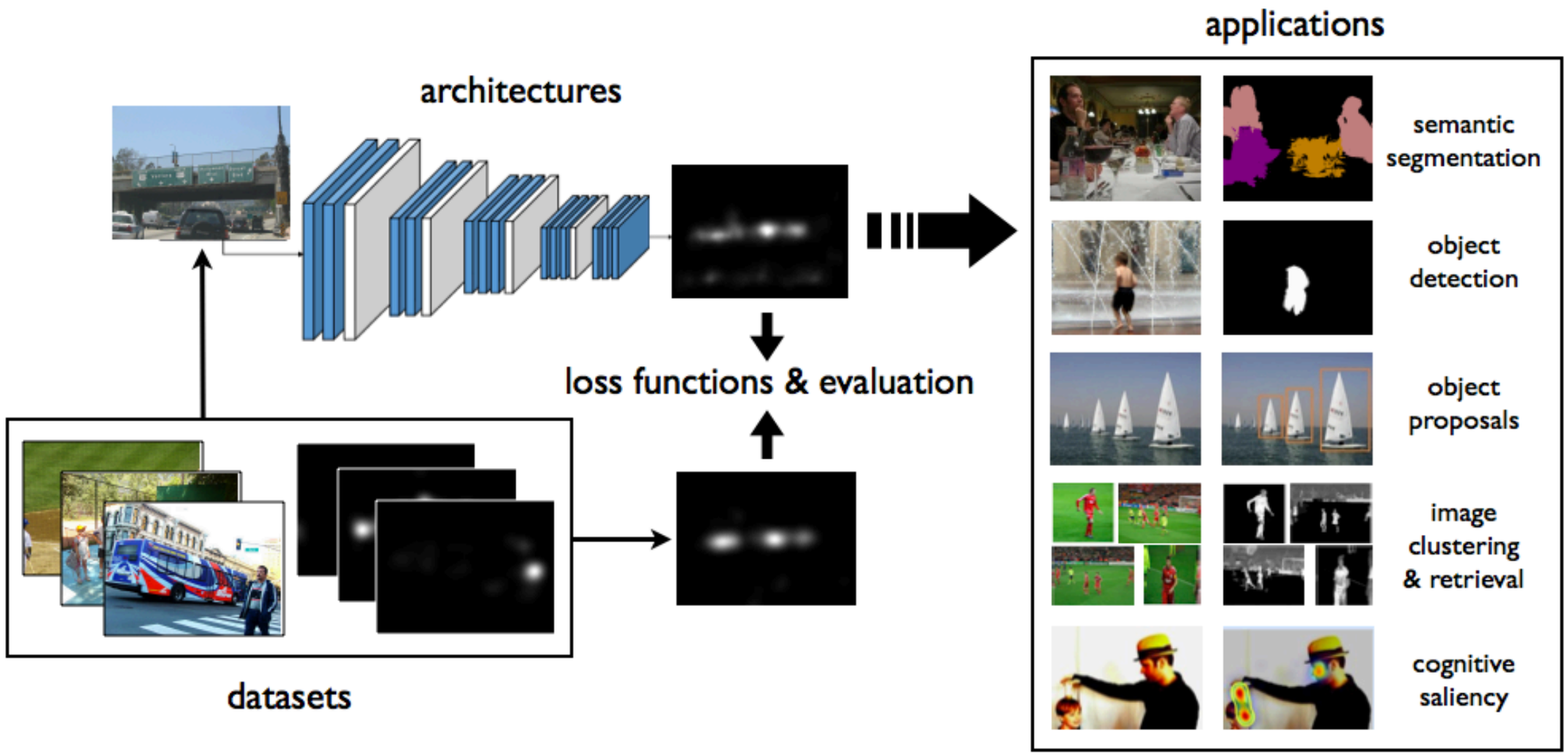


New directions in saliency research:

Developments in architectures, datasets, and evaluation

Tutorial Overview

Zoya Bylinskii & Tilke Judd, ECCV Oct. 8, 2016

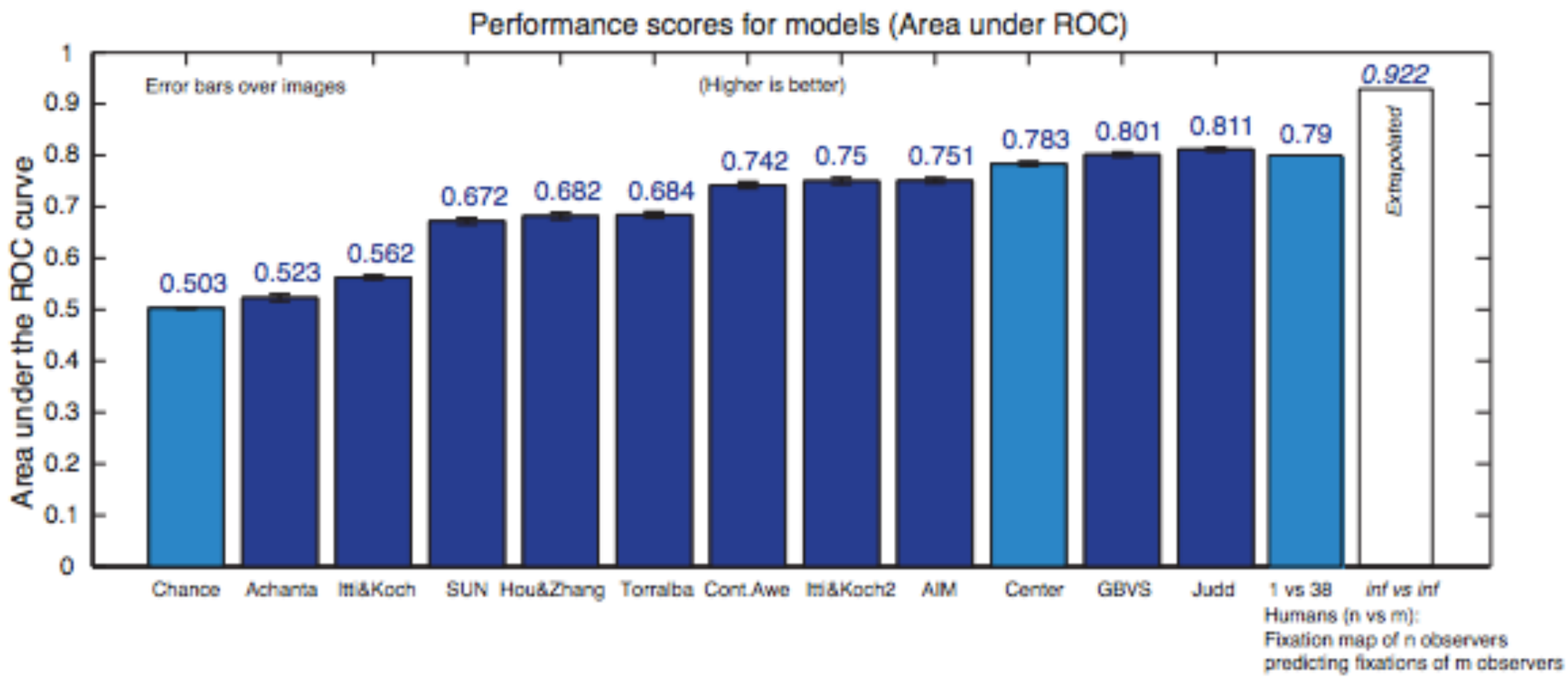
MIT Saliency Benchmark Team

Tutorial organized by:

Ali Borji
Center for Research in Computer Vision, University of Central Florida

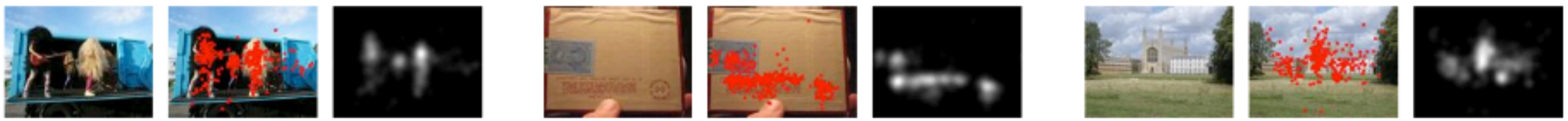
Zoya Bylinskii
Computer Science and Artificial Intelligence Lab, MIT

Tilke Judd
Product Manager, Google
Previously: MIT



Judd et al. "A Benchmark of Computational Models of Saliency to Predict Human Fixations" [MIT Tech Report 2012]

mit saliency benchmark



home | results | datasets | submission | downloads

mit300 | cat2000

mit saliency benchmark results: mit300

images

300 benchmark images (the fixations from 39 viewers per image are not public such that no model can be trained using this data set).

model performances

Model Visualizations

66 models, **5** baselines, **8** metrics, and counting...

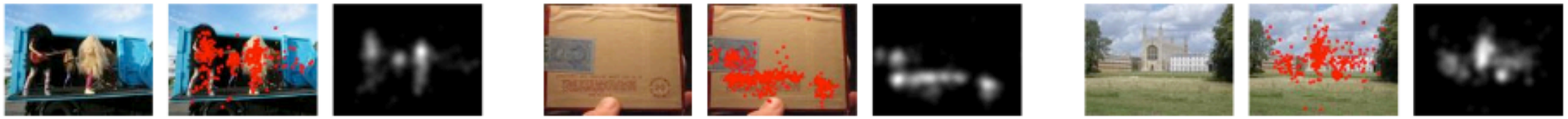
Performance numbers prior to September 25, 2014.

Matlab code for the metrics we use.

