

Where are we looking?

Predicting human gaze using deep networks

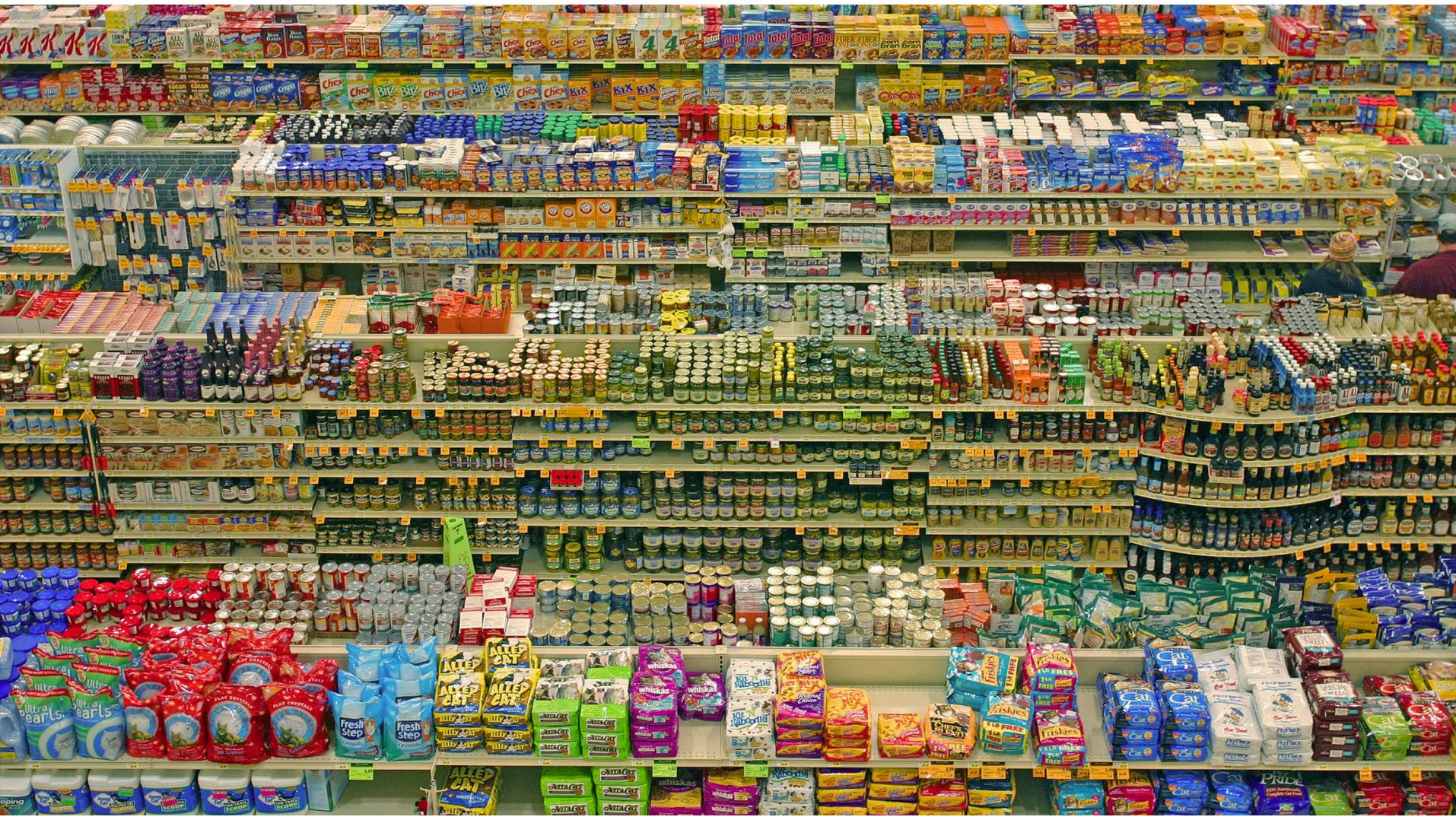
Oliver Eberle

Bernstein Center for Computational Neuroscience
Berlin

Prof. Bethge

Matthias Kümmerer
University of Tübingen

oliver.eberle@bccn-berlin.de
oeberle.blogspot.com
github.com/Kreiswolke
@EberleOliver





How to study human gaze?



How to study human gaze?

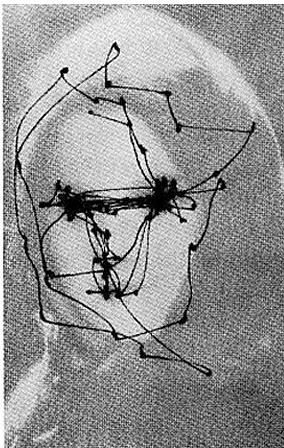
Eye-tracking devices record
where our eyes are looking at



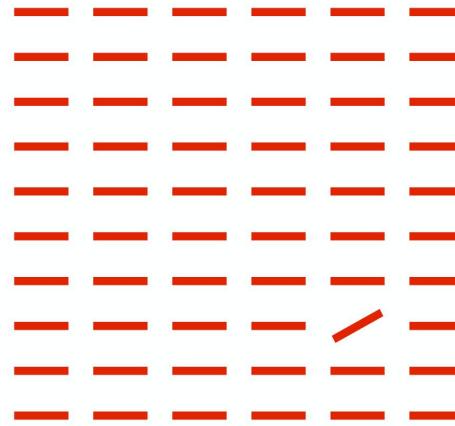
Eye-tracking in research



Yarbus, 1967



Treismann & Gelade,
1980



Judd et al., 2009

Eye-tracking in industry

The screenshot shows the Guardian Unlimited homepage from October 5, 2007. A heatmap overlay highlights areas of high visual attention, primarily focusing on the main news articles, the sidebar, and the search bar area.

Header: Sign in | Register | Text larger | smaller | Go to: Guardian Unlimited home | Go

Title: GuardianUnlimited

Date: October 5 2007 | Last updated less than one minute ago

Navigation: Home | Arts | Books | Business | Comment is free | Dating | Education | Environment | Film | Football | Jobs | Life & Style | Media | Society | Sport | Technology | Talk | Travel | UK news | World news

Content:

- Budget statement: "Budget statement on how to deliver pre-budget"
- Business back: "Brown pauses while substance and style collide"
- David Cameron: "Cameron's all hit growth"
- Pakistan: "Bhutto deal"
- US: "US to meet Burma's leaders"
- More news: UN, Nepal, Kidnapped, North Korea, Marines
- Latest videos: Diana inquest, Simon Hoggart, Cronenberg chat, Alex James's Cheese Diaries, Emma Thompson on sex trafficking
- Music review: Sweet enough

Right Sidebar:

- Search bar: Search | Go
- Guardian Unlimited | Web
- Webfeed
- guardianjobs: Search all jobs | Go | Upload your CV and get jobs by email
- Find a date: Guardian Singles - over 65,000 members online
- Buy one get one FREE: Luxury Le Creuset for only £94.99
- Halloween Fright Lite
- On this site: Audio reports, Cartoons, Comment, leaders, letters and corrections, Crosswords, Email services, G2 features, In pictures, Interactive guides, News quiz, Notes and Queries, Obituaries



What visual features could be of interest?

What visual features could be of interest?



