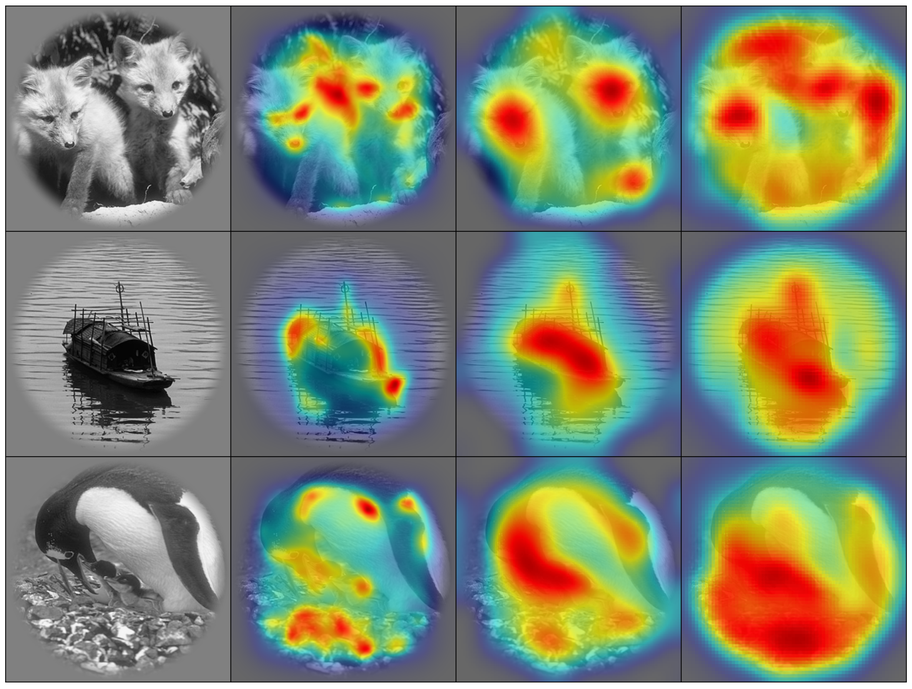
The predicted response is the normalized weighted sum of the feature map with the pRF of each cortical location:

$$p = \frac{\sum_{i=1}^{N} w_i \cdot S_i}{\sum_{j=1}^{M} w_j} \qquad w_i = \exp\left(-\frac{(x_i - x_c)^2 + (y_i - y_c)^2}{2\sigma^2}\right)$$

3. Evaluation

We compare the predictive power of each of the three feature maps, by assessing the **accuracy** of the identification and a **confidence** score:

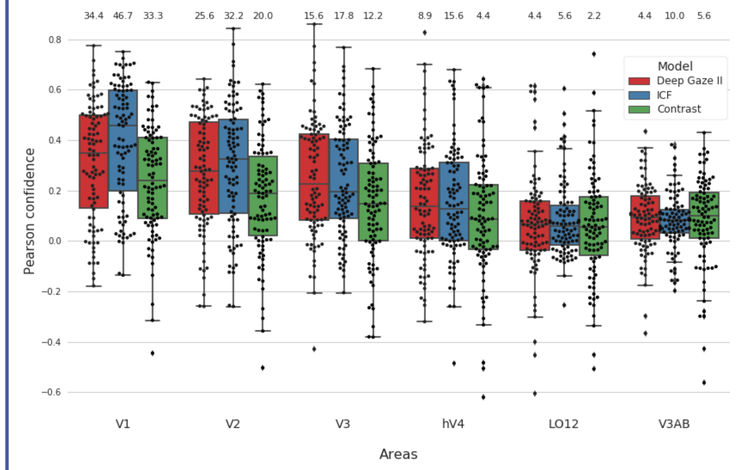
$$c_i = \frac{-\sum_{j=1}^{K} corr(p_j, m_i) - corr(p_i, m_i)}{K}$$



CONCLUSIONS

• The best identification performance is observed on V1 and the performance gradually decreases on higher areas - this is perhaps due to increased pRF sizes and decreased spatial accuracy on higher areas.
• Saliency maps provide higher predictive power than contrast in V1, V2 and V3 and no worse in higher areas - this suggests that saliency may contain richer information present in the measured responses.

RESULTS



• Each dot shows the prediction **confidence** on one image with respect to the measured response on one subject.
• The identification **accuracy** is shown on top of each box.
• The correlation of the predictions with the measured responses and the identification accuracy decrease on higher cortical regions.
• In all the cortical regions, saliency is able to identify the presented stimulus with higher accuracy than contrast.
• The intensity-contrast (ICF) model shows higher correlation than the deep features (Deep Gaze) model.

Martin et al., 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics.
Zuiderbaan et al., 2017. Image identification from brain activity using the population receptive field model.
Dumoulin & Wandell, 2008. Population receptive field estimates in human visual cortex.
Kümmerer et al., 2017. Understanding low- and high-level contributions to fixation prediction.