Sorted by: AUC-Judd metric

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]		0.88 (0.84)	0.46 (0.43)	3.98 (4.52)	0.86 (0.83)	0.72 (0.77)	0.52 (0.45)	1.29 (1.16)	0.96 (1.04)	first tested: 26/11/2015 last tested: 13/09/2016 maps from authors (model without center bias in parentheses)	
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	
Deep Spatial Contextual Long-term Recurrent Convolutional	Nian Liu et al. [unpublished]		0.87	0.68	2.17	0.79	0.72	0.80	2.35	0.95	first tested: 16/06/2016 last tested: 27/07/2016 maps from authors	

mit saliency benchmark



home | results | datasets | submission | downloads

mit300 | cat2000

mit saliency benchmark results: mit300

images

300 benchmark images (the fixations from 39 viewers per image are not public such that no model can be trained on this data set).

model performances

Model Visualizations

Performance numbers prior to September 25, 2014.

Matlab code for the metrics we use.

Model Name	Published	Code	AUC-Judd [?]	SIM [?]	EMD [?]	AUC-Borji [?]	sAUC [?]	CC [?]	NSS [?]	KL [?]	Date tested [key]	Sample [img]
Bas infinite humans [?]			0.92	1	0	0.88	0.81	1	3.29	0		
DC	Matthias Kümmerer, Lucas Theis, Matthias Bethge. Deep Gaze I: Boosting Saliency Prediction with Feature Maps Trained on ImageNet [arxiv 2014]		0.88 (0.84)	0.46 (0.43)	4.98 (4.52)	0.8					first tested: 26/11/2015 last tested: 13/09/2016 maps from authors (model without center bias in parentheses)	
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	
Deep Spatial Contextual Long-term Recurrent Convolutional	Nian Liu et al. [unpublished]		0.87	0.68	2.17	0.79	0.72	0.80	2.35	0.95	first tested: 16/06/2016 last tested: 27/07/2016 maps from authors	

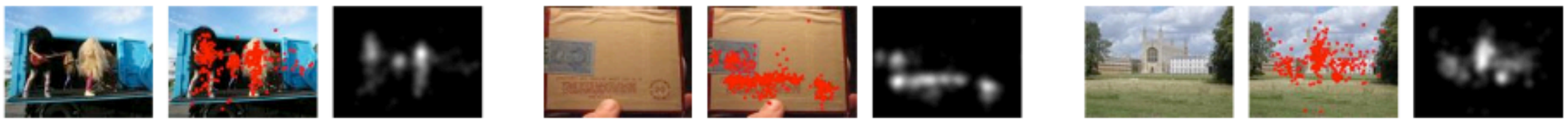
2011

- 10 saliency models
- 3 baslines
- 3 metrics (AUC, SIM, EMD)
- 1 dataset

2016

- 66 saliency models
- 5 baslines
- 8 metrics
- 2 datasets

mit saliency benchmark



home | results | datasets | submission | downloads

mit300 | cat2000

mit saliency benchmark results: mit300

images

300 benchmark images (the fixations from 39 viewers per image are not public such that no model can be trained using this data set).

model performances

Model Visualizations

66 models, **5** baselines, **8** metrics, and counting...

Performance numbers prior to September 25, 2014.

Matlab code for the metrics we use.

Sorted by: AUC-Judd metric

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]		0.88 (0.84)	0.46 (0.43)	3.98 (4.52)	0.86 (0.83)	0.72 (0.77)	0.52 (0.45)	1.29 (1.16)	0.96 (1.04)	first tested: 26/11/2015 last tested: 13/09/2016 maps from authors (model without center bias in parentheses)	
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	
Deep Spatial Contextual Long-term Recurrent Convolutional	Nian Liu et al. [unpublished]		0.87	0.68	2.17	0.79	0.72	0.80	2.35	0.95	first tested: 16/06/2016 last tested: 27/07/2016 maps from authors	

tracking progress with the **MIT Saliency Benchmark**

home | results | datasets | submission | downloads

mit300 | cat2000

mit saliency benchmark results: mit300

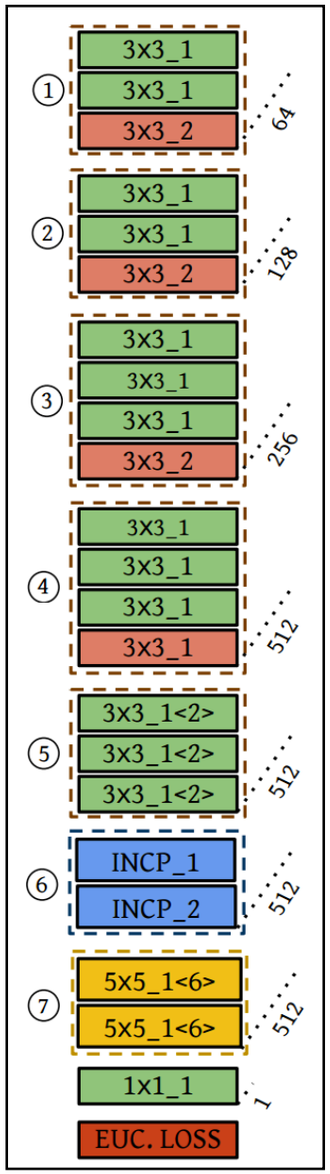
- 66 models evaluated on MIT benchmark (as of Sept 30, 2016)
- top **8/10** are neural networks
- neural networks make up **25%** of all submissions (17 models since 2014)

large boosts in performance in recent years driven by neural network architectures

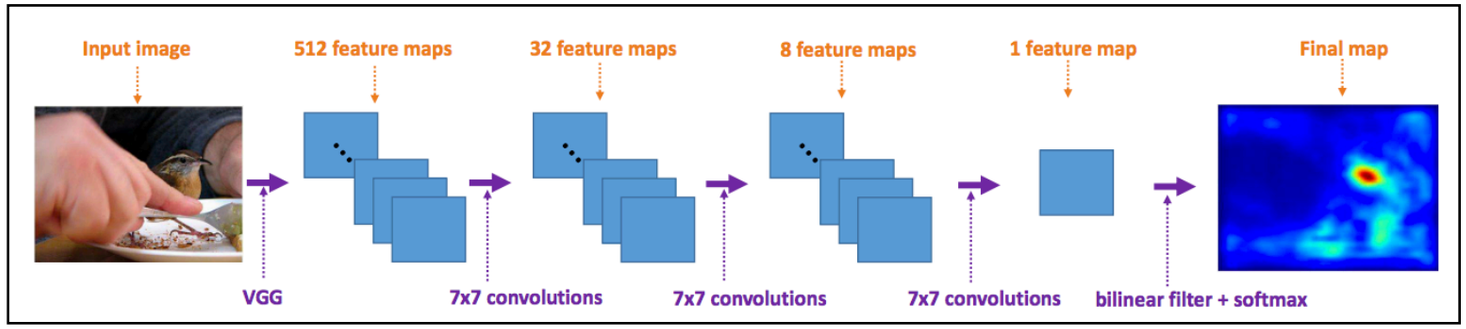
Model Name	Published	Code	AUC-Judd [?]	AUC-Borji [?]	[?]	[?]	[?]	[?]	CC [?]	NSS [?]	KL [?]	Date tested [key]	Sample [img]
[?]										3.29	0		
Deep Gaze 2	Matthias Kümmerer, Lucas Theis, Matthias Bethge. Deep Gaze I: Boosting Saliency Prediction with Feature Maps Trained on ImageNet [arxiv 2014]		0.88 (0.84)	0.46 (0.43)	3.98 (4.52)	0.86 (0.83)	0.72 (0.77)	0.52 (0.45)	1.29 (1.16)	0.96 (1.04)	first tested: 26/11/2015 last tested: 13/09/2016 maps from authors (model without center bias in parentheses)	
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	
Deep Spatial Contextual Long-term Recurrent Convolutional	Nian Liu et al. [unpublished]		0.87	0.68	2.17	0.79	0.72	0.80	2.35	0.95	first tested: 16/06/2016 last tested: 27/07/2016 maps from authors	

What are the common architectures used for neural network models of saliency?

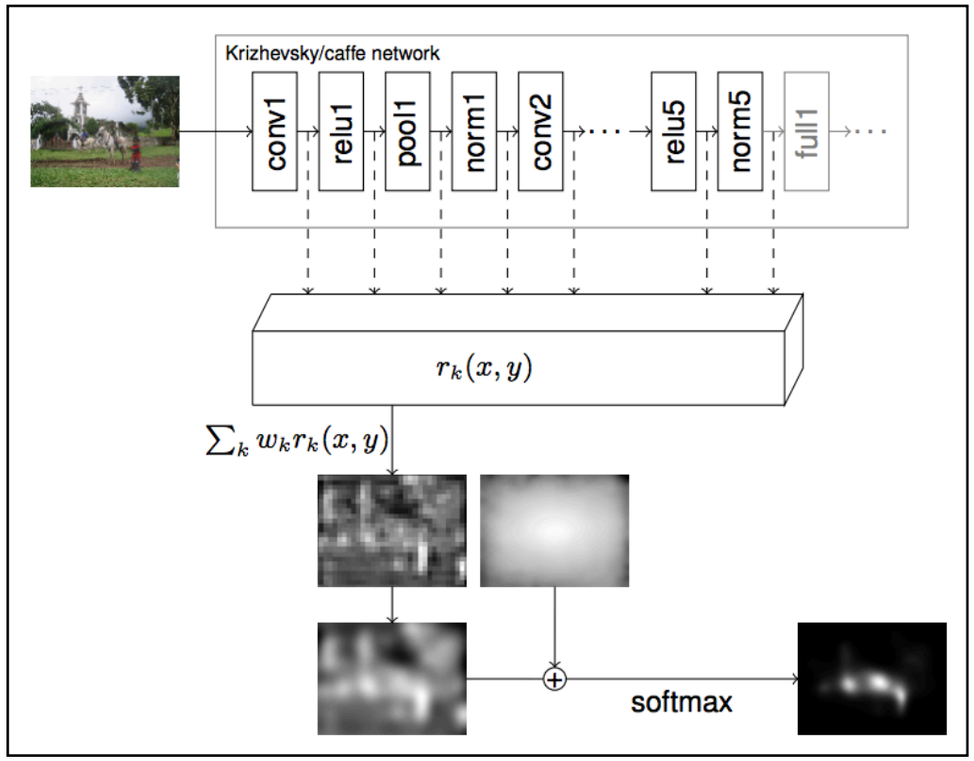
How are these models trained?