color contrast



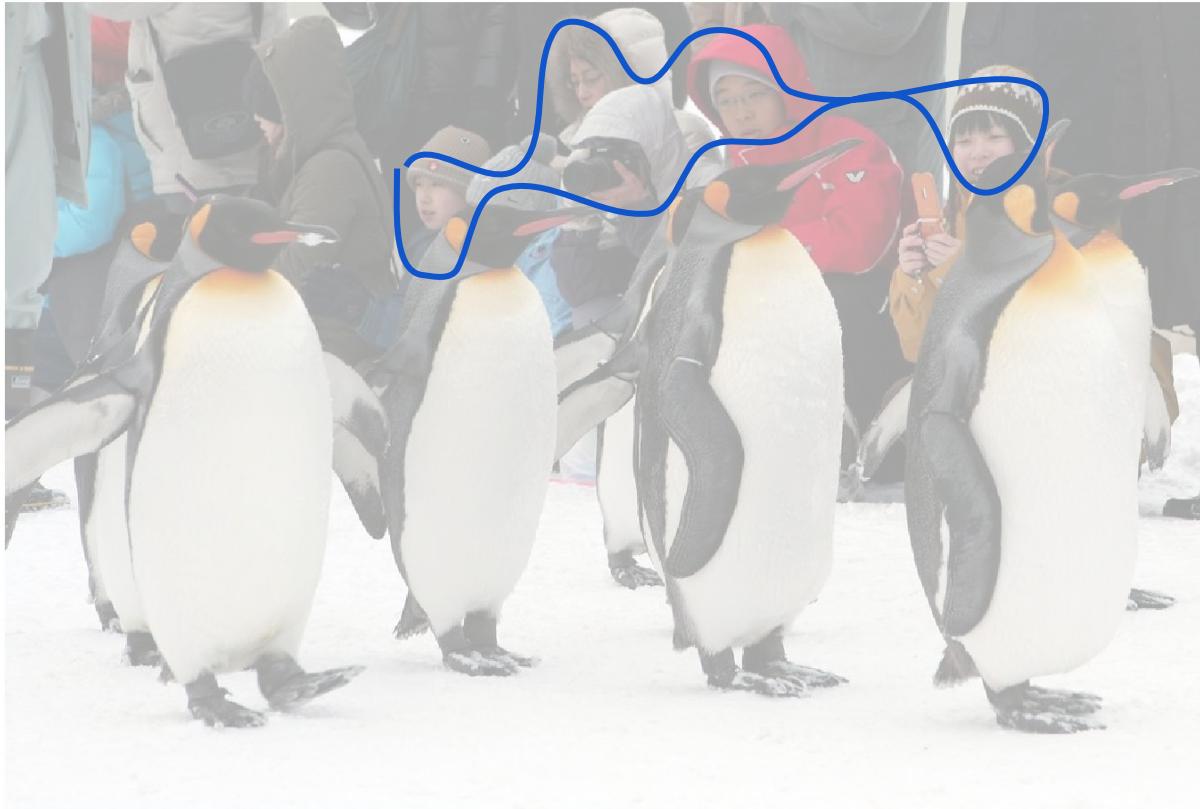
orientation contrast



faces and objects



human faces

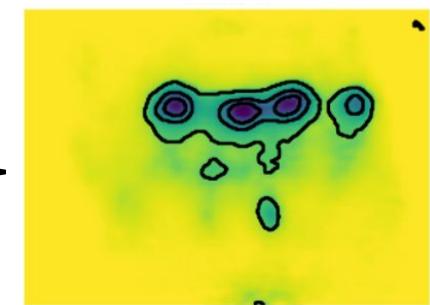
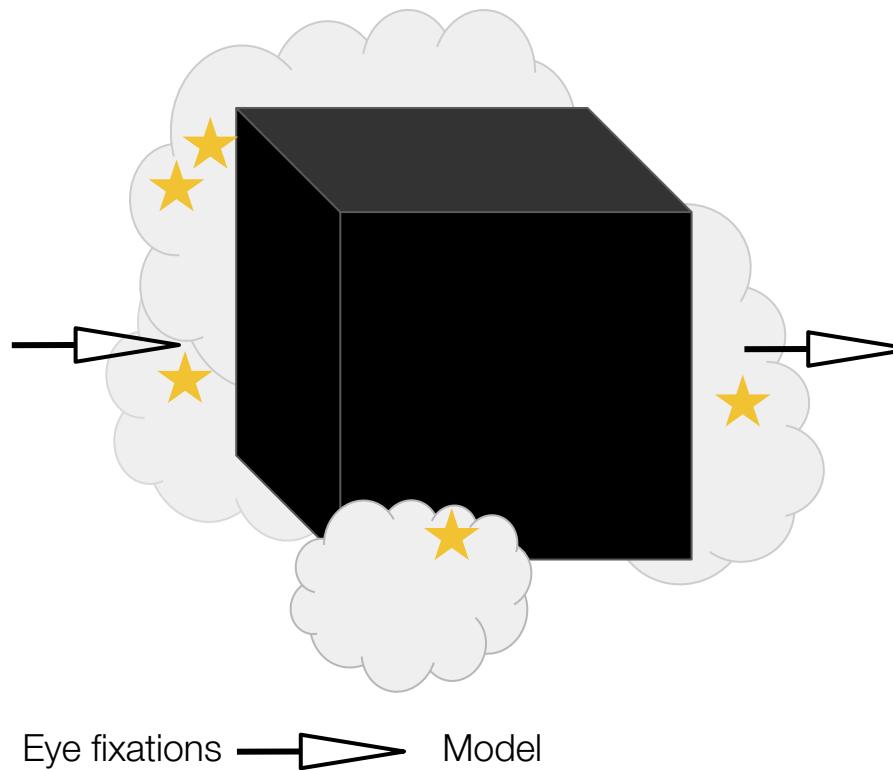


