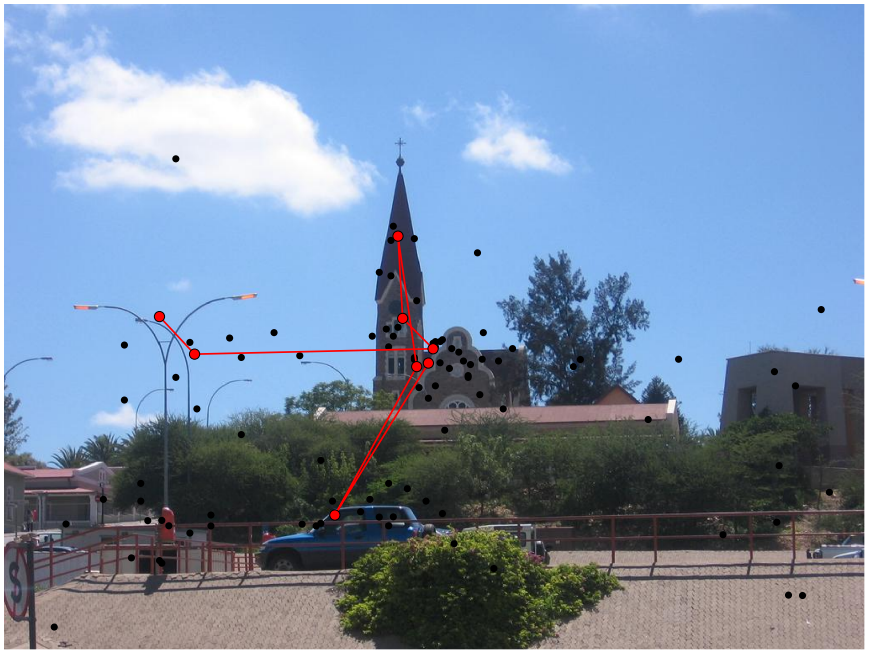
Probabilistic Evaluation of Saliency Models

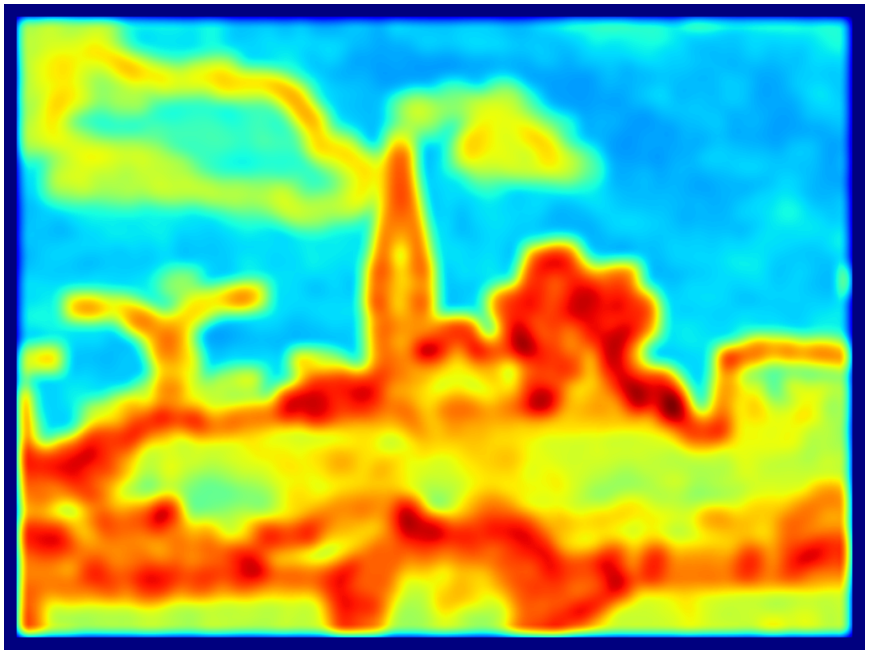
Matthias Kümmerer Matthias Bethge

Centre for Integrative Neuroscience, University of Tübingen, Germany

October 8, 2016

Saliency Maps





Model Name	Published	Code	AUC-Judd [?]	SIM [?]	EMD [?]	AUC-Borji [?]	sAUC [?]	CC [?]	NSS [?]	KL [?]	Date tested [key]	Sample [img]
Baseline: infinite humans [?]			0.92	1	0	0.88	0.81	1	3.29	0		
Deep Gaze 2	Matthias Kümmerer, Lucas Theis, Matthias Bethge. **Deep Gaze I: Boosting Saliency Prediction with Feature Maps Trained on ImageNet [arxiv 2014]**		0.87	0.46	4.00	0.86	0.76	0.51	1.29	0.97	first tested: 26/11/2015 last tested: 22/02/2016 maps from authors	
SALICON	Xun Huang, Chengyao Shen, Xavier Boix, Qi Zhao		0.87	0.60	2.62	0.85	0.74	0.74	2.12	0.54	first tested: 19/11/2014 last tested: 15/11/2015 maps from authors	
DeepFix	Srinivas S S Kruthiventi, Kumar Ayush, R. Venkatesh Babu **DeepFix: A Fully Convolutional Neural Network for predicting Human Eye Fixations [arXiv 2015]**		0.87	0.67	2.04	0.80	0.71	0.78	2.26	0.63	first tested: 02/10/2015 last tested: 02/10/2015 maps from authors	
Probability Distribution Prediction (PDP)	Anonymous.		0.85	0.60	2.58	0.80	0.73	0.70	2.05	0.92	first tested: 05/11/2015 last tested: 05/11/2015 maps from authors	
Deep Gaze 1	Matthias Kümmerer, Lucas Theis, Matthias Bethge. **Deep Gaze I: Boosting Saliency Prediction with Feature Maps Trained on ImageNet [arxiv 2014]**		0.84	0.39	4.97	0.83	0.66	0.48	1.22	1.23	first tested: 02/10/2014 last tested: 15/11/2015 maps from authors	

mit.saliency.edu

How to judge saliency model performance?

- Common intuition: Higher saliency corresponds to more fixations

- Common intuition: Higher saliency corresponds to more fixations
- Saliency is operationalised by measuring fixation densities:

$$p(x, y)$$

History of Probablistic Modelling in Saliency

- Vincent et al: Do we look at lights? Vis.Cog. 2009
- Barthelmé et al: Modelling fixation locations using spatial point processes, JoV 2013
- Kümmerer et al: Information-theoretic model comparison unifies saliency metrics, PNAS 2015
- Kümmerer et al: DeepGaze I, ICLR Worshop 2015

Information theory provides a principled and accepted way to assess how well a model predicts the true density

- The average *log-likelihood* of a model is

$$\sum_{i}^{N} \frac{1}{N} \log \hat{p}(x_i, y_i \mid I_i)$$

for fixations (x_i, y_i) on images I_i and a model $\hat{p}(x, y \mid I)$

- The average *log-likelihood* of a model is

$$\sum_{i}^{N} \frac{1}{N} \log \hat{p}(x_i, y_i \mid I_i)$$

for fixations (x_i, y_i) on images I_i and a model $\hat{p}(x, y \mid I)$

- *Information gain*: Average log-likelihood relative to baseline

$$\mathrm{IG}(\hat{p} \| p_{\mathrm{bl}}) = \frac{1}{N} \sum_{i}^{N} \log \hat{p}(x_i, y_i \mid I_i) - \frac{1}{N} \sum_{i}^{N} \log p_{\mathrm{bl}}(x_i, y_i)$$