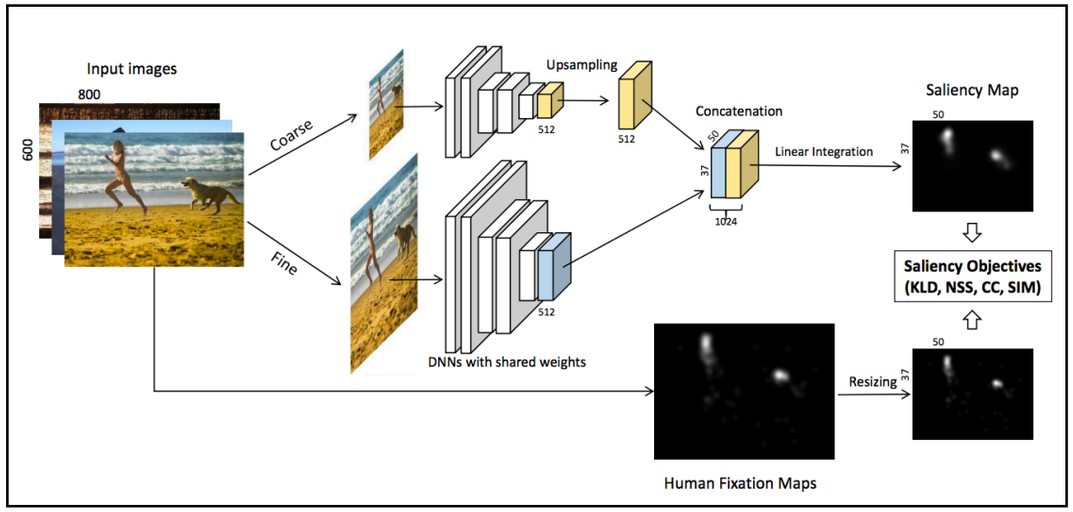
DeepFix - Kruthiventi et al. [ArXiv 2015]



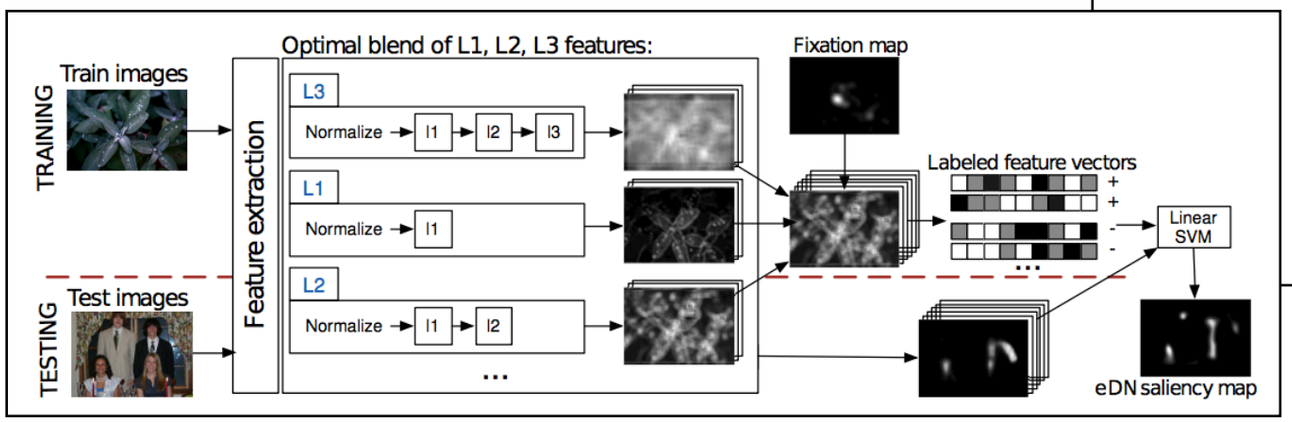
PDP - Jetley et al. [CVPR 2016]



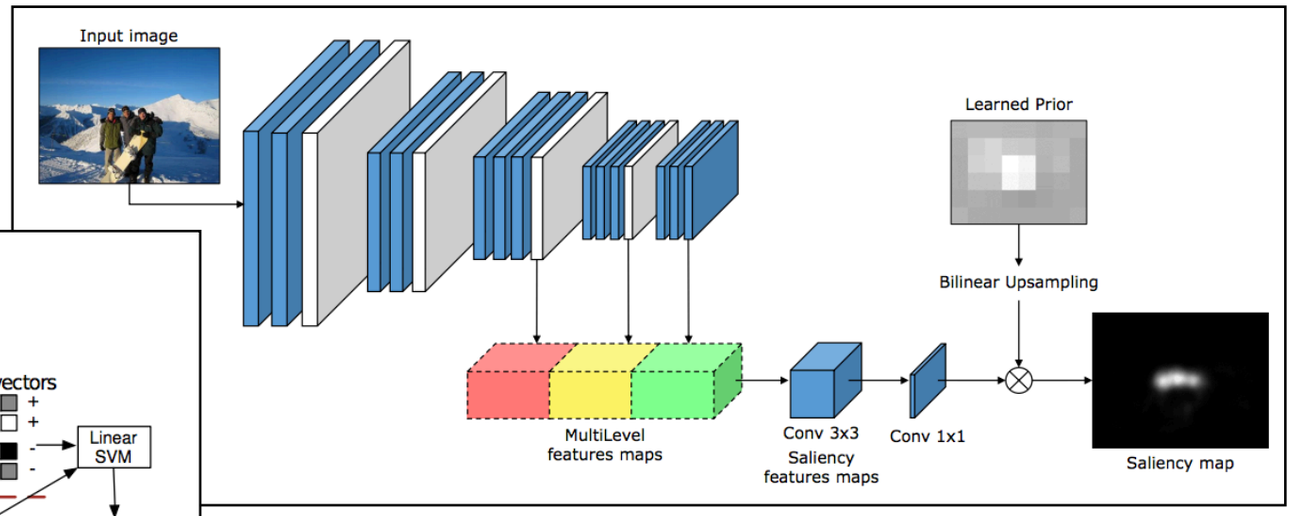
DeepGaze - Kümmerer et al. [ICLR 2015 workshop]



SALICON - Huang et al. [ICCV 2015]



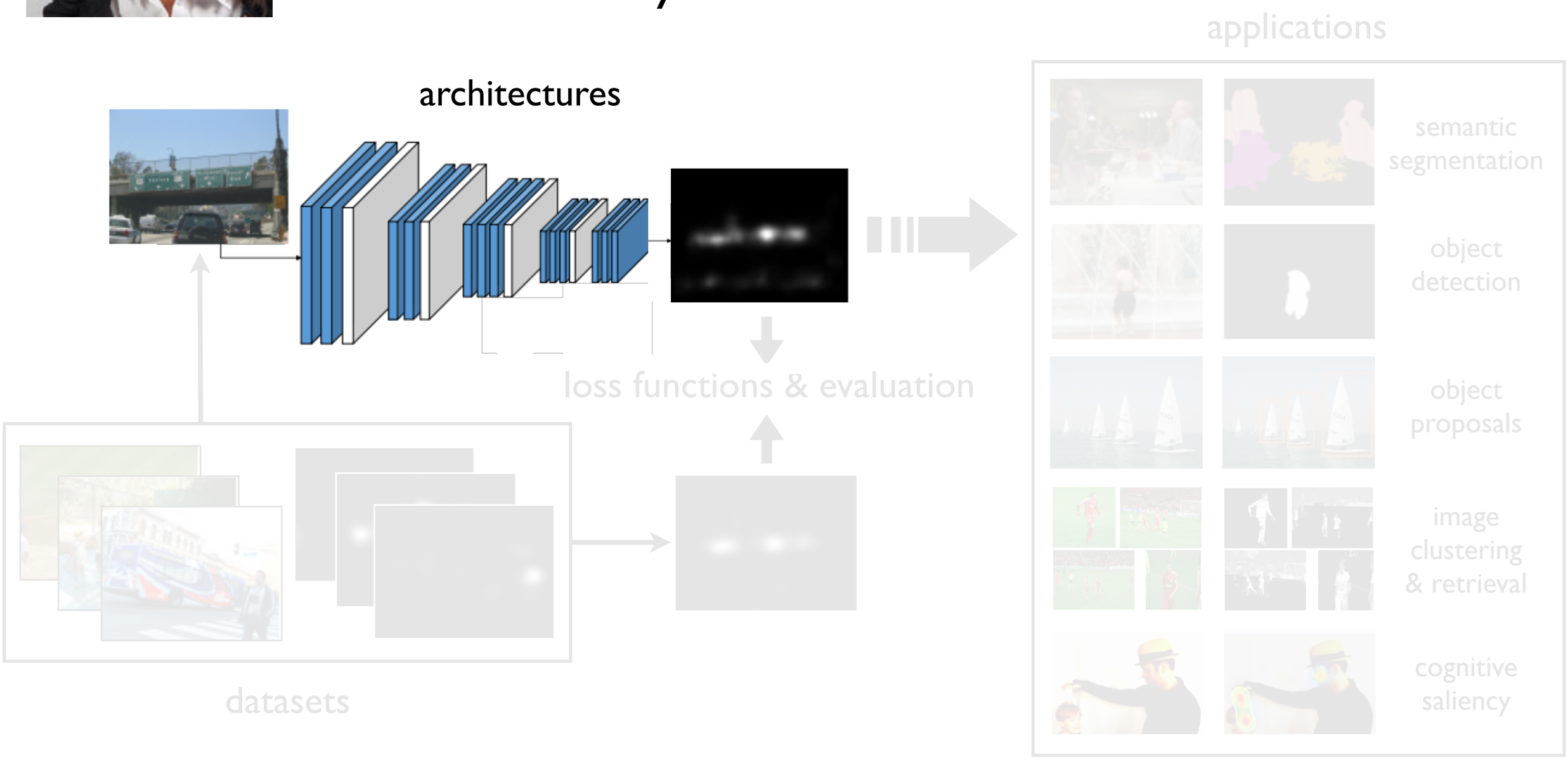
eDN - Vig et al. [CVPR 2014]