Experimental results



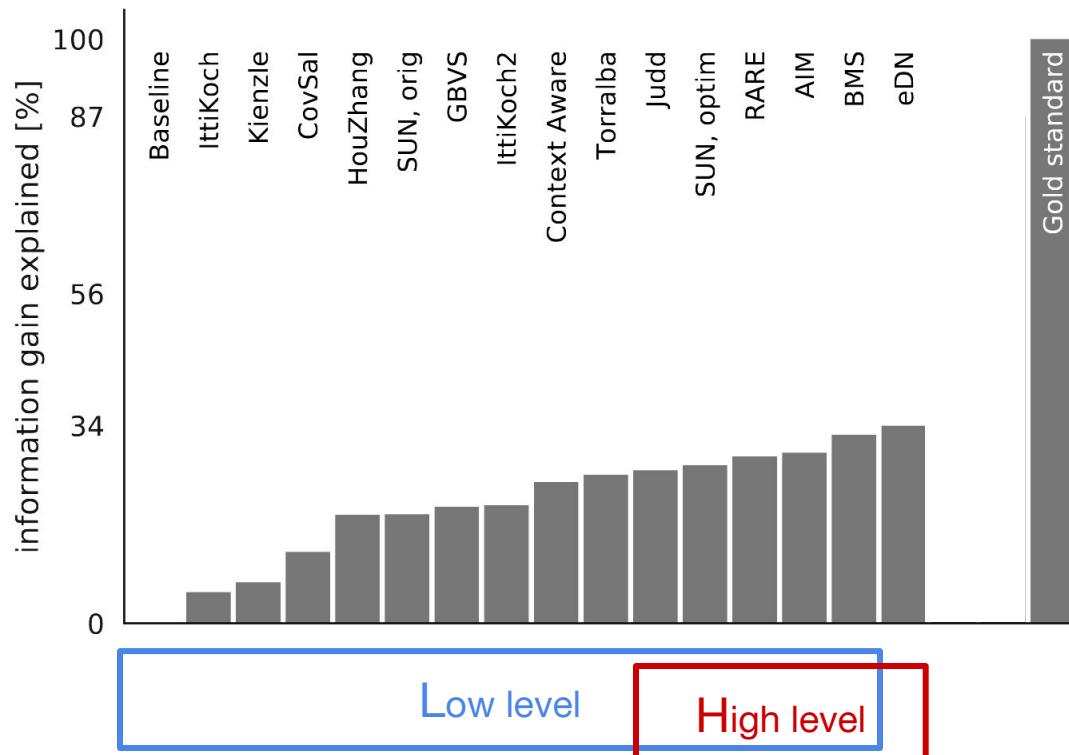


Input image

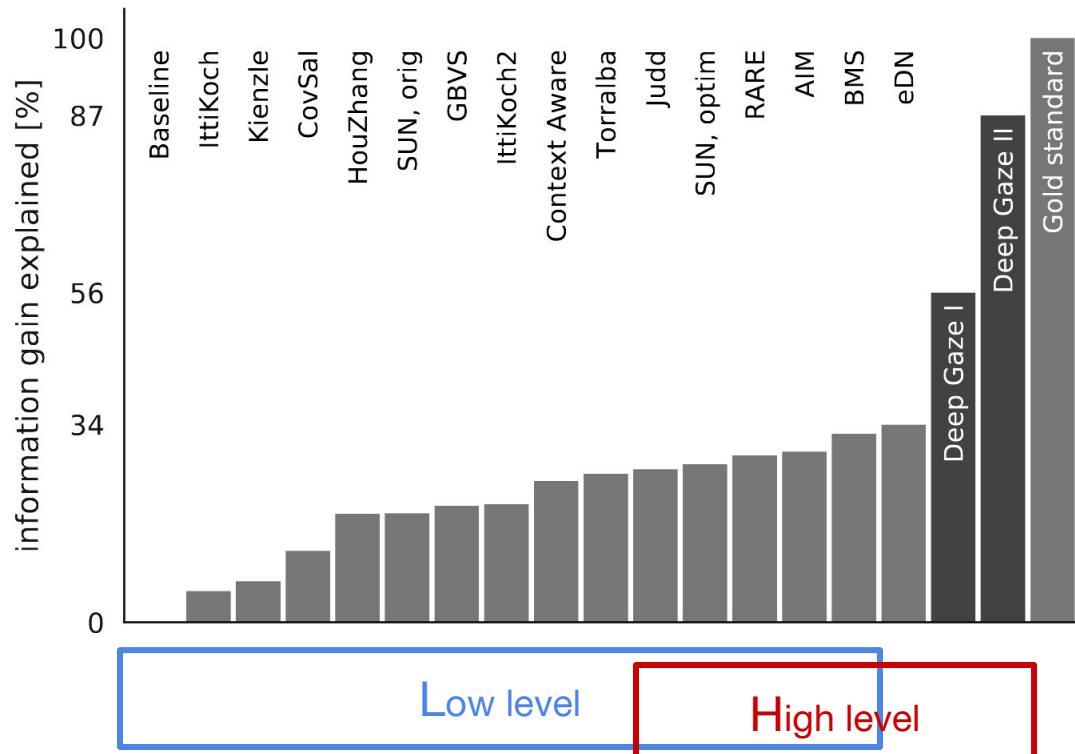


Saliency map

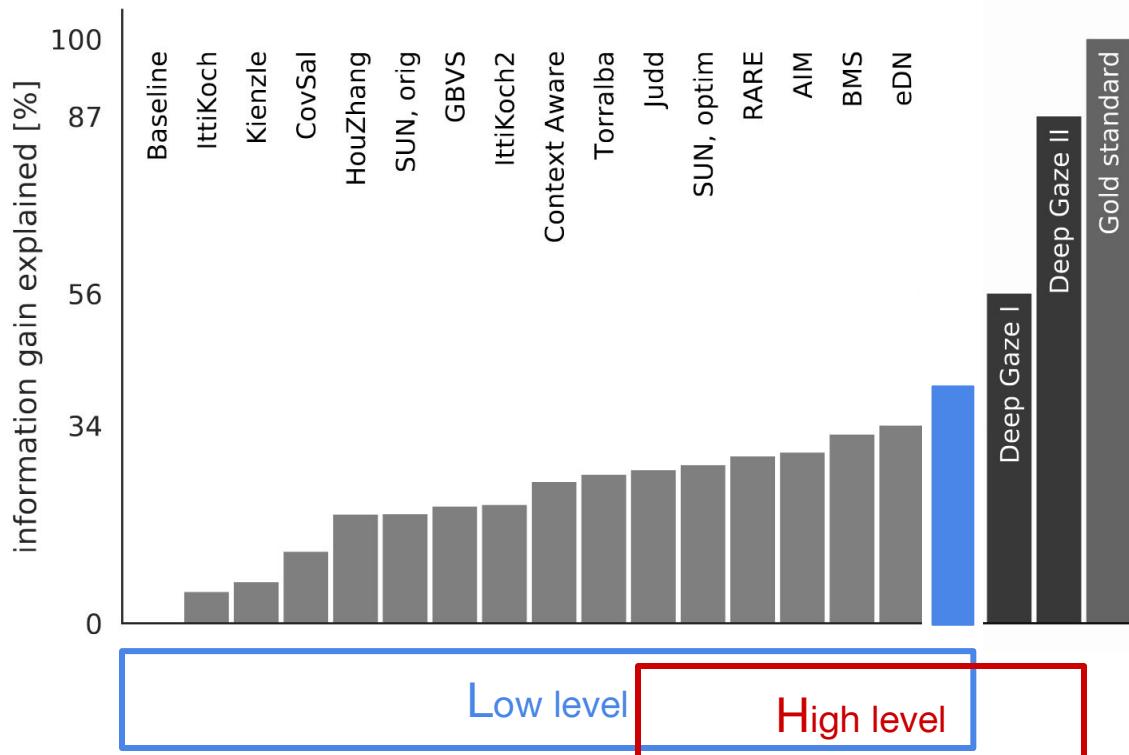
A zoo of models



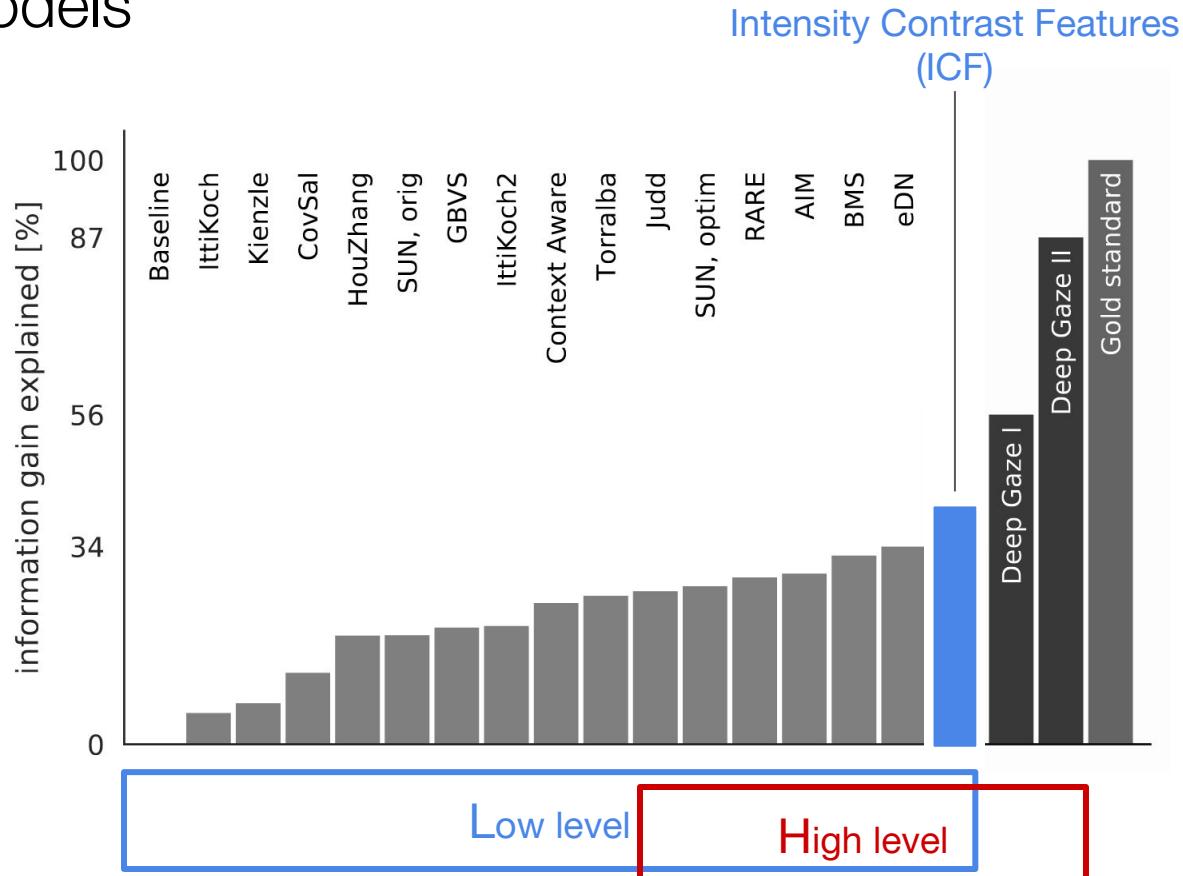
A zoo of models



A zoo of models

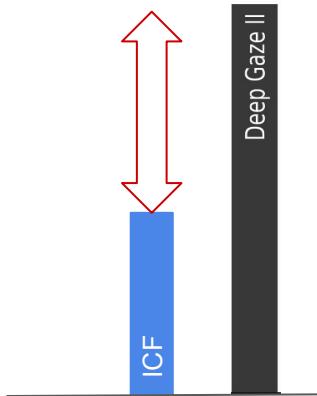