- The average *log-likelihood* of a model is

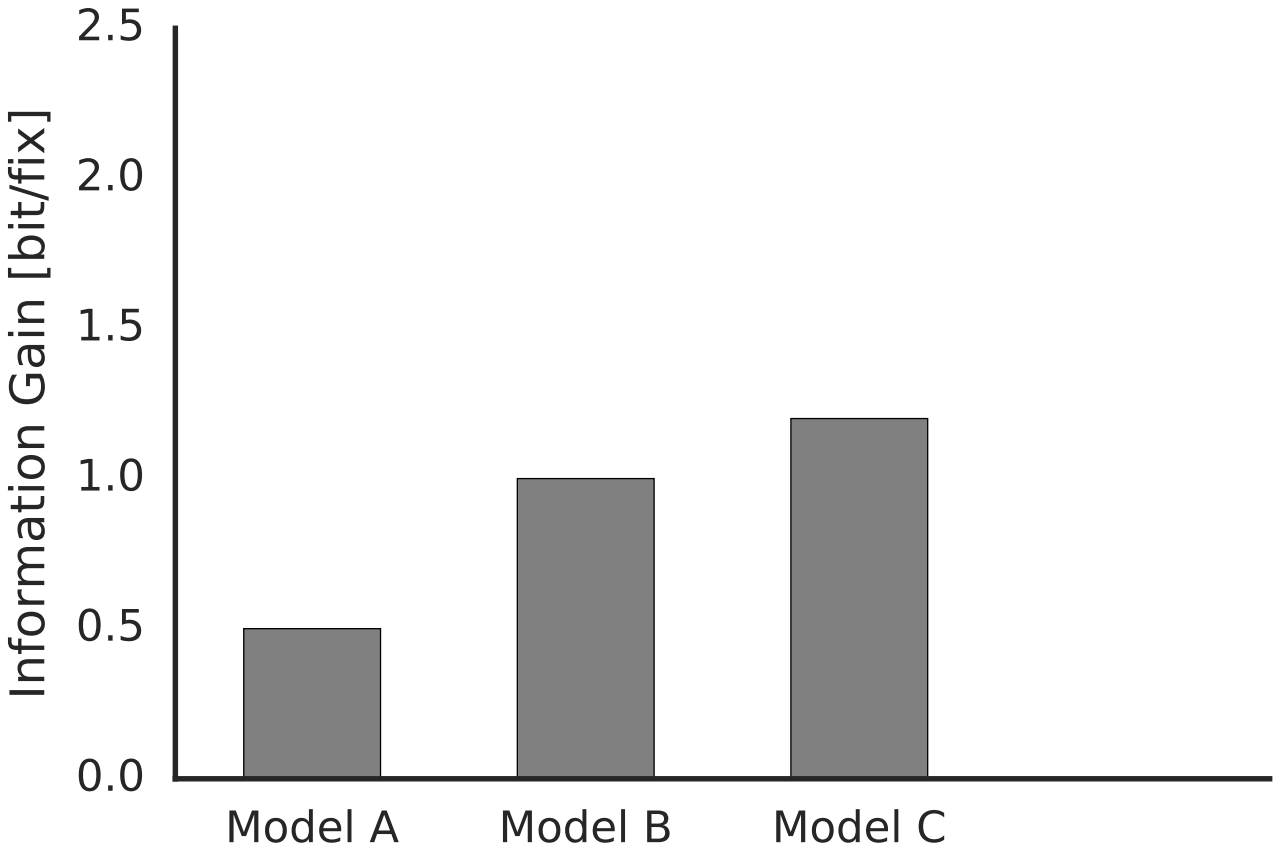
$$\sum_{i}^{N} \frac{1}{N} \log \hat{p}(x_i, y_i \mid I_i)$$

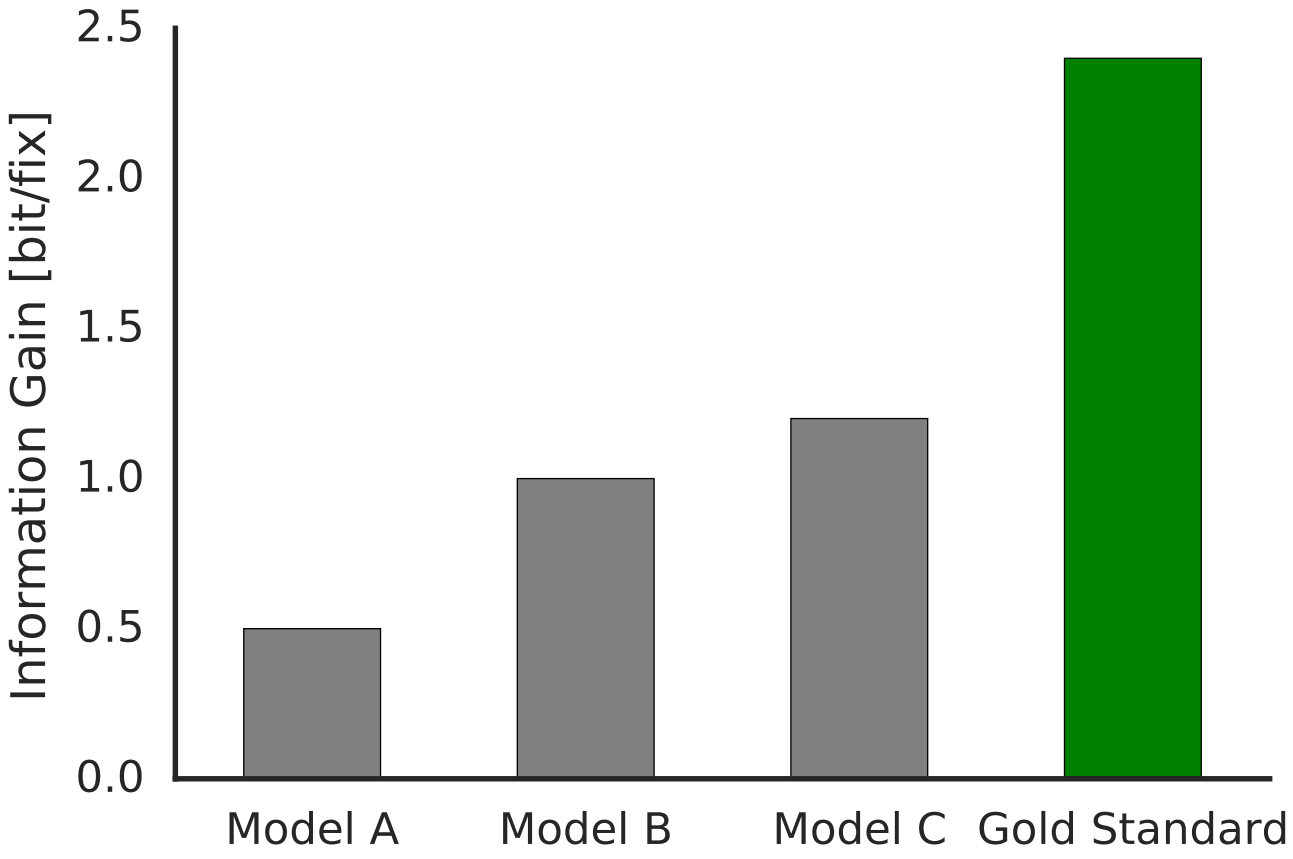
for fixations (x_i, y_i) on images I_i and a model $\hat{p}(x, y \mid I)$