ML-Net - Cornia et al. [ICPR 2016]

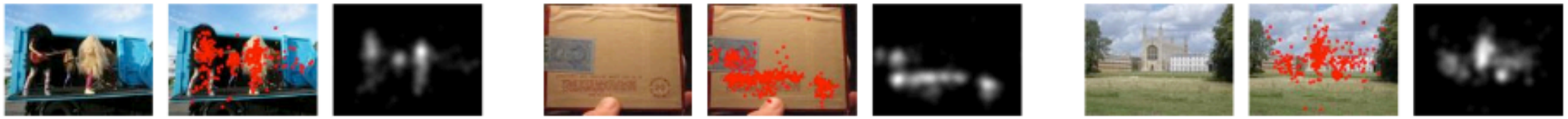
Tutorial overview

"Deep networks for saliency map prediction"
Naila Murray



tracking progress with the **MIT Saliency Benchmark**

tracking progress with the **MIT Saliency Benchmark**



home | results | datasets | submission | downloads

mit300 | cat2000

mit saliency benchmark results: mit300

images

300 benchmark images (the fixations from 39 viewers per image are not public such that no model can be trained using this data set).

model performances

Model Visualizations

66 models, 5 baselines, 8 metrics, and [illegible]...

Performance numbers prior to September 25, 2014.

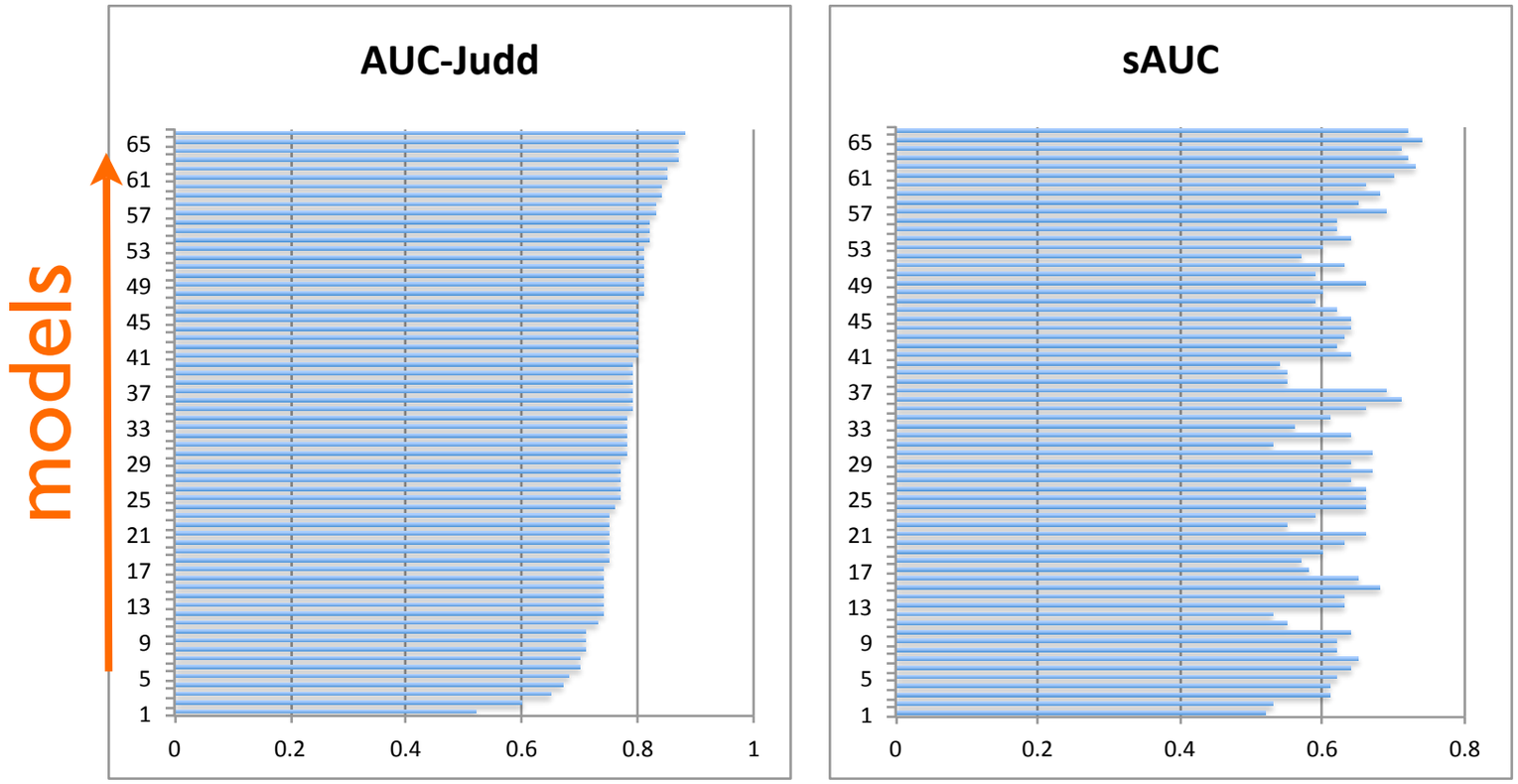
Matlab code for the metrics we use.

Sorted by: AUC-Judd metric

- 66 models evaluated on MIT benchmark
- using 8 evaluation metrics
- on 2 datasets (MIT300, CAT2000)

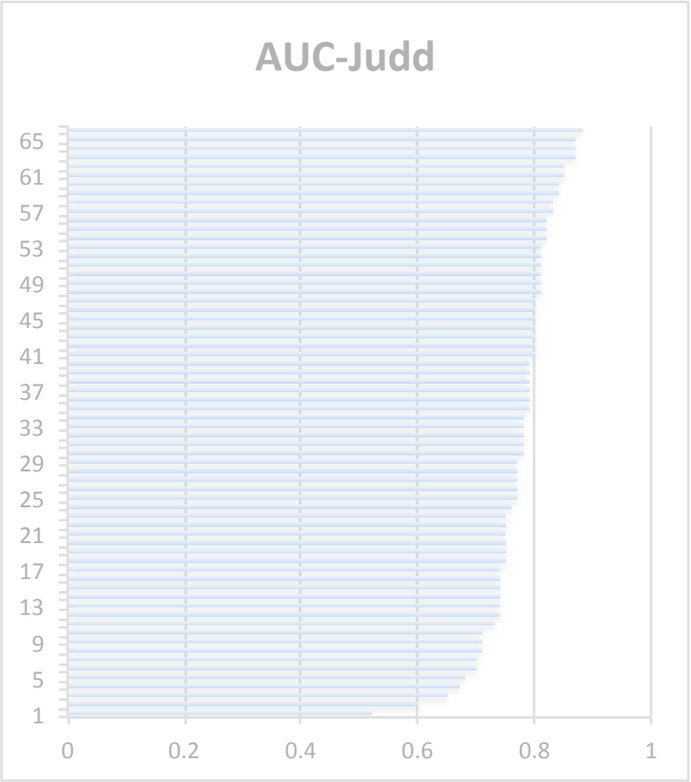
Model Name	Published	Code	AUC-Judd [?]	SIM [?]	EMD [?]	AUC-Borji [?]	sAUC [?]	CC [?]	NSS [?]	KL [?]	Date tested [key]	Sample [img]
Baseline: infinite humans [?]			0.92	1	0	0.88	0.81	1	3.29	0		
Deep Gaze 2	Matthias Kümmerer, Lucas Theis, Matthias Bethge. Deep Gaze I: Boosting Saliency Prediction with Feature Maps Trained on ImageNet [arxiv 2014]		0.88 (0.84)	0.46 (0.43)	3.98 (4.52)	0.86 (0.83)	0.72 (0.77)	0.52 (0.45)	1.29 (1.16)	0.96 (1.04)	first tested: 26/11/2015 last tested: 13/09/2016 maps from authors (model without center bias in parentheses)	
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	
Deep Spatial Contextual Long-term Recurrent Convolutional	Nian Liu et al. [unpublished]		0.87	0.68	2.17	0.79	0.72	0.80	2.35	0.95	first tested: 16/06/2016 last tested: 27/07/2016 maps from authors	

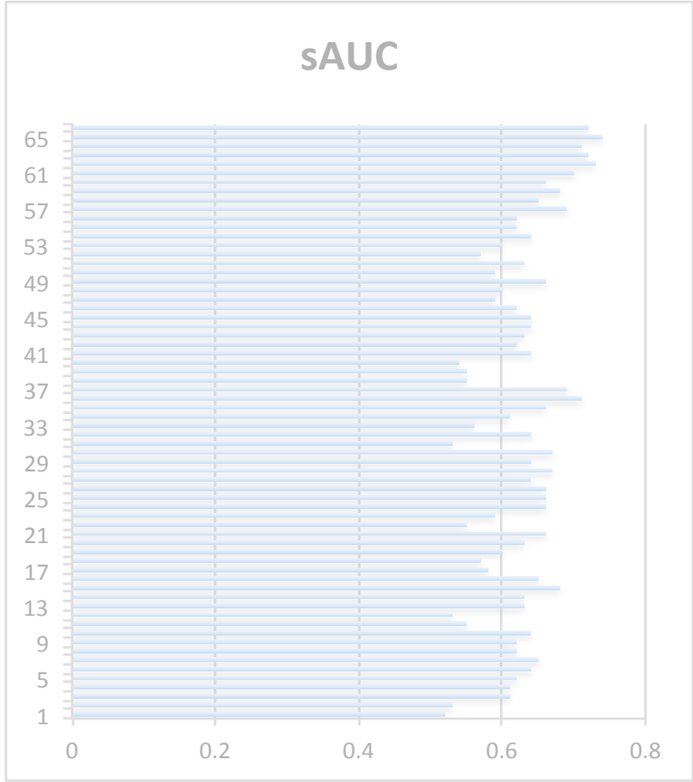
tracking progress with the **MIT Saliency Benchmark**