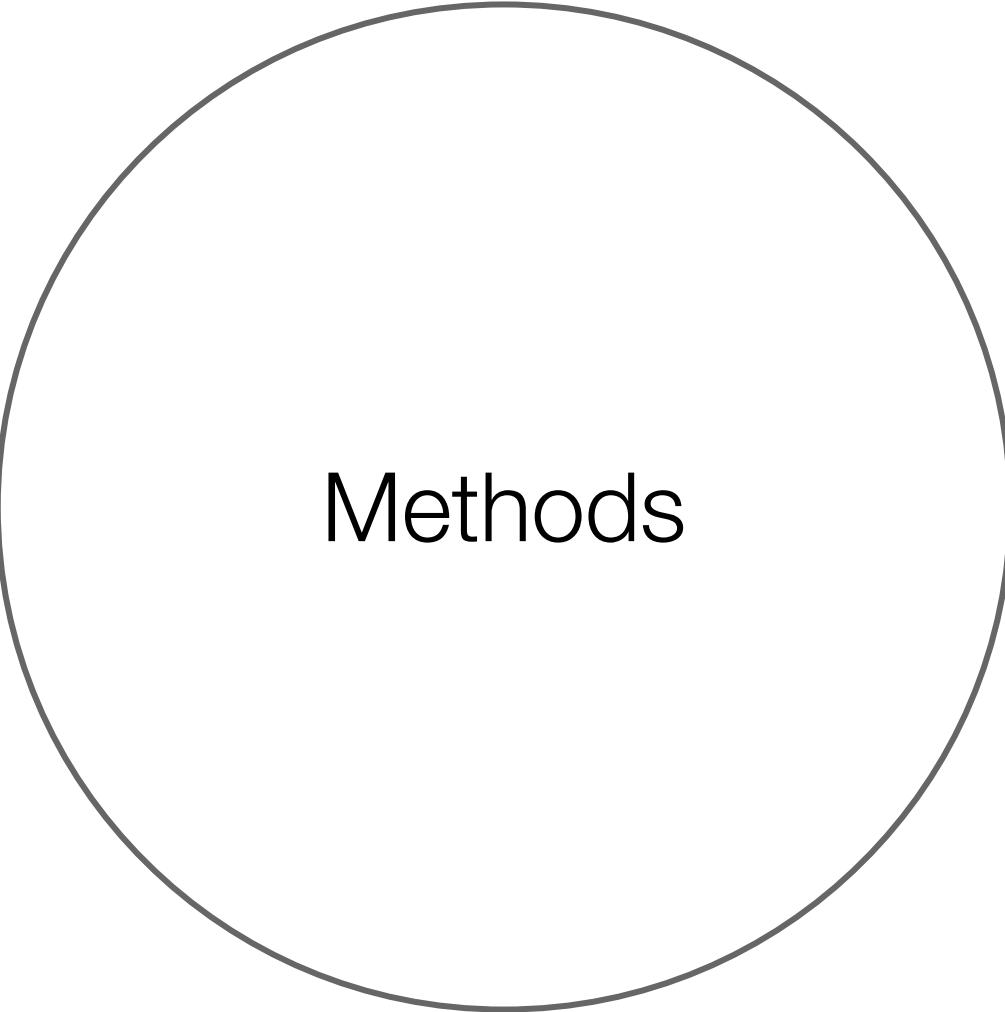
A zoo of models



Research questions

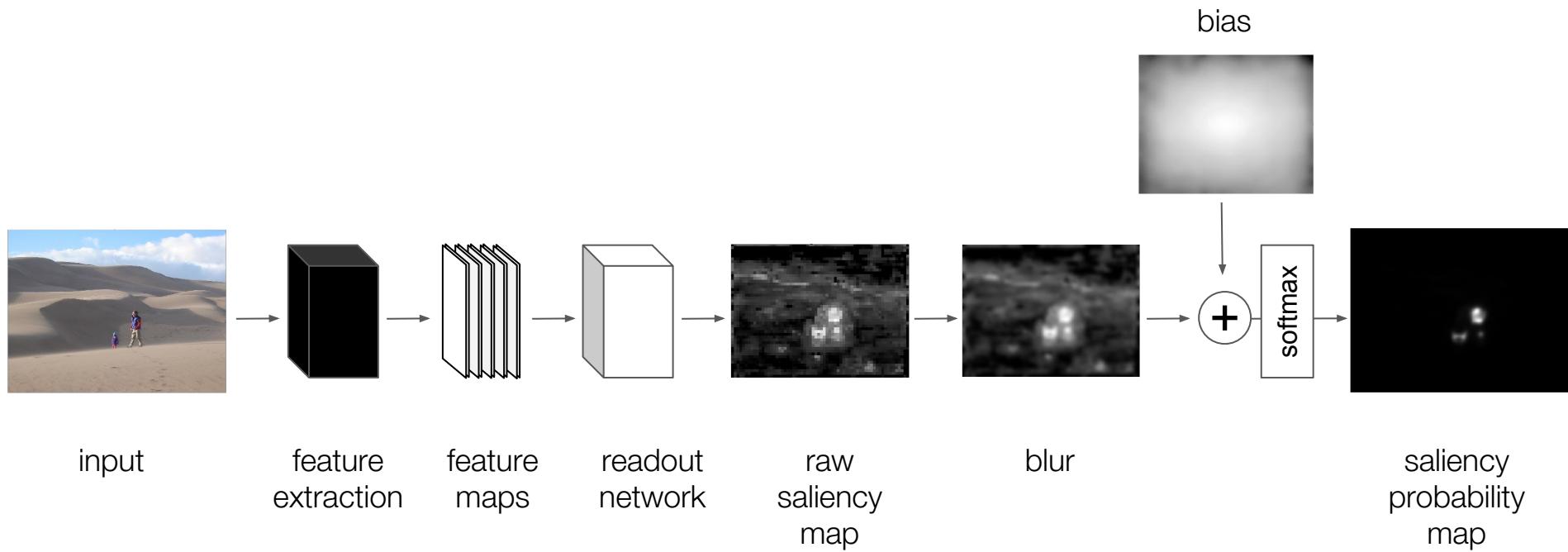
- What features drive saliency?
- How much saliency can be explained by low-level versus high-level features?
- Further reduction of the performance gap between high and low-level models?