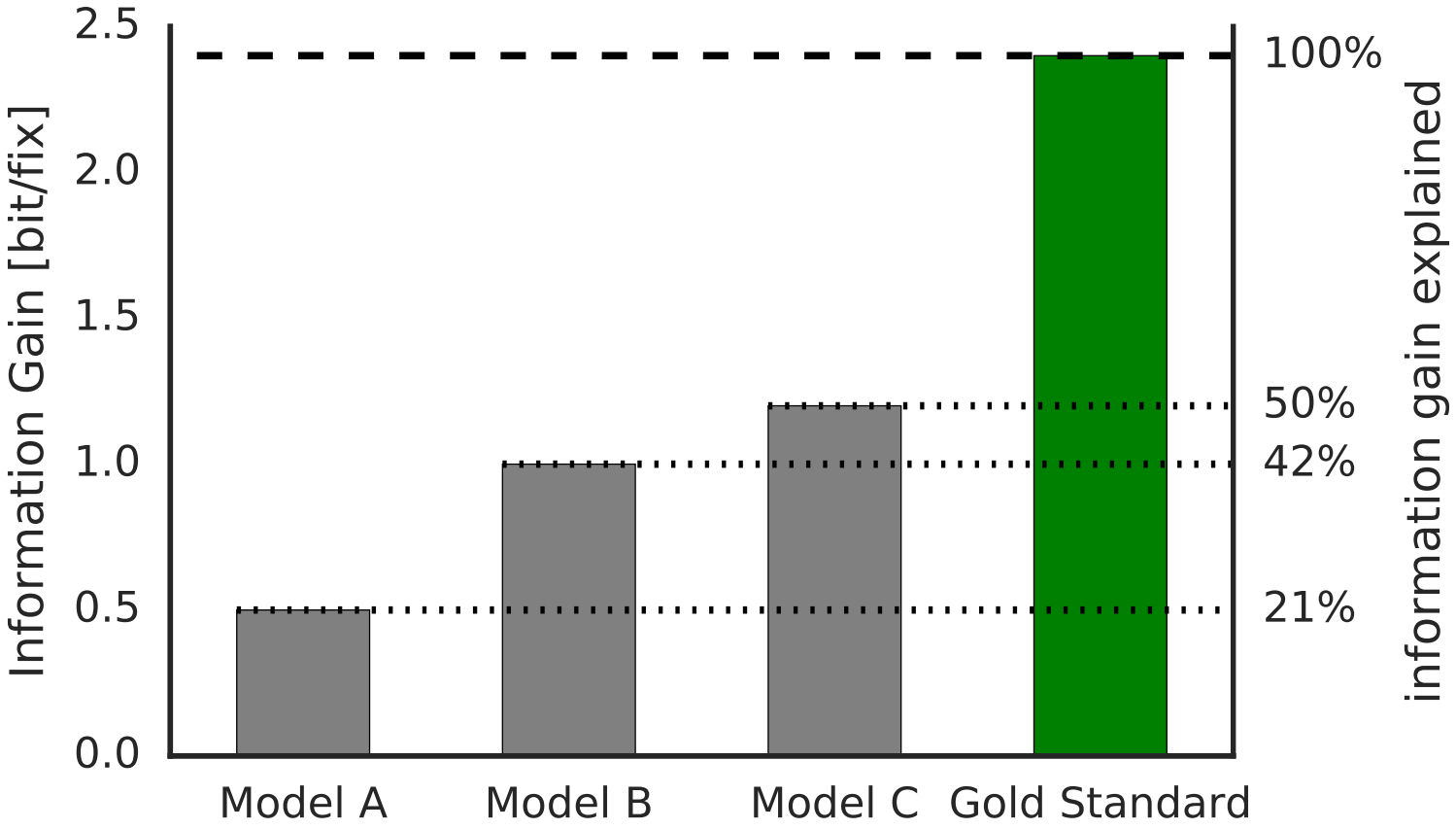
- *Information gain*: Average log-likelihood relative to baseline

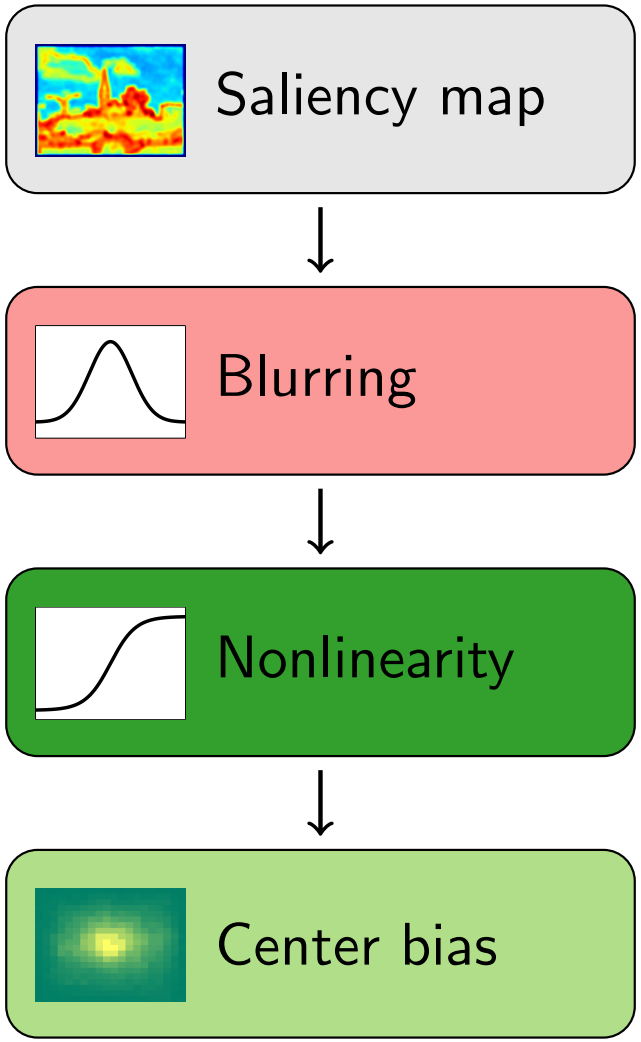
$$\mathrm{IG}(\hat{p} \| p_{\mathrm{bl}}) = \frac{1}{N} \sum_{i}^{N} \log \hat{p}(x_i, y_i \mid I_i) - \frac{1}{N} \sum_{i}^{N} \log p_{\mathrm{bl}}(x_i, y_i)$$

- Interpretation: In a game of 20 questions, how many questions does the model save compared to baseline when trying to find the location of a fixation


Information Gain [bit/fix]
2.5
2.0
1.5
1.0
0.5
0.0
Model A
Model B
Model C


Information Gain [bit/fix]
2.5
2.0
1.5
1.0
0.5
0.0
Model A
Model B
Model C
Gold Standard