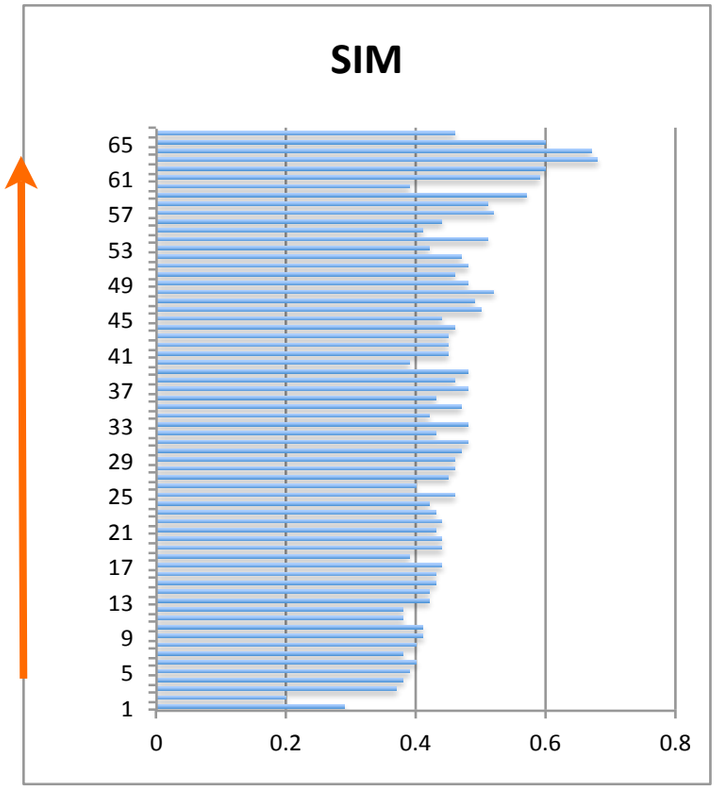
gradual improvement
over time
<u>but</u> saturating

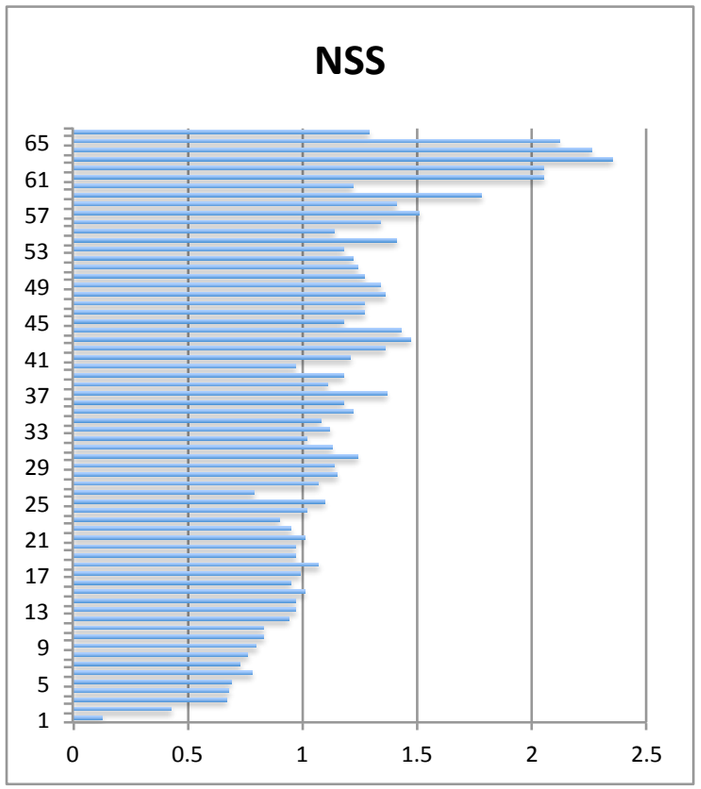
tracking progress with the **MIT Saliency Benchmark**