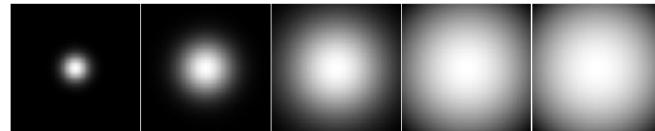
Methods

The model's basic architecture

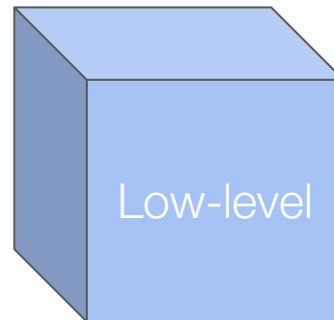


Models

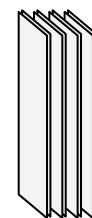
Intensity Contrast
Features (ICF)



input



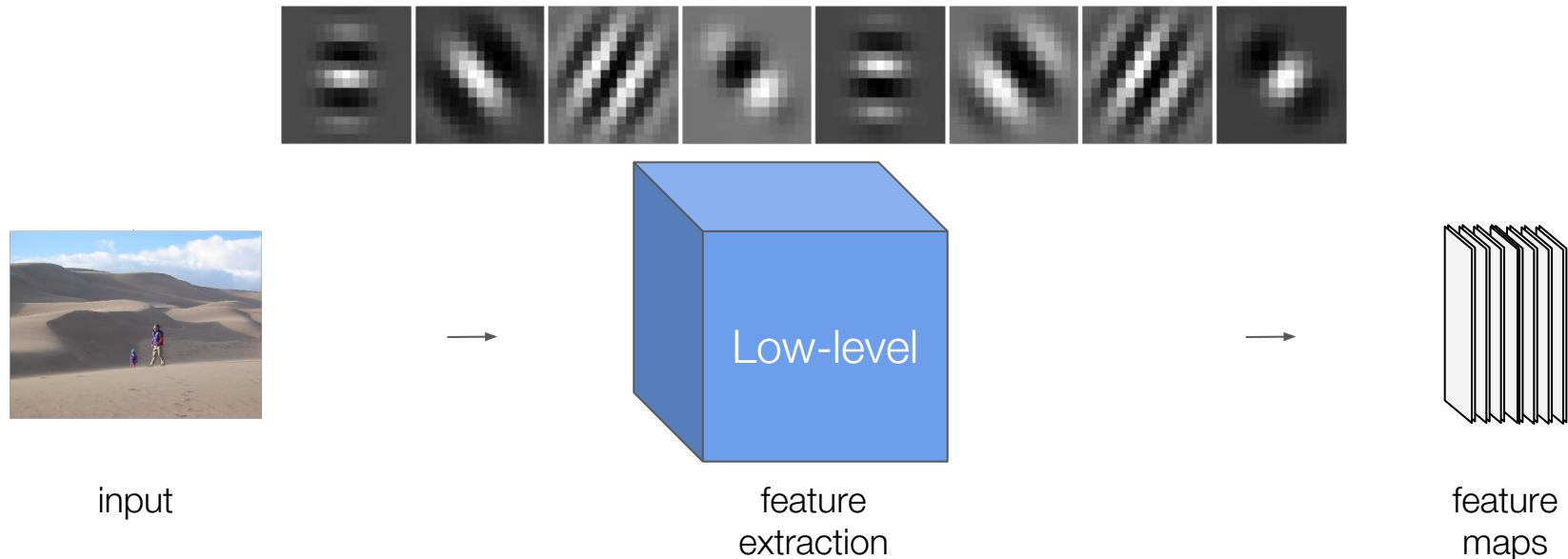
feature
extraction



feature
maps

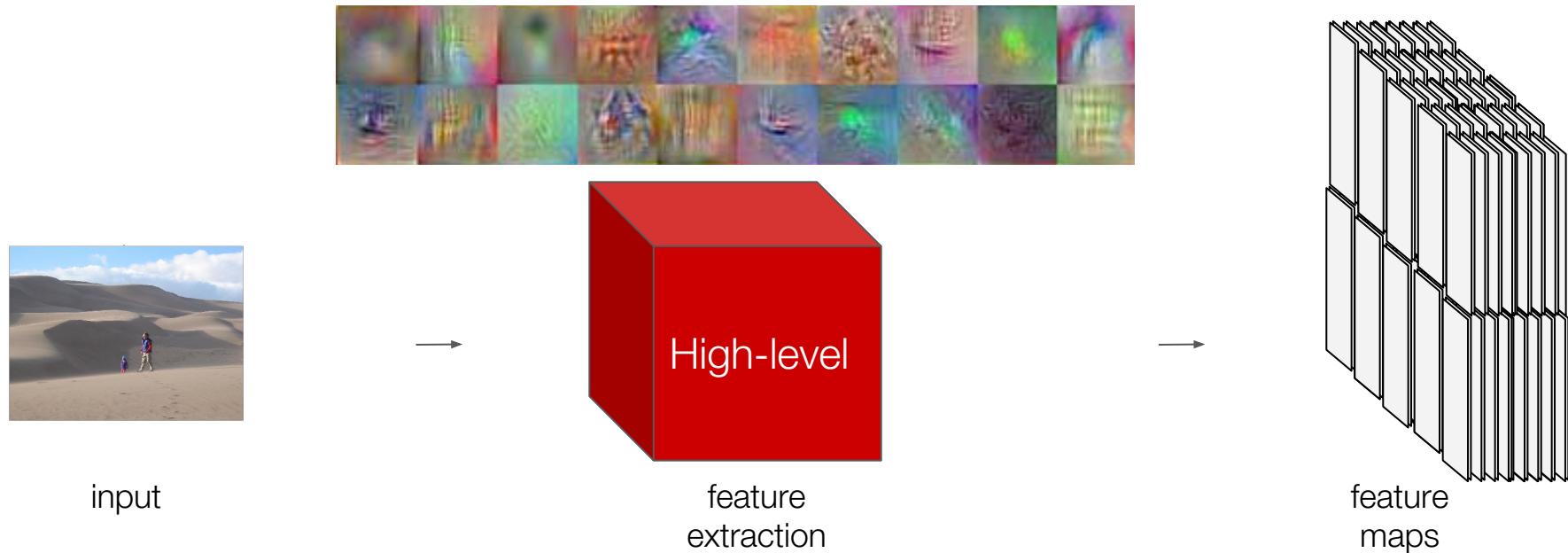
Models

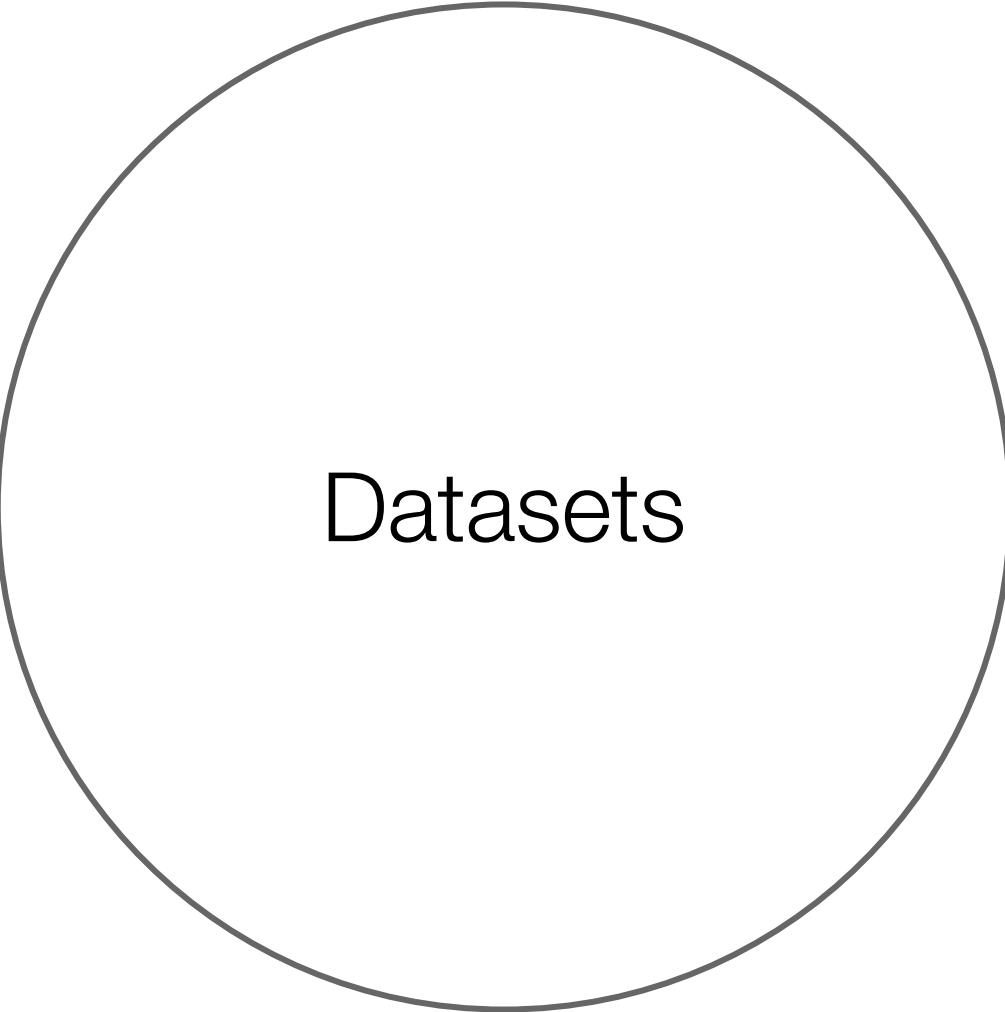
Gabor Model



Models

Deep Gaze II



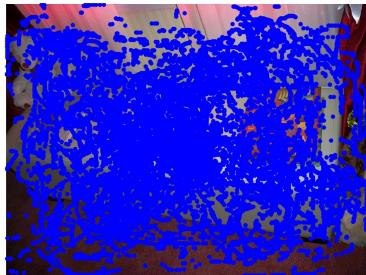
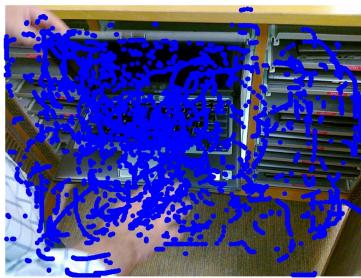
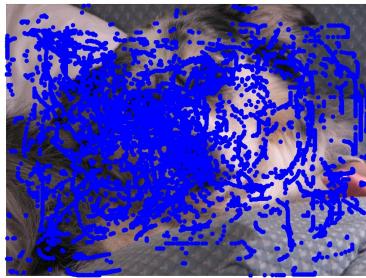


Datasets

Datasets

SALICON

10k images
mouse tracking

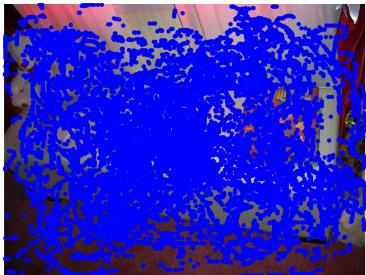
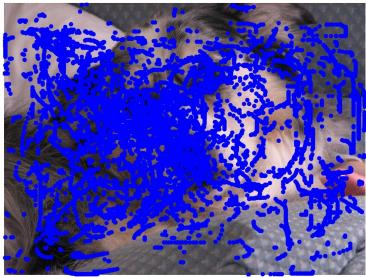


MIT1003

Datasets

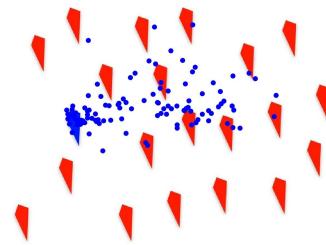
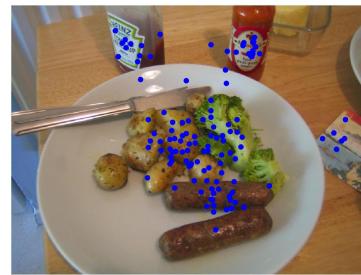
SALICON

10k images
mouse tracking



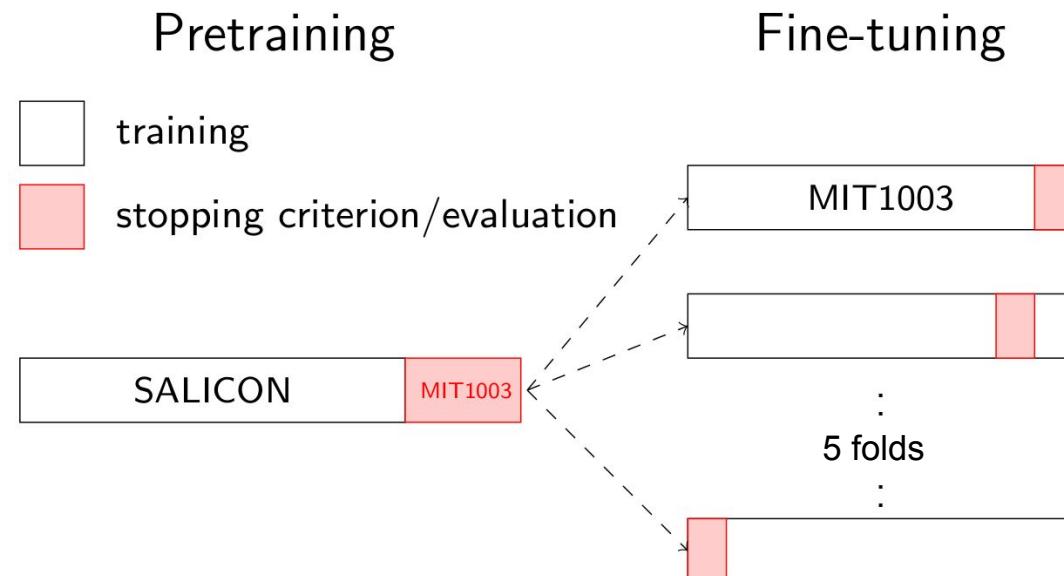
MIT1003

1k images
eye tracking



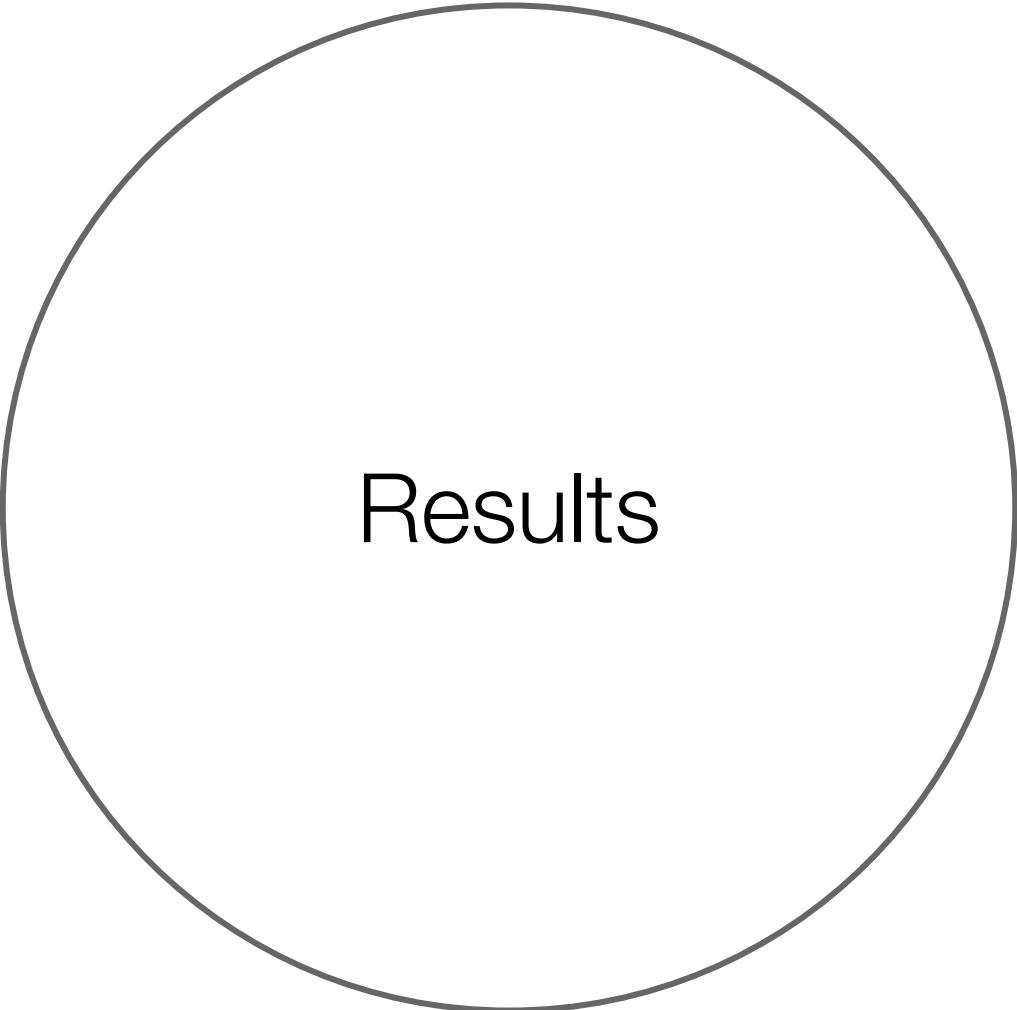
Pipeline

Loss



$$\frac{1}{N} \sum_i \log(p(x_i, y_i | I_i))$$

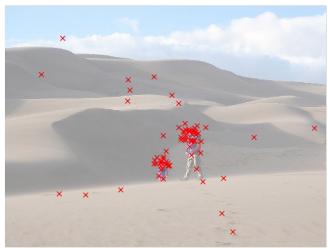
N Number of images
 x, y Pixel locations
 I Image