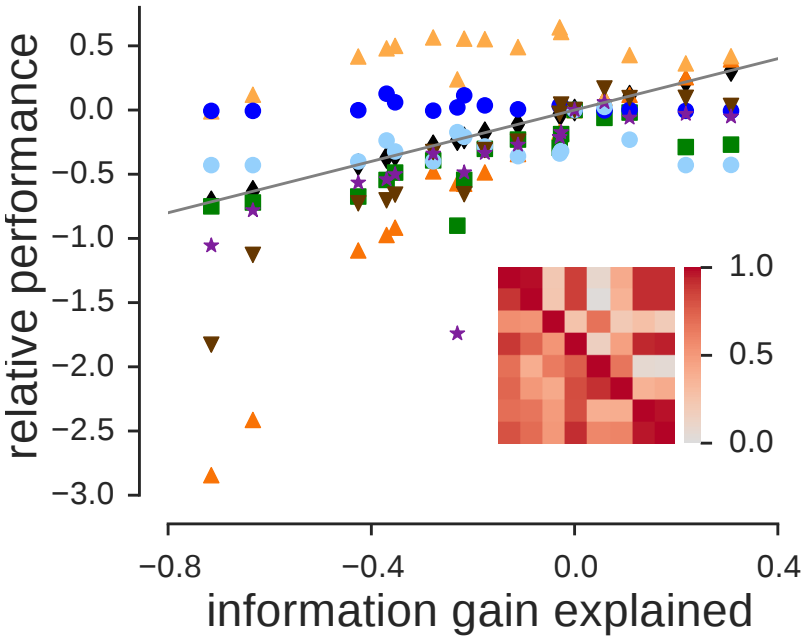

2.5
2.0
1.5
1.0
0.5
0.0
Information Gain [bit/fix]
100%
50%
42%
21%
information gain explained
Model A
Model B
Model C
Gold Standard



http://github.com/matthias-k/pysaliency

Resolving the metric inconsistencies

.

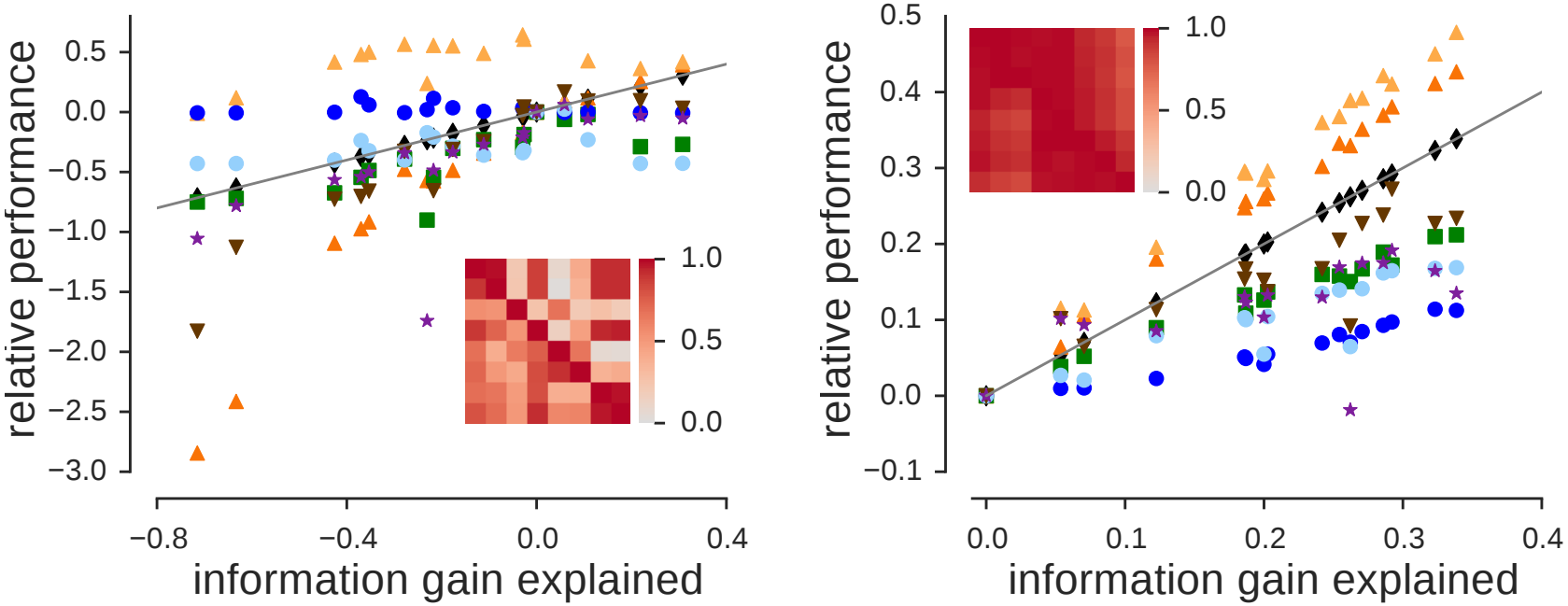


- ♦ information gain explained
- ▲ AUC wrt. uniform
- ▲ AUC wrt. center bias
- ■ i.b. D_{KL}
- ● f.b. D_{KL} wrt. center bias
- ● f.b. D_{KL} wrt. uniform
- ▼ NSS
- ★ CC

Kümmerer et al., PNAS 2015

Resolving the metric inconsistencies

.