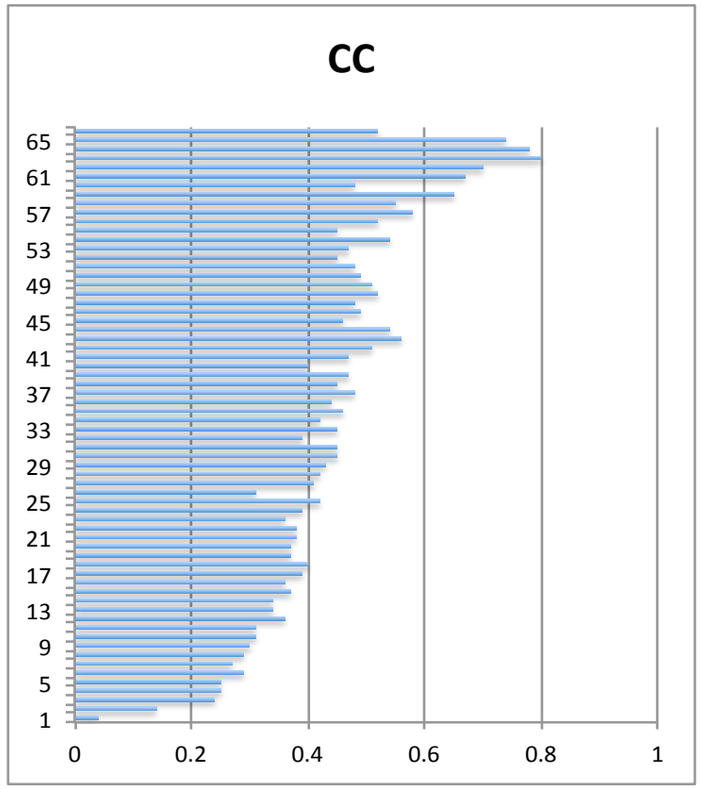




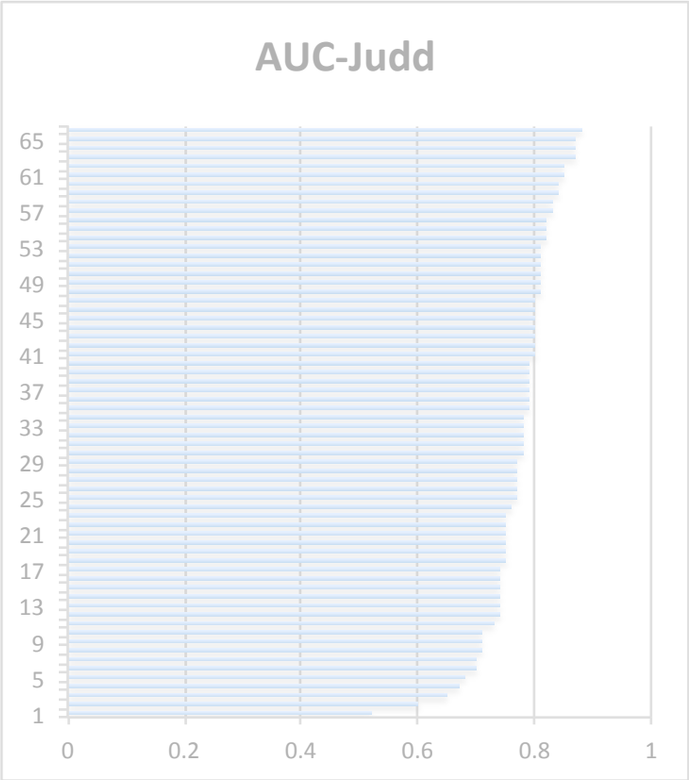
models

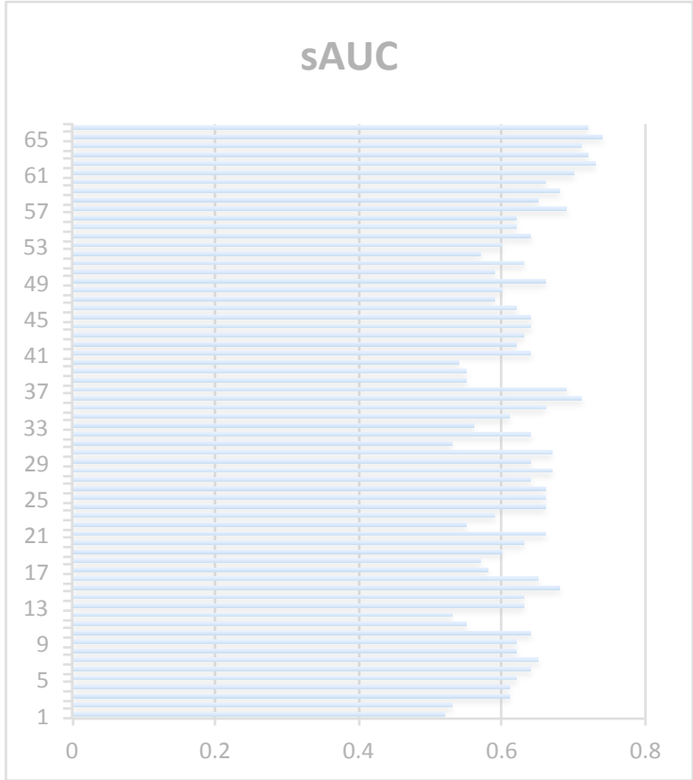






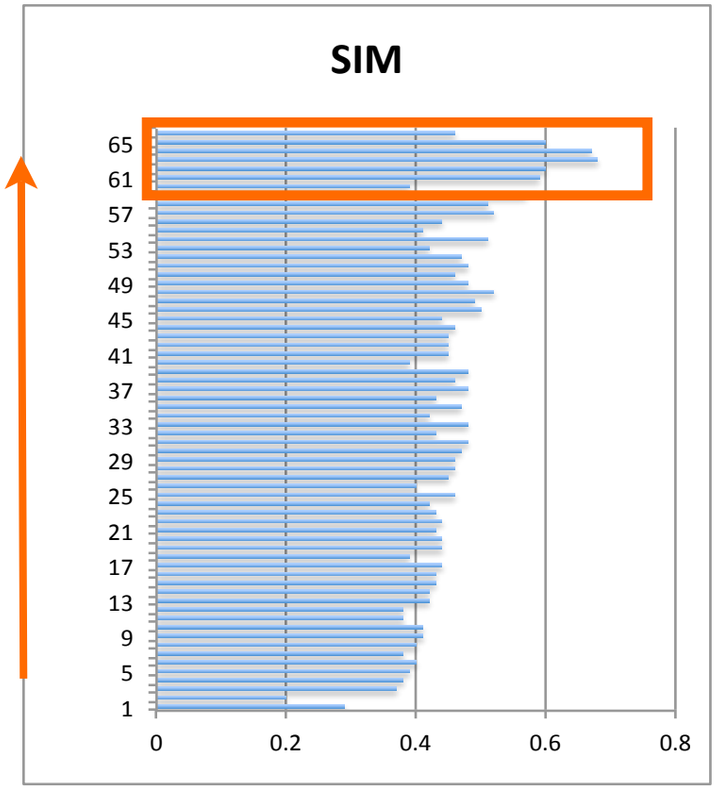
tracking progress with the **MIT Saliency Benchmark**

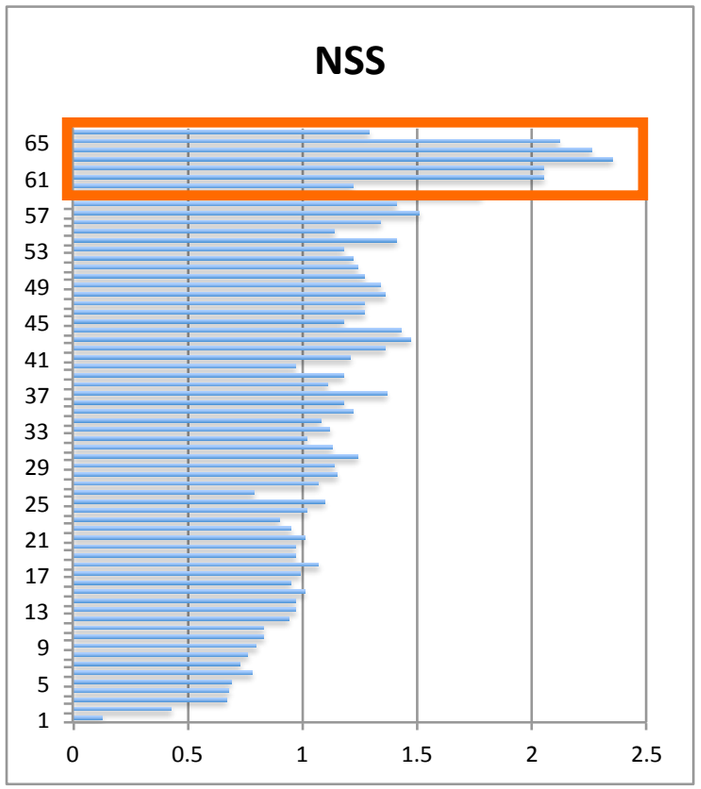


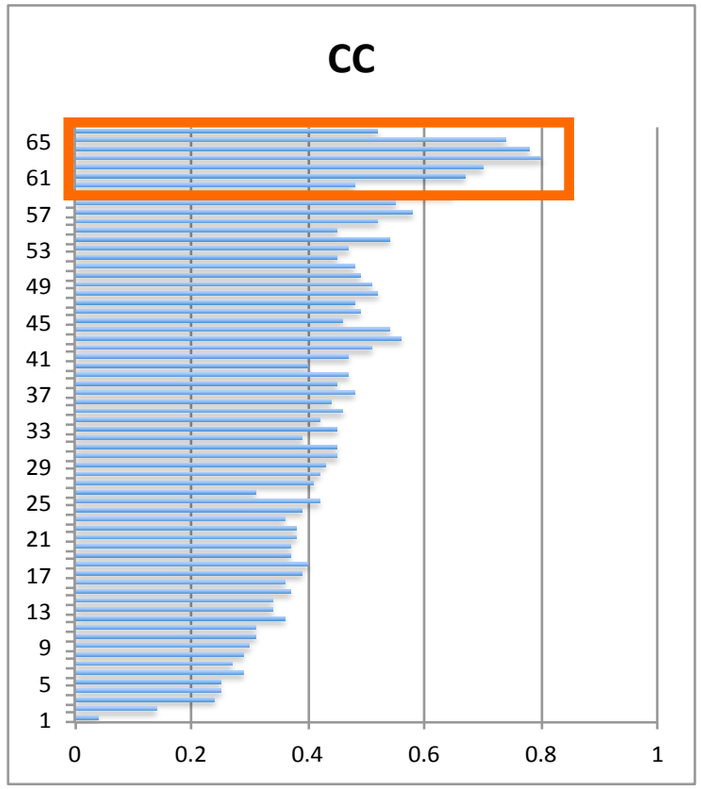


large boosts in performance in recent years driven by neural network architectures

models



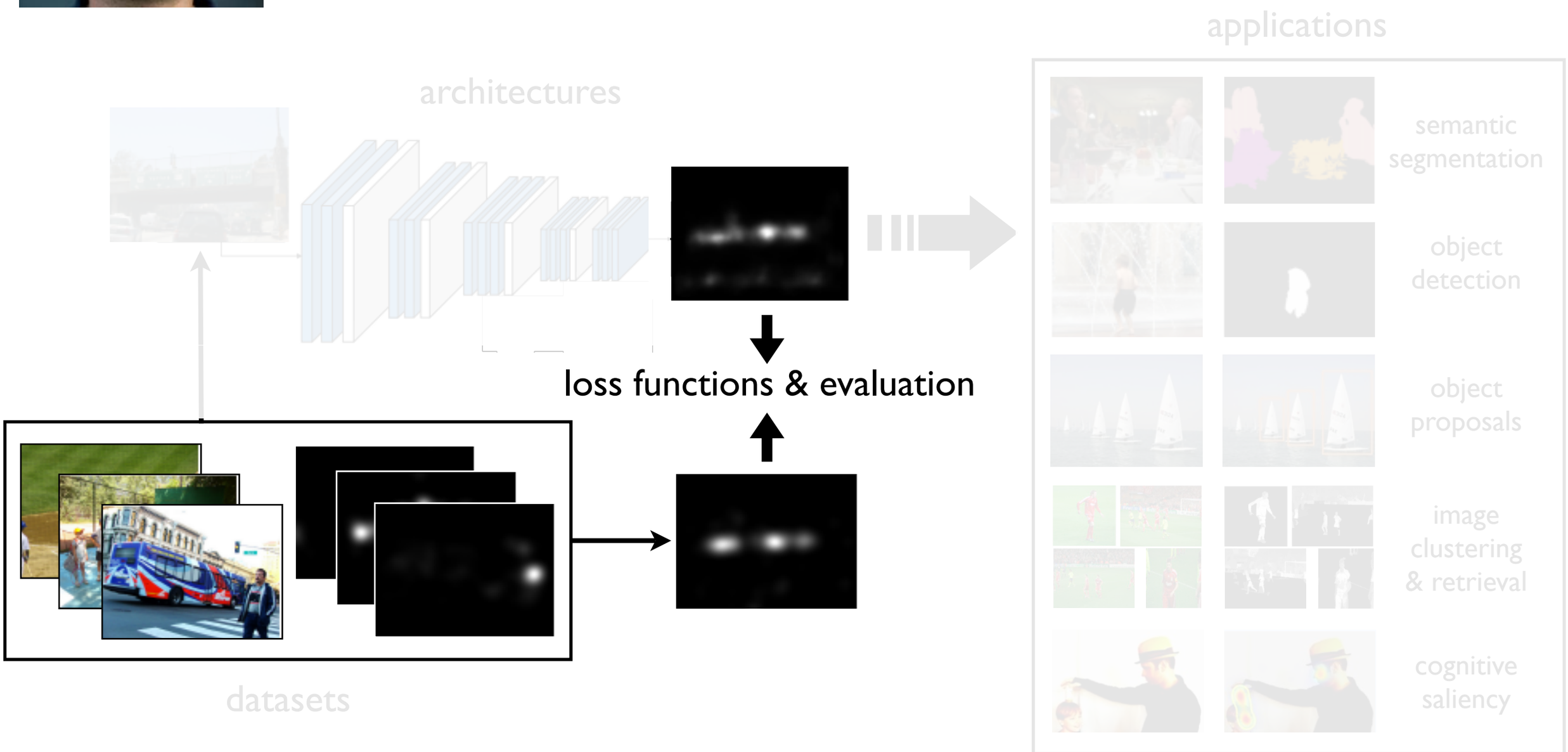




As model performances continue to approach ground truth, how can we have more meaningful evaluation?