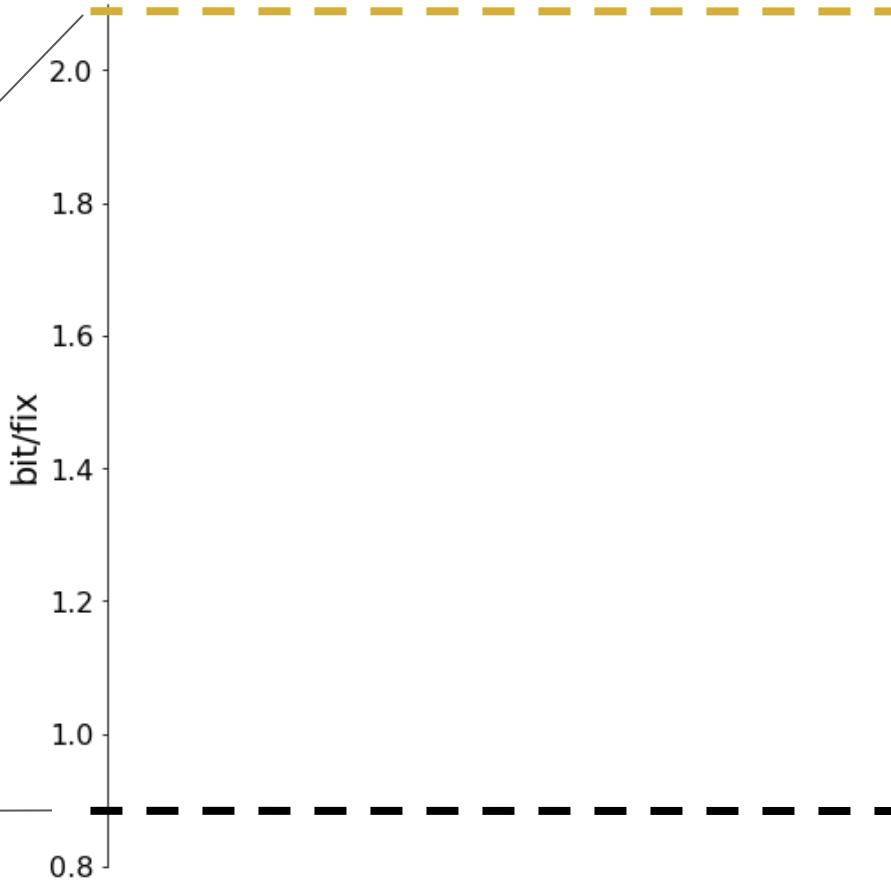
Results

Benchmarks

Experimental
Gold Standard

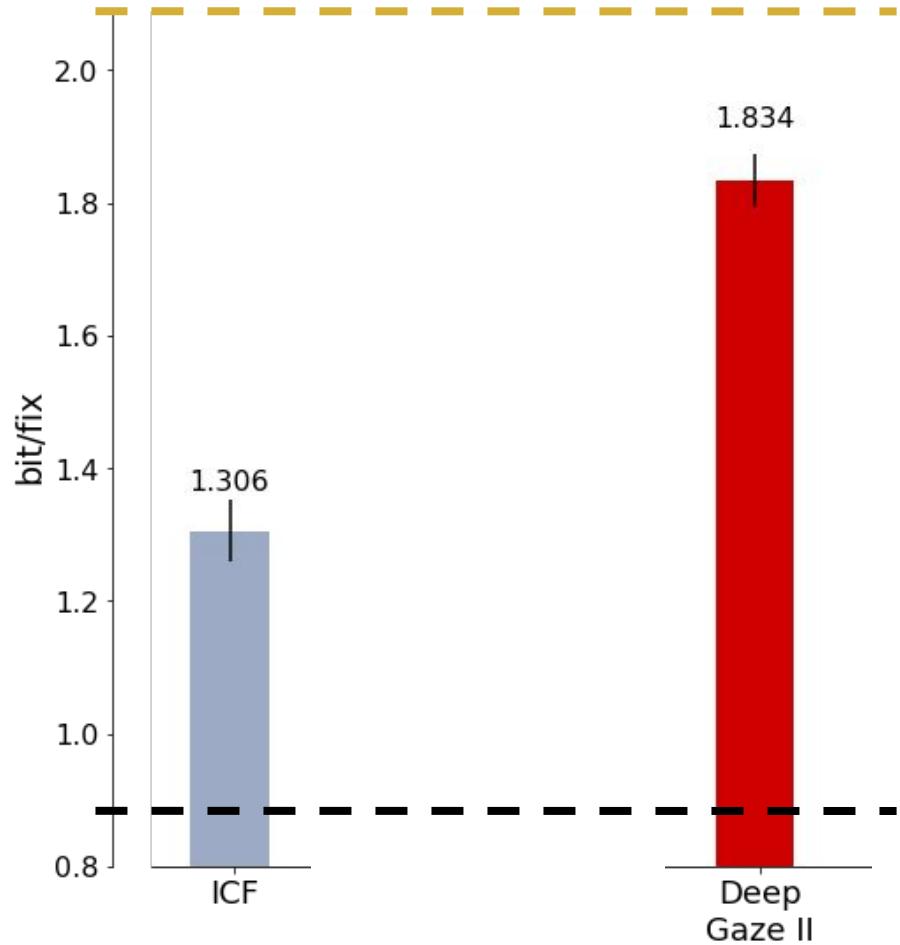


Centerbias



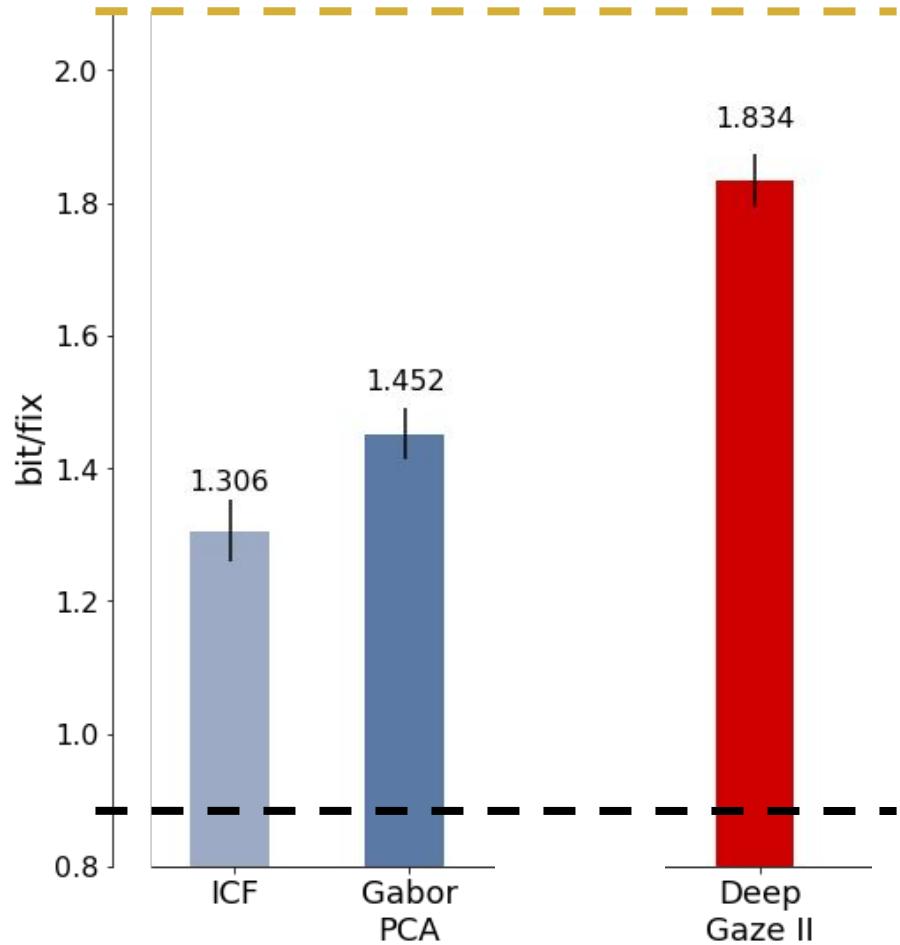
Benchmarks

- Reproduced results of Kümmerer et al.



