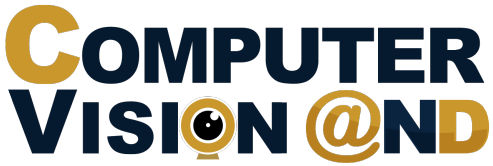
The Limits and Potentials of Deep Learning for Facial Analysis