Tutorial overview

"Evaluating saliency models in a probabilistic framework"
Matthias Kümmerer

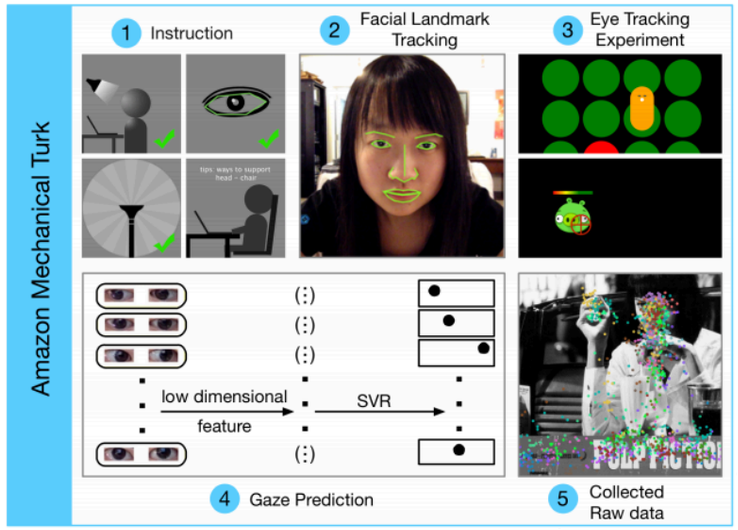


New **data collection** methodologies

How can we collect enough training data
for neural network models of saliency?

New **data collection** methodologies

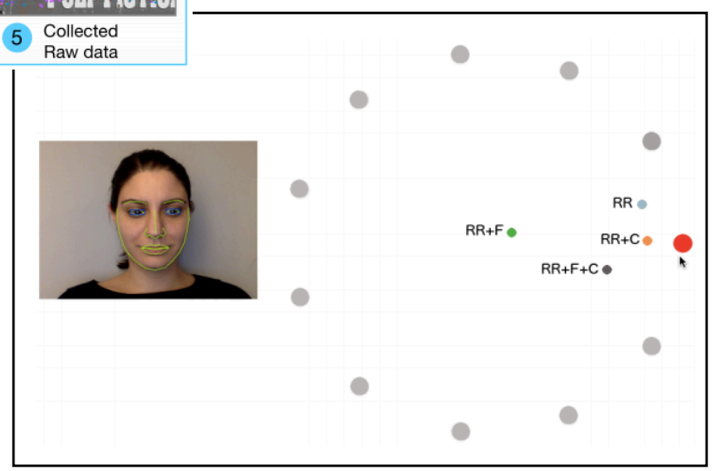
Webcam-based

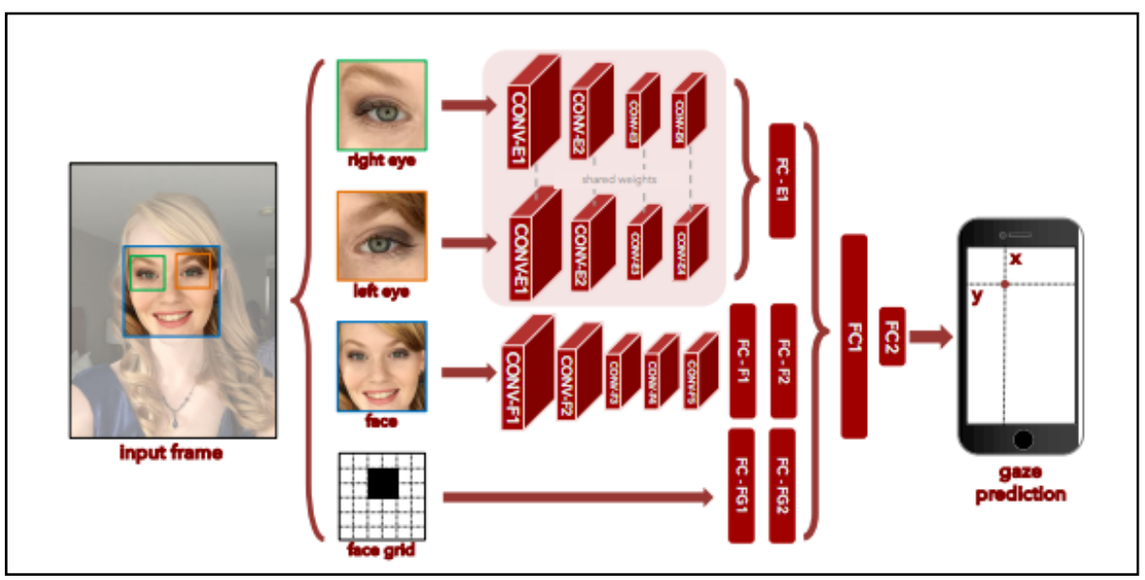


Xu et al. [ArXiv 2015]

iSUN

Papoutsaki et al. [IJCAI 2016]

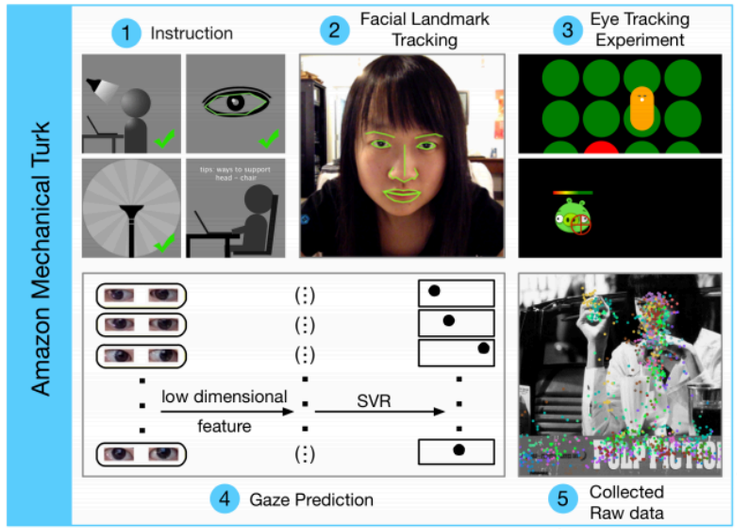




Krafka et al. [CVPR 2016]

New **data collection** methodologies

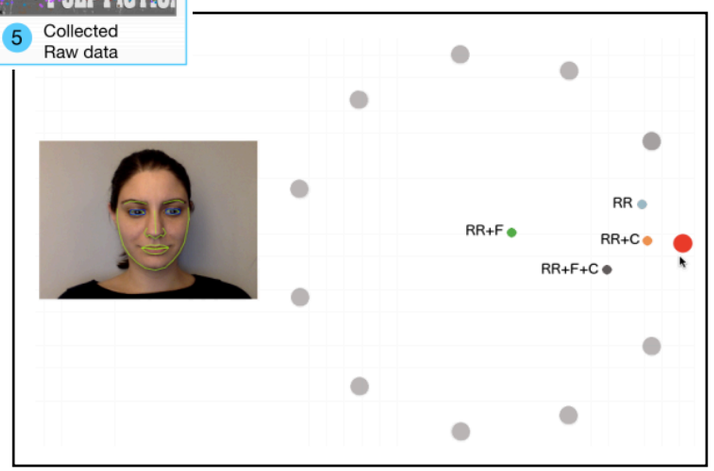
Webcam-based

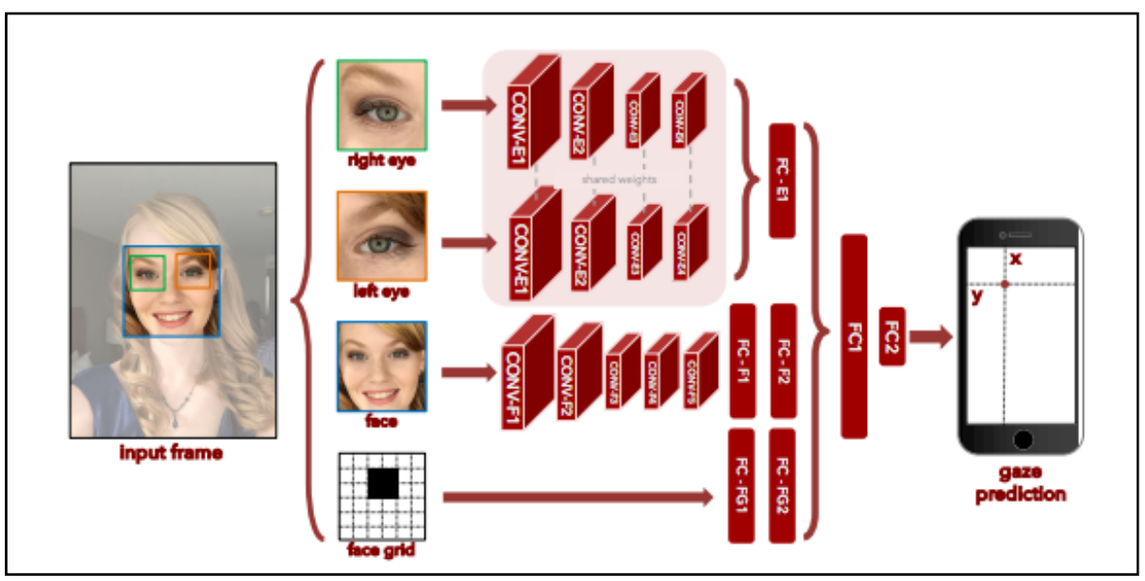


Xu et al. [ArXiv 2015]

iSUN

Papoutsaki et al. [IJCAI 2016]