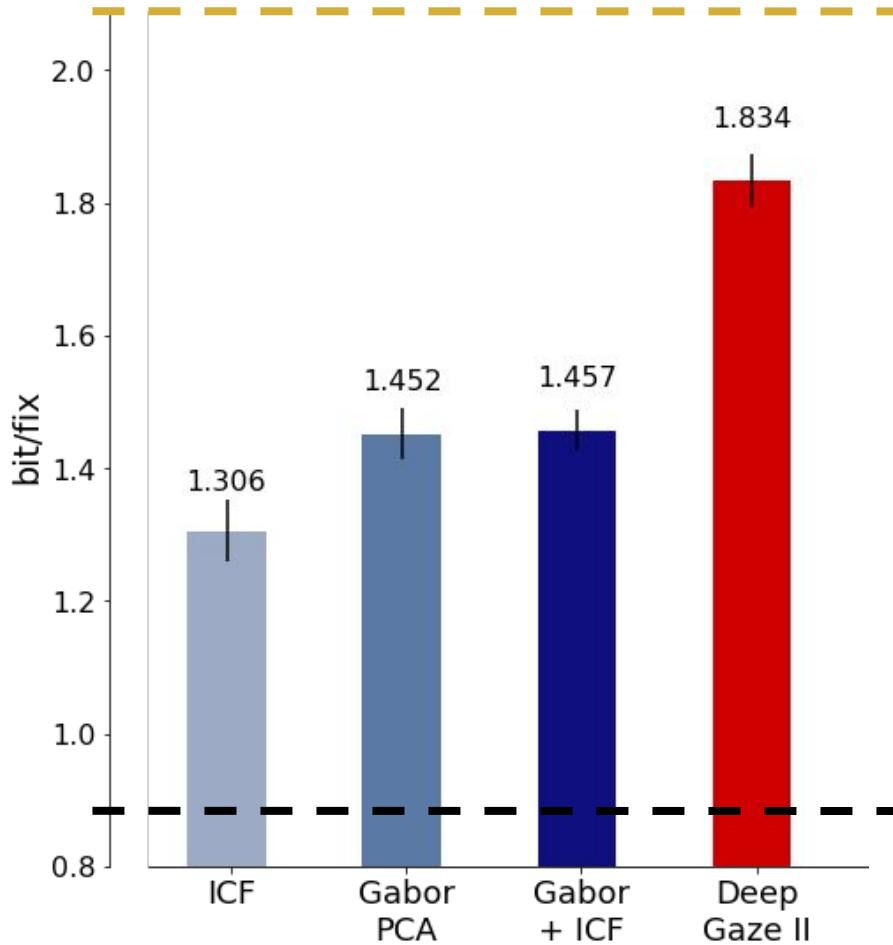
Benchmarks

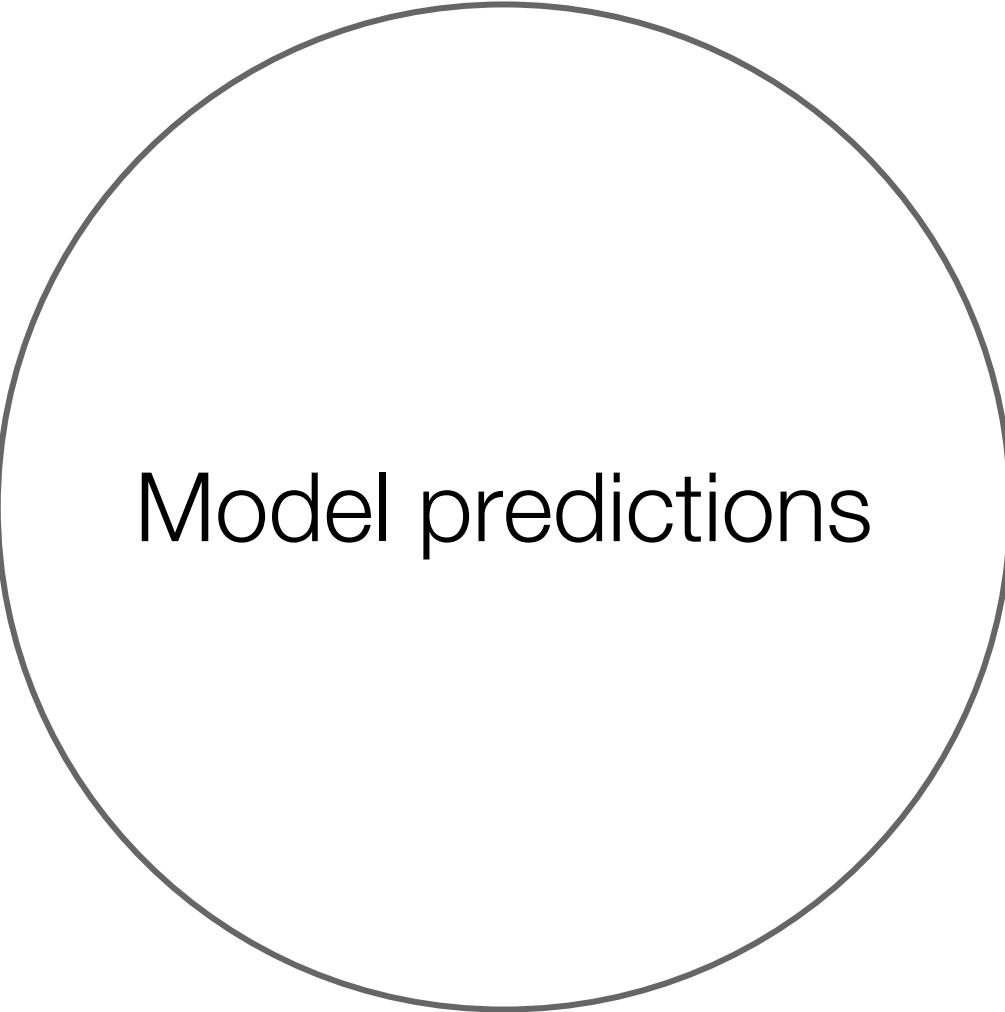
- Gabor features outperform Intensity Contrast Features (ICF)



Benchmarks

- Gabor features outperform Intensity Contrast Features (ICF)
- Combining ICF + Gabor features yields no improvement





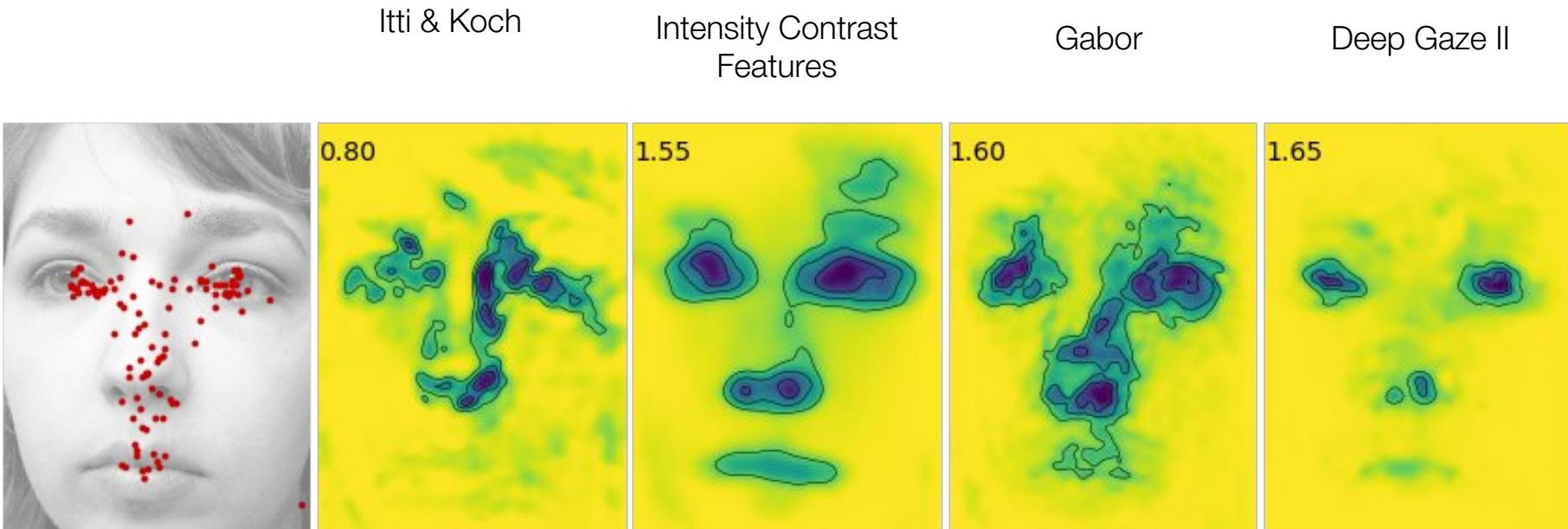
Model predictions

Predictions



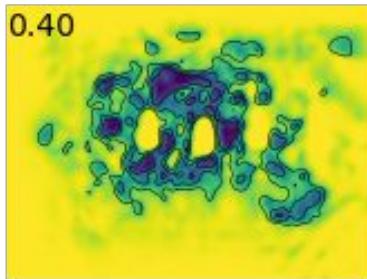
in bits/fix
the higher the better

Predictions

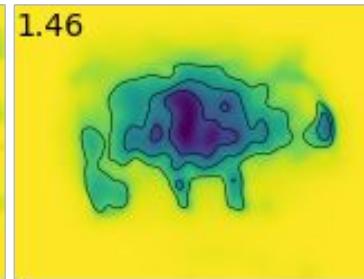


Predictions

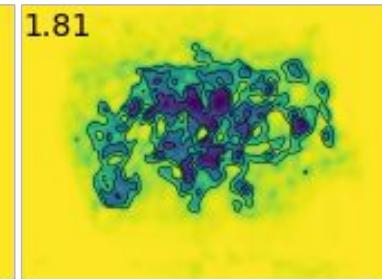
Itti & Koch



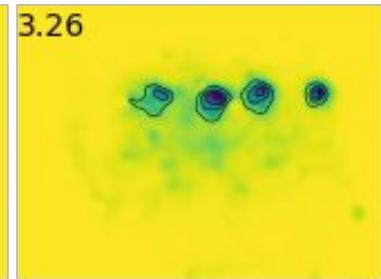
Intensity Contrast
Features



Gabor

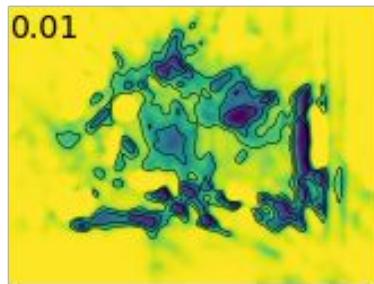


Deep Gaze II

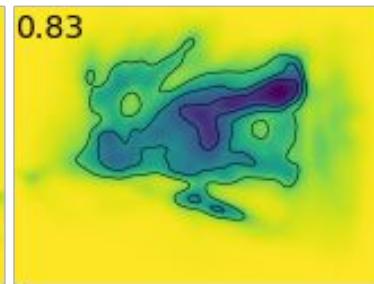


Predictions

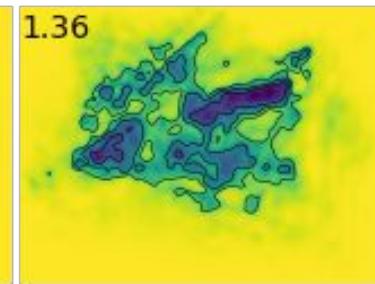
Itti & Koch



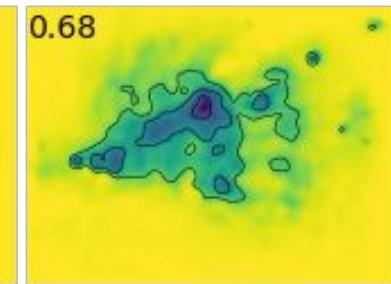
Intensity Contrast
Features



Gabor

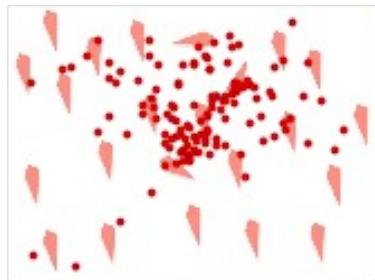


Deep Gaze II

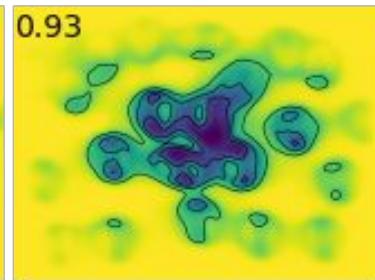


Predictions

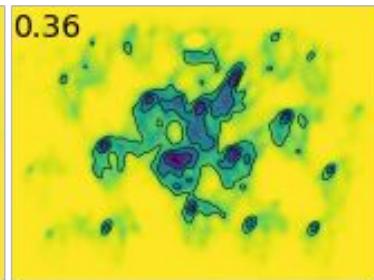
Itti & Koch



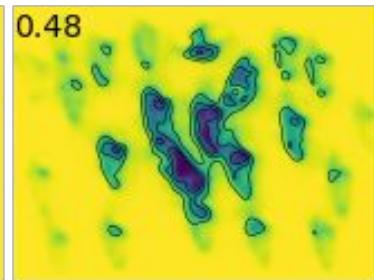
Intensity Contrast
Features



Gabor

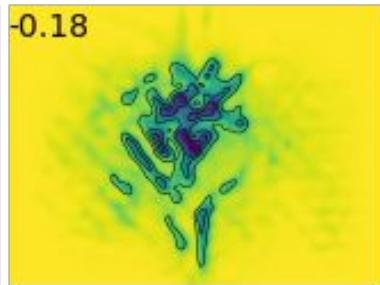
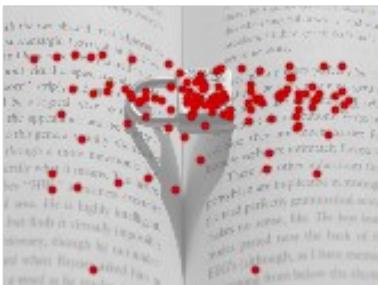


Deep Gaze II

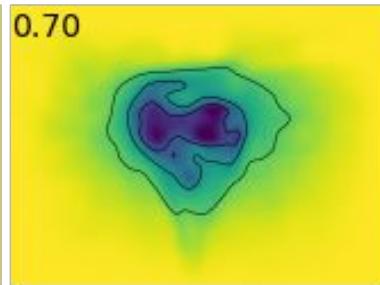


Predictions

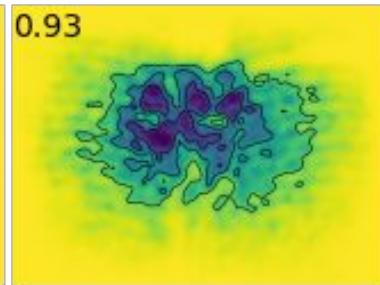
Itti & Koch



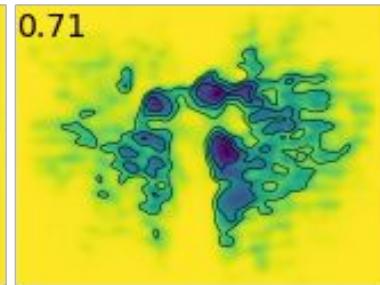
Intensity Contrast
Features



Gabor

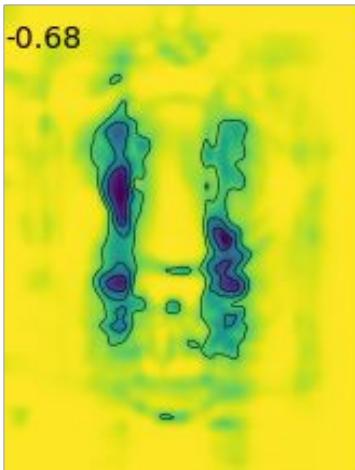


Deep Gaze II

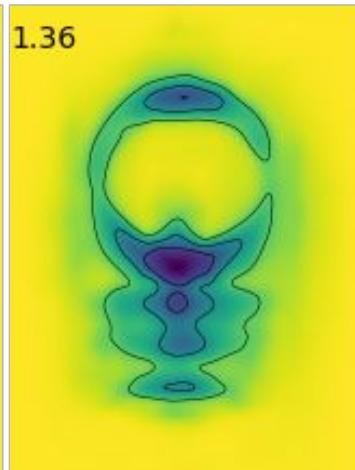


Predictions

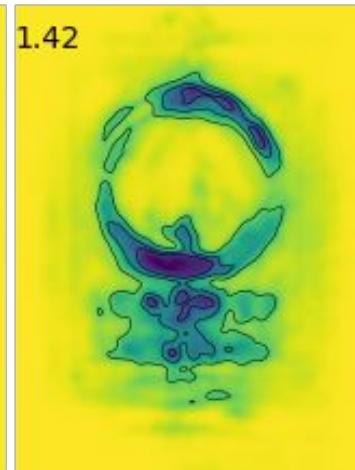
Itti & Koch



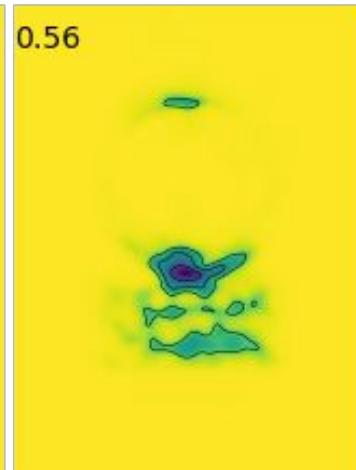
Intensity Contrast
Features



Gabor



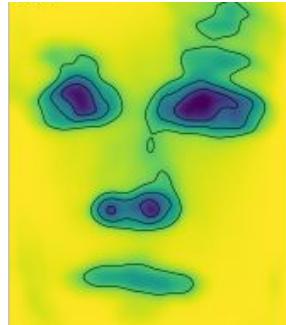
Deep Gaze II



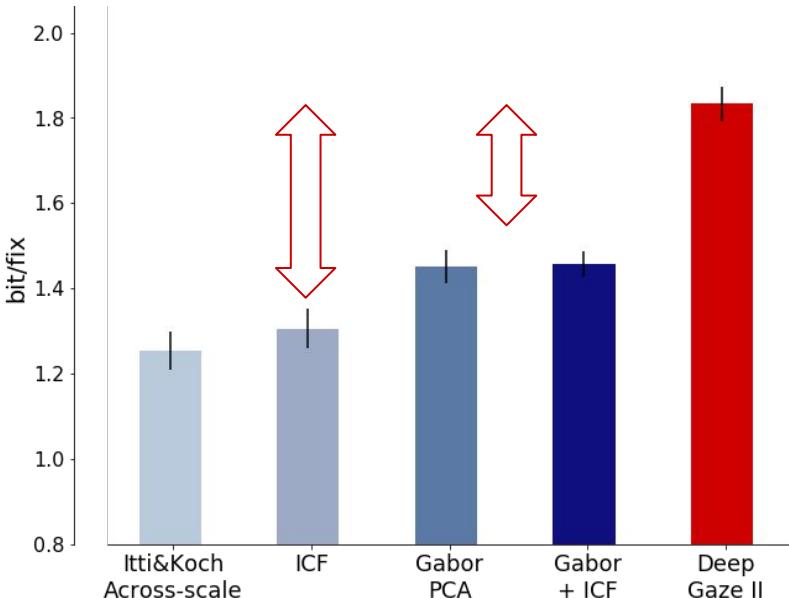


Summary & Discussion

Take-home



- Humans look at faces, text and objects **and** low-level feature contrast
- Saliency is driven both by high-level **and** low-level features
- Beware: Classic models perform poorly



Sources

Laurent Itti, Christof Koch, and Ernst Niebur. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(11):1254–1259, November 1998.

Ming Jiang, Shengsheng Huang, Juanyong Duan, and Qi Zhao. Salicon: Saliency in context. In *CVPR*, pages 1072–1080. IEEE Computer Society, 2015.

Tilke Judd, Krista Ehinger, Frédo Durand, and Antonio Torralba. Learning to predict where humans look. In *IEEE International Conference on Computer Vision (ICCV)*, 2009.

C. Koch and S. Ullman. Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry. *Human Neurobiology*, 4:219–227, January 1985.

M. Kümmeler, T. Wallis, and M. Bethge. How close are we to understanding image-based saliency? *arXiv*, Sep 2014.

M. Kümmeler, T. Wallis, and M. Bethge. Information-theoretic model comparison unifies saliency metrics. *Proceedings of the National Academy of Science*, 112(52):16054–16059, Oct 2015.

Matthias Kümmeler, Lucas Theis, and Matthias Bethge. Deep Gaze I: Boosting Saliency Prediction with Feature Maps Trained on ImageNet. *arXiv:1411.1045 [cs, q-bio, stat]*, November 2014. arXiv: 1411.1045.

Matthias Kümmeler, Thomas S. A. Wallis, and Matthias Bethge. DeepGaze II: Reading fixations from deep features trained on object recognition. *arXiv:1610.01563 [cs, q-bio, stat]*, October 2016. arXiv: 1610.01563.

A. M. Treisman and G. Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12:97–136, 1980.

Links

Evaluate own images on Deep Gaze II
<https://deepgaze.bethgelab.org/>

pysalency toolbox by M. Kümmeler
<https://github.com/matthias-k/pysalency>

My code for this project (work in progress)
<https://github.com/Kreiswolke/HighLowSaliency/tree/master/SaliencyModels>

MIT saliency benchmark challenge
<http://saliency.mit.edu/home.html>



Thanks for your attention,

questions?