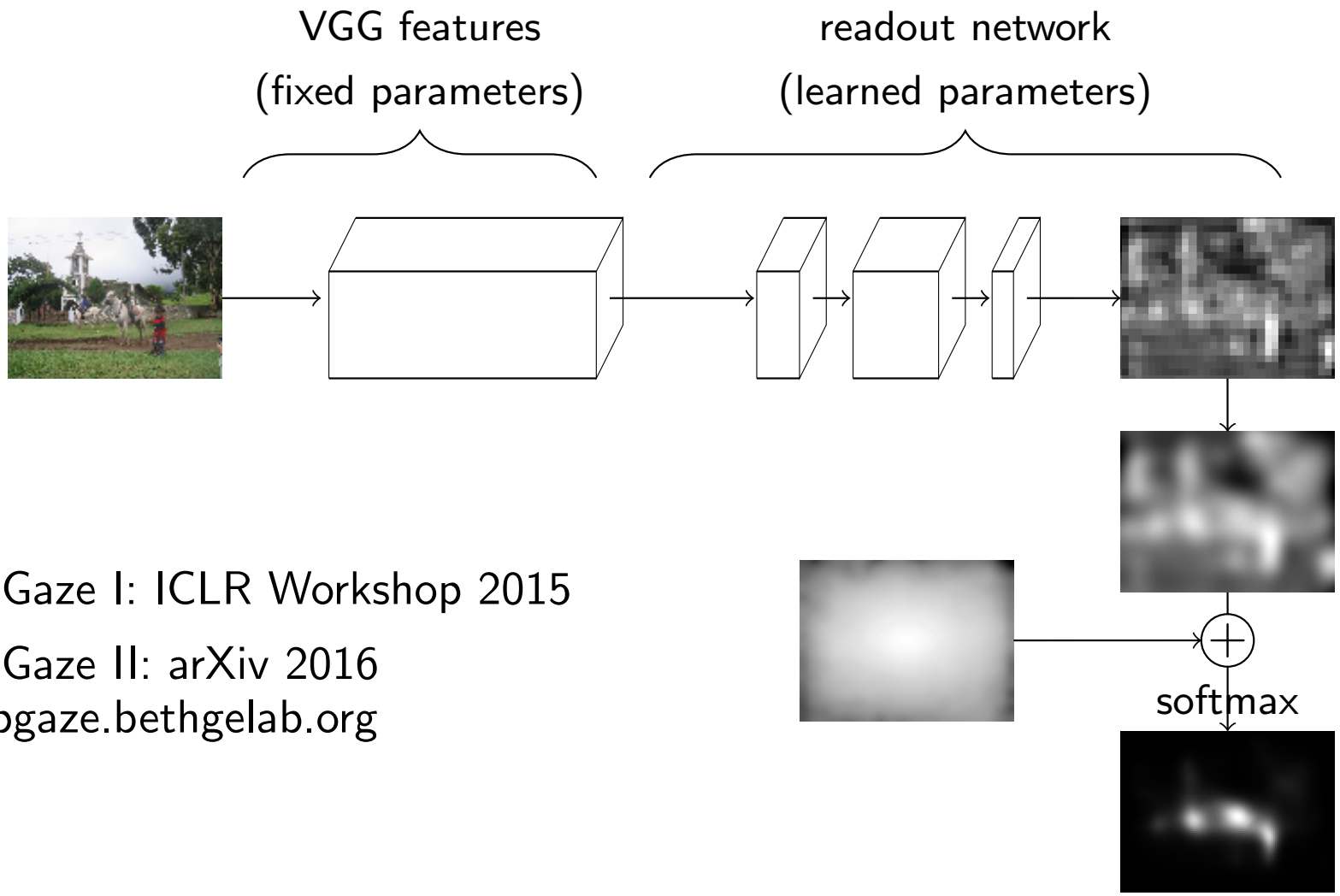


- information gain explained
- AUC wrt. uniform
- AUC wrt. center bias
- i.b. D_{KL}
- f.b. D_{KL} wrt. center bias
- f.b. D_{KL} wrt. uniform
- NSS
- CC

Kümmerer et al., PNAS 2015

DeepGaze

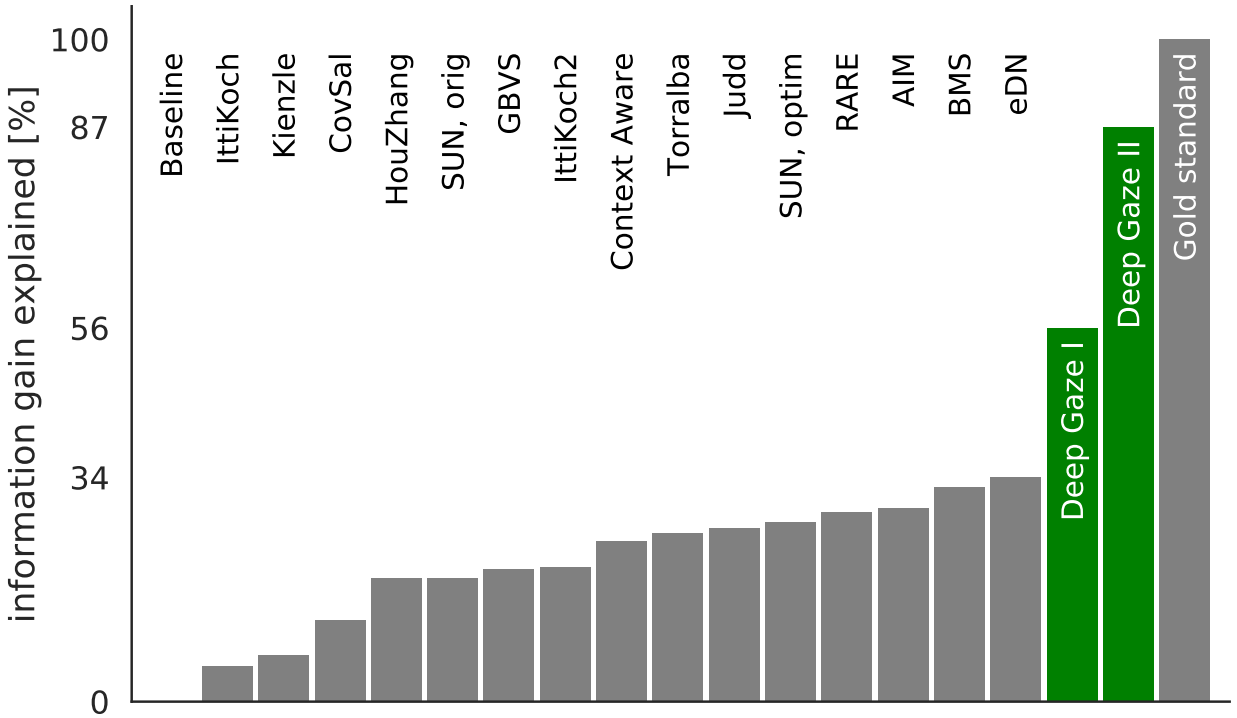
Deep Gaze II: Model architecture



DeepGaze I: ICLR Workshop 2015

DeepGaze II: arXiv 2016
deepgaze.bethgelab.org

Deep Gaze: Performance



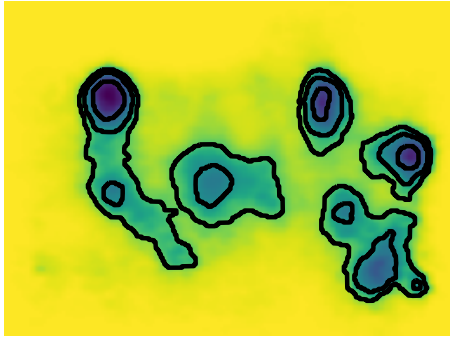
Analysing Probabilistic Models

stimulus with ground truth

Analysing Probabilistic Models

stimulus with ground truth

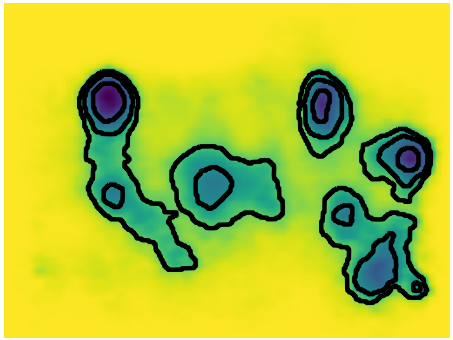
model predictions



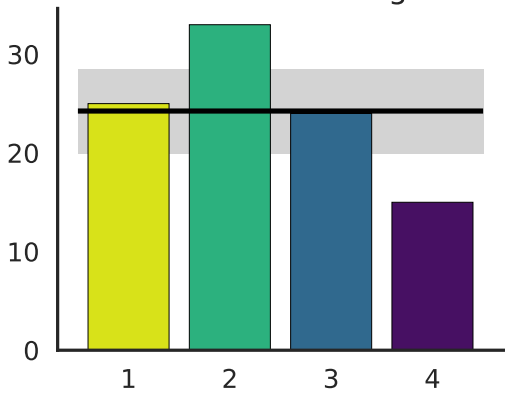
Analysing Probabilistic Models

stimulus with ground truth

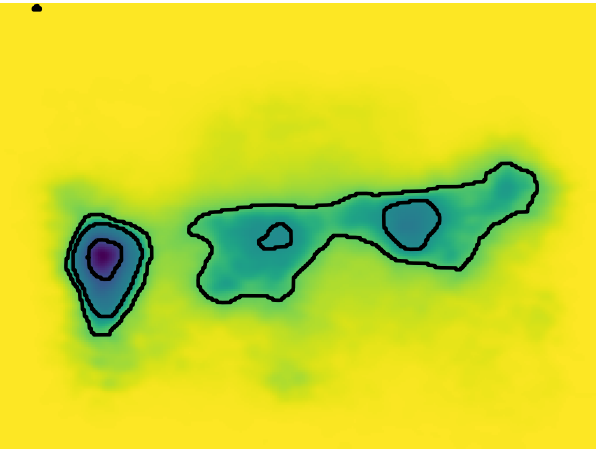
model predictions



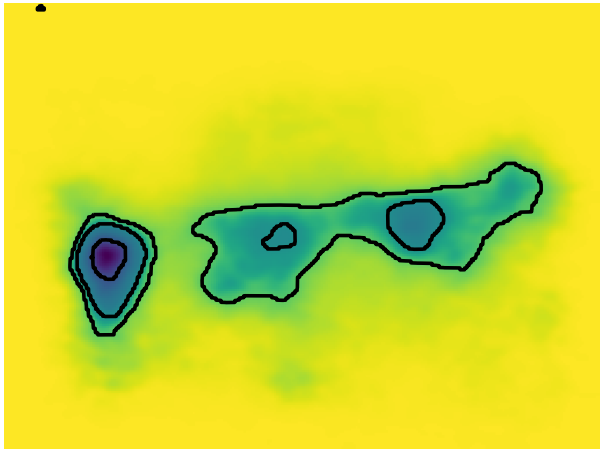
fixation counts for image areas

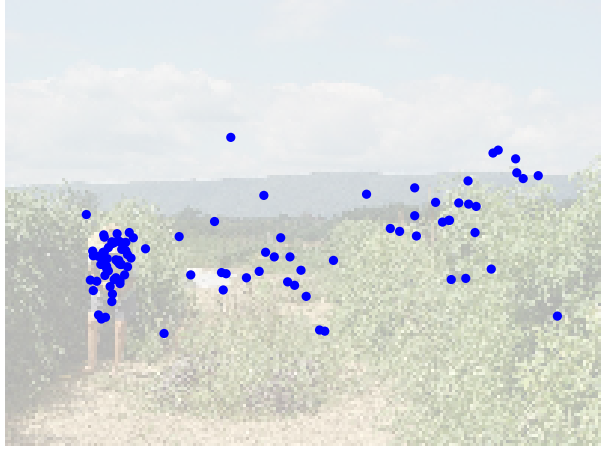


Analysing Probabilistic Models

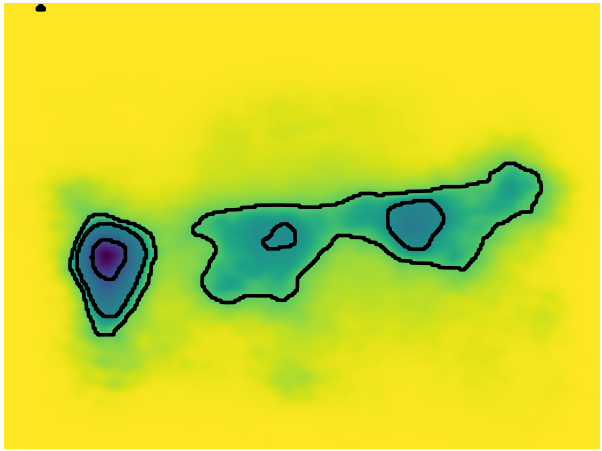


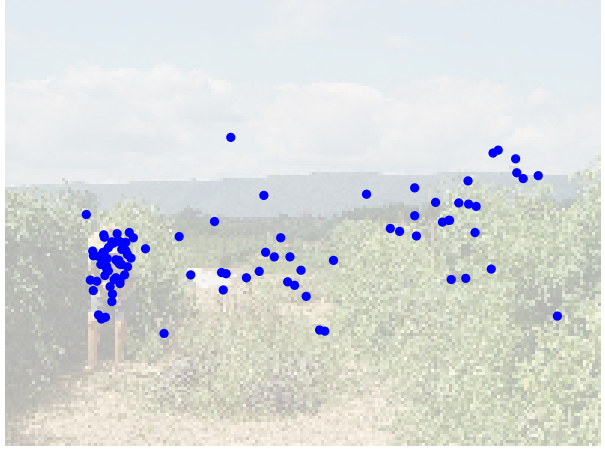
Analysing Probabilistic Models

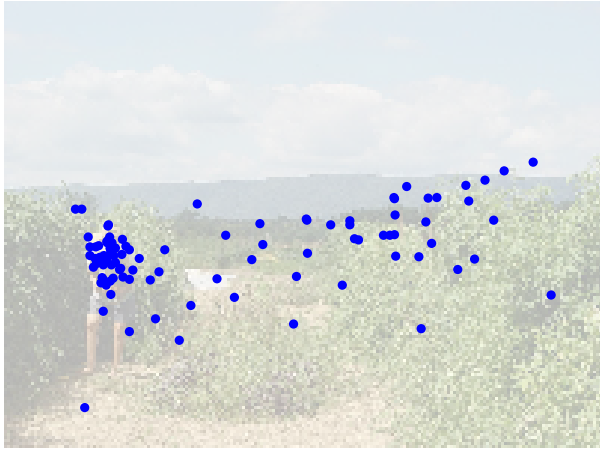




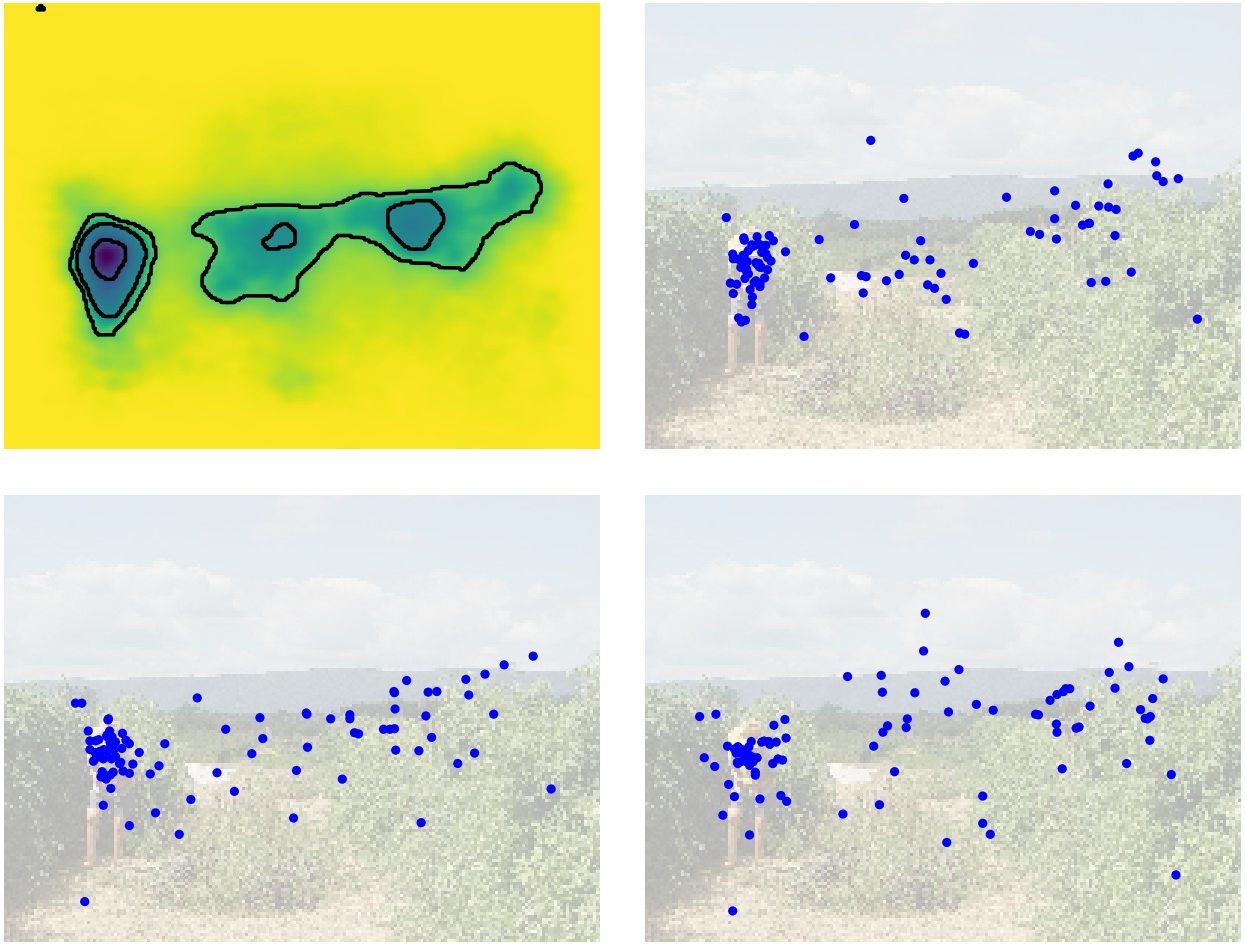
Analysing Probabilistic Models



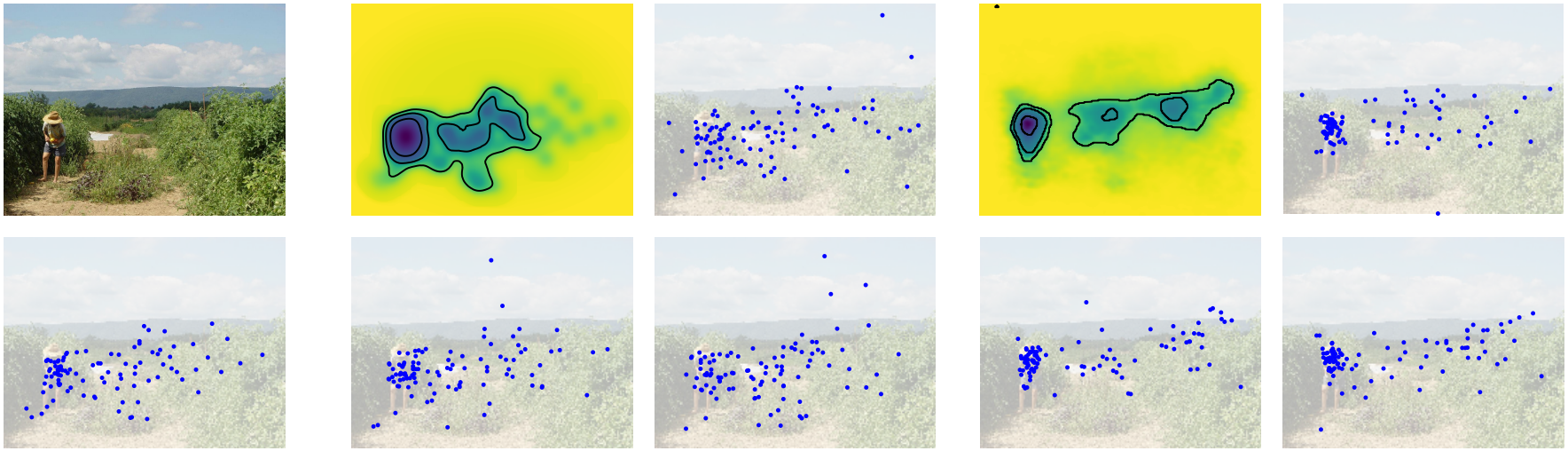




Analysing Probabilistic Models

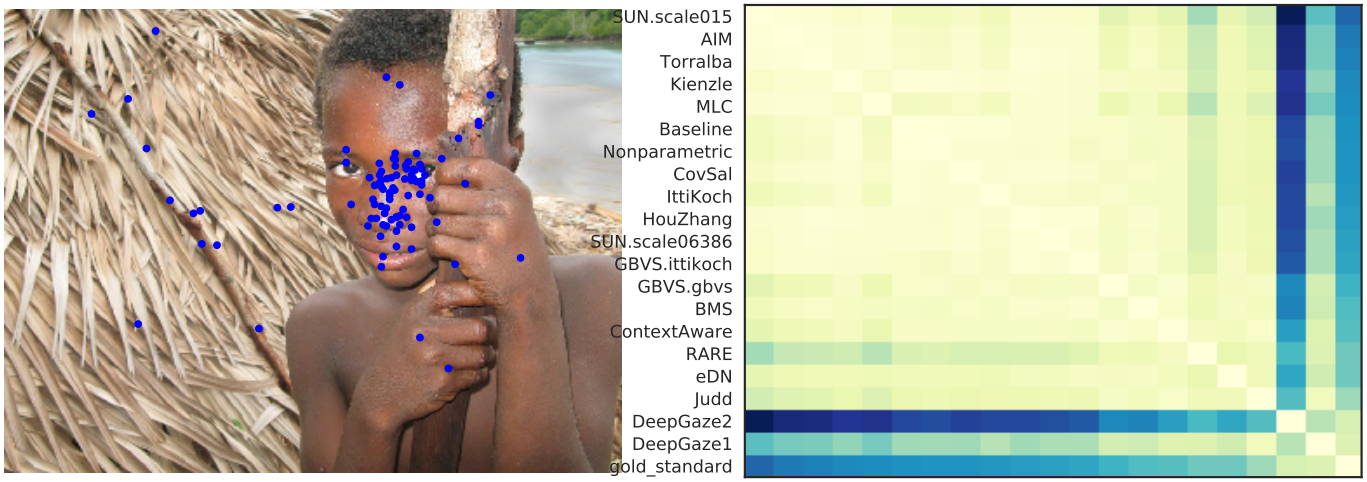


Analysing Probabilistic Models

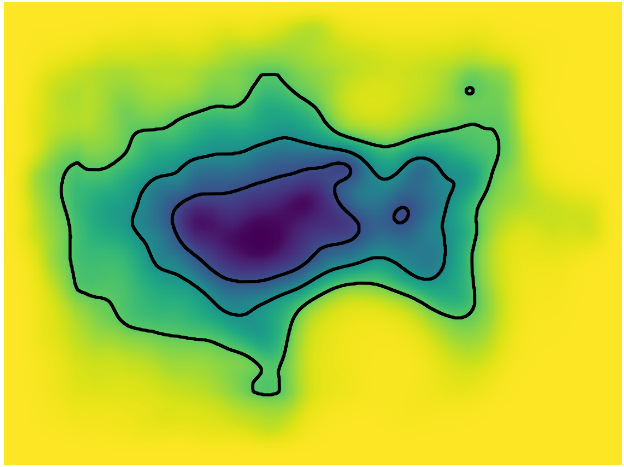


Analysing Probabilistic Models

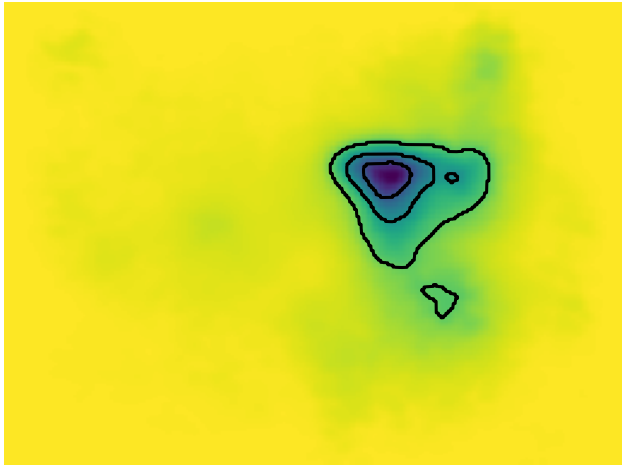
p_DG(left)-p_MLC(left)=-0.5977329390906019



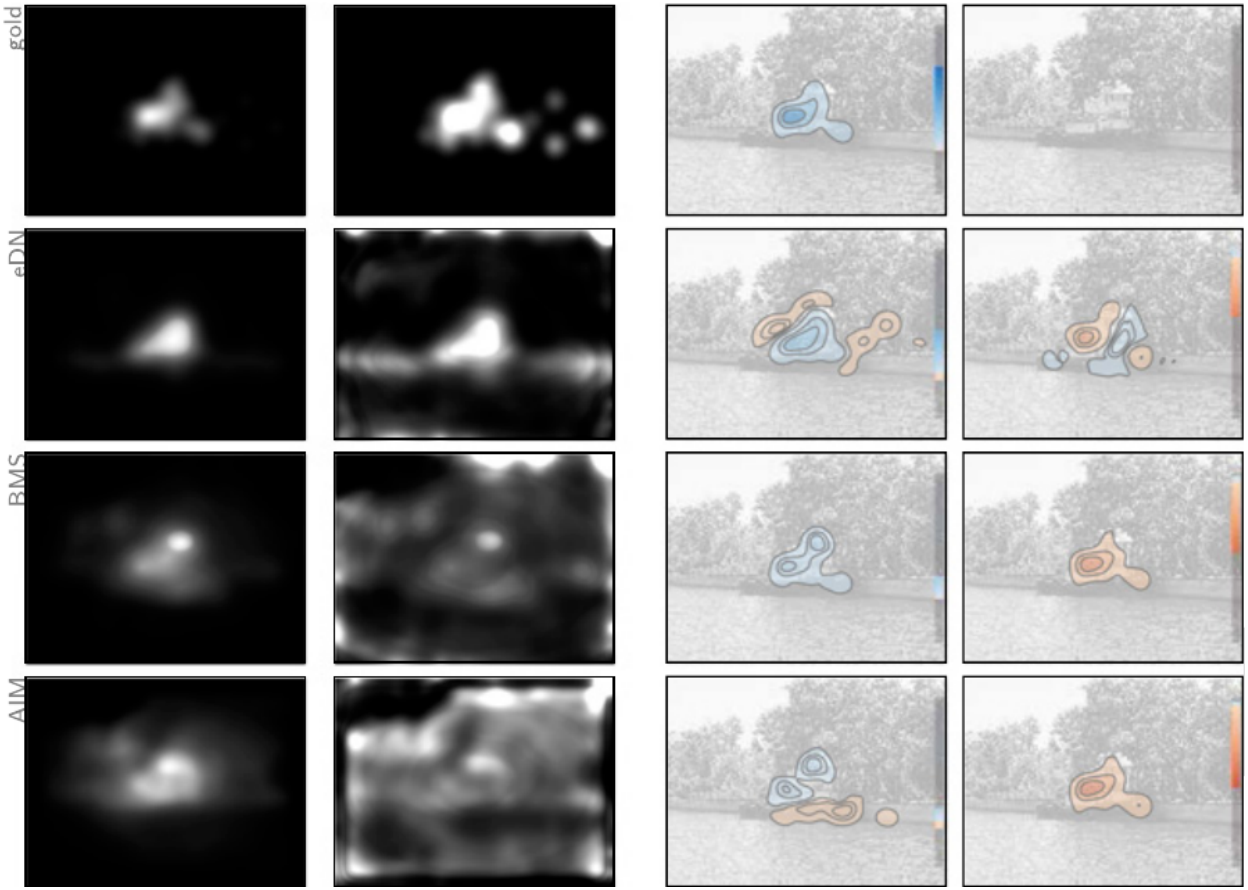
Mean-Luminance-Contrast



DeepGaze2



Analysing Probabilistic Models

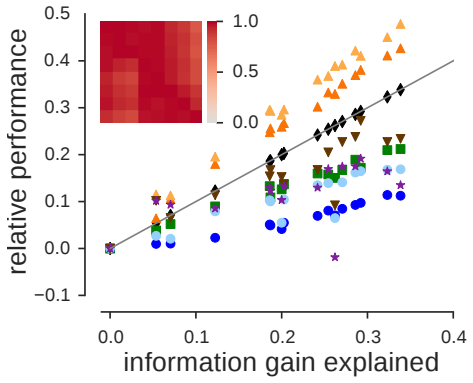


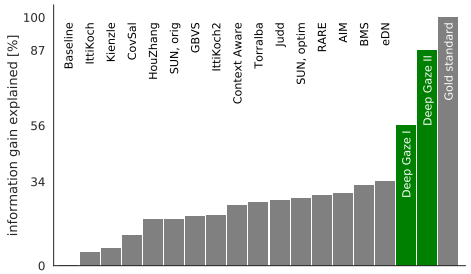
Saliency Benchmarking

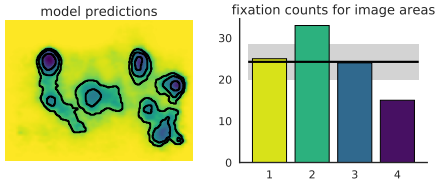
- Define submission format; encourage people to hand in probabilistic models
- Have a principled way to use different saliency maps for different metrics, e.g.
  - with/without centre bias (AUC vs sAUC)
  - match empirical saliency histogramm for CC
- Evaluate information gain / information gain explained
- Publish the centre bias
- What to do about classical models?

Summary

- Phrasing saliency models probabilistically allows to resolve the inconsistency between different metrics, making benchmarking more interpretable (PNAS 2015)
- By using probabilistic modeling and optimizing for information gain, we were able to improve the state-of-the-art in fixation prediction (ICLR Workshop 2015; arXiv 2016)
- Probabilistic modelling gives us new analysis techniques to quantify where and how models fail, and to visualize the limitations of our datasets







Thank you

- Matthias Bethge
- Tom Wallis
- Lucas Theis
- Bethgelab

CIN
Werner Reichardt
Centre for Integrative Neuroscience

bccn tübingen

EBERHARD KARLS UNIVERSITÄT TÜBINGEN