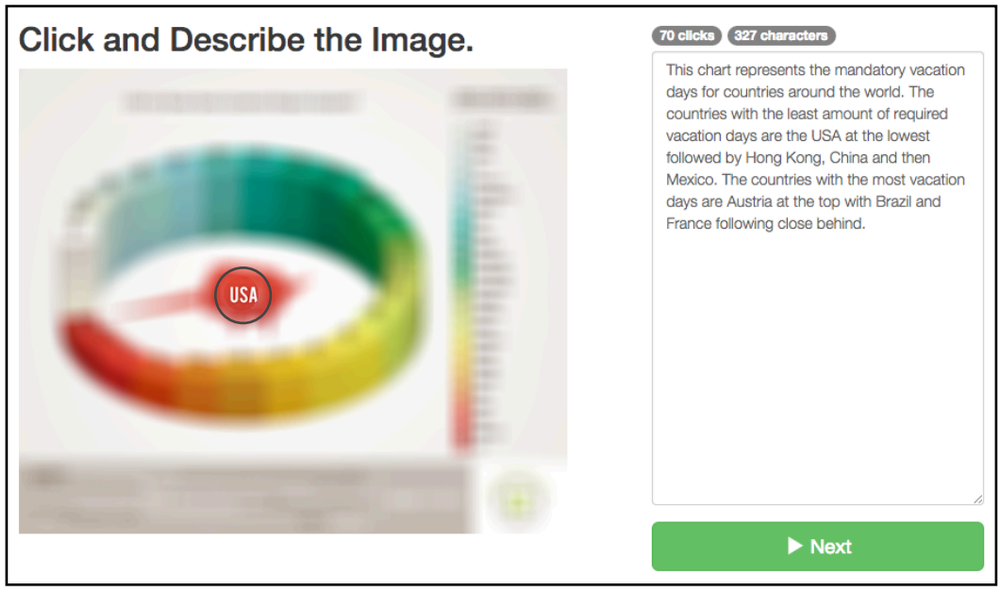


Krafka et al. [CVPR 2016]

Click-based

Jiang et al. [CVPR 2015]

SALICON



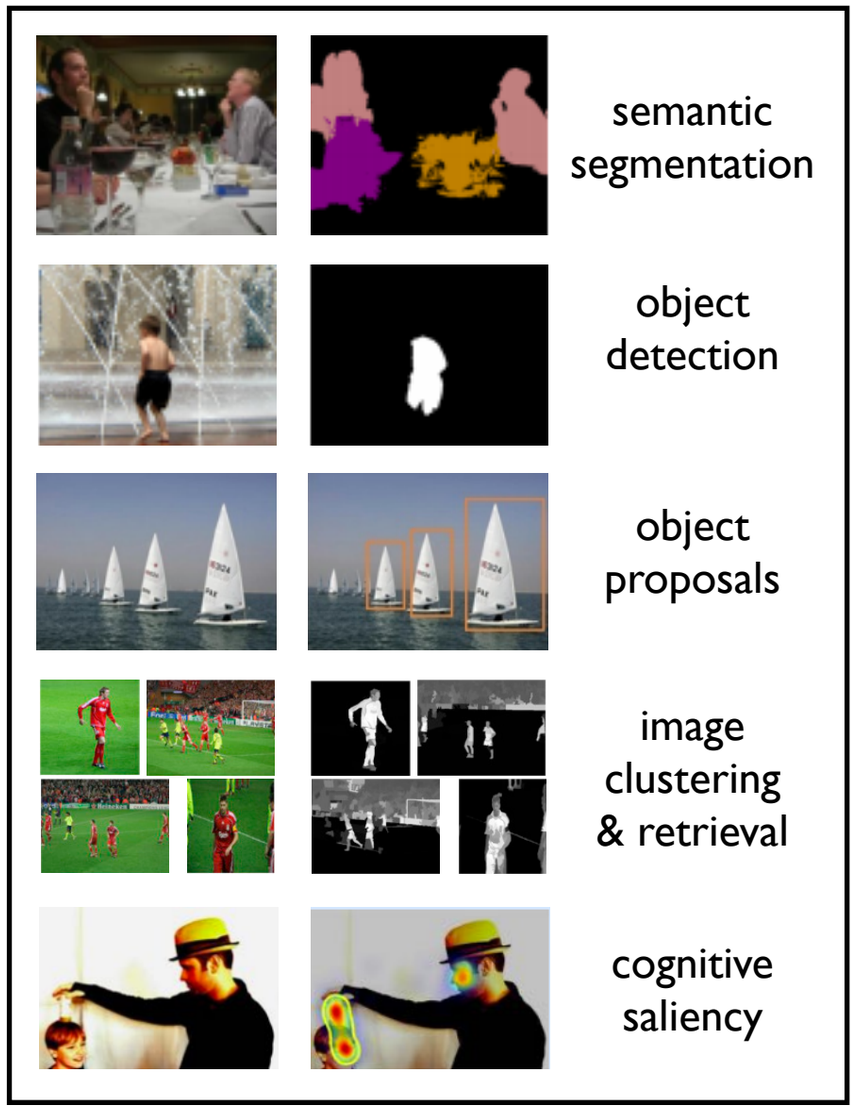
Kim et al. [CHI EA 2015]

What kind of new applications are possible with state-of-the-art saliency architectures and big data?

Tutorial overview

"Saliency for image understanding and manipulation"
Ming-Ming Cheng

applications



Is saliency solved?

Tutorial overview

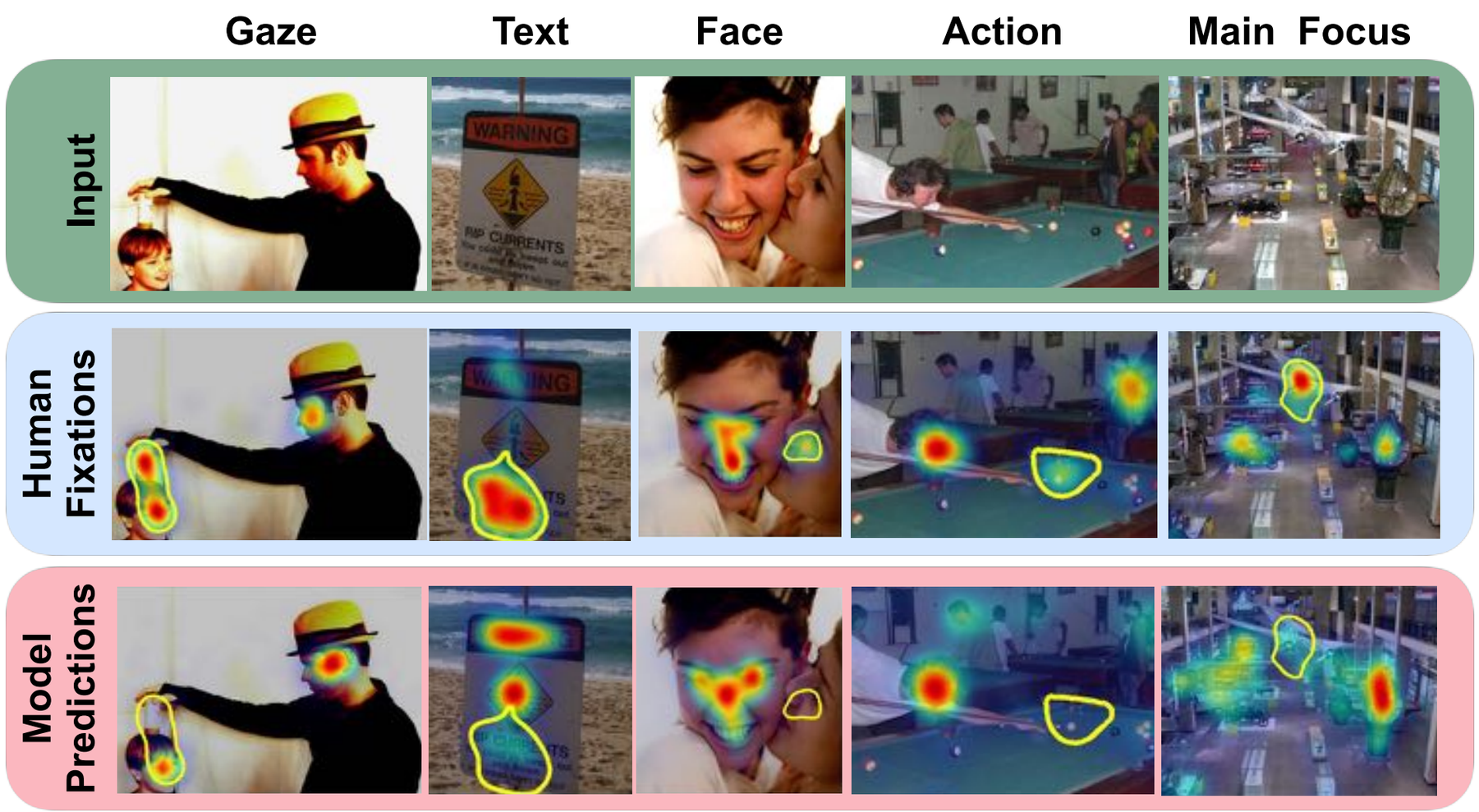
"Towards cognitive saliency"
Zoya Bylinskii

Is sali solved?

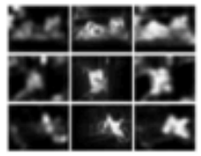
NO!

Tutorial overview

"Towards cognitive saliency"
Zoya Bylinskii



Tutorial Schedule

9:00 - 9:15		Intro & Overview: New directions in saliency Tutorial organizers
9:15 - 9:45		Deep networks for saliency map prediction Naila Murray, Xerox Research Centre Europe
9:45 - 10:15		Evaluating saliency models in a probabilistic framework Matthias Kümmerer, University of Tuebingen
10:15 - 10:45		Saliency for image understanding and manipulation Ming-Ming Cheng, Nankai University
10:45 - 11:00		Coffee break
11:00 - 11:30		Towards cognitive saliency Zoya Bylinskii, Massachusetts Institute of Technology
11:30 - 12:00		Research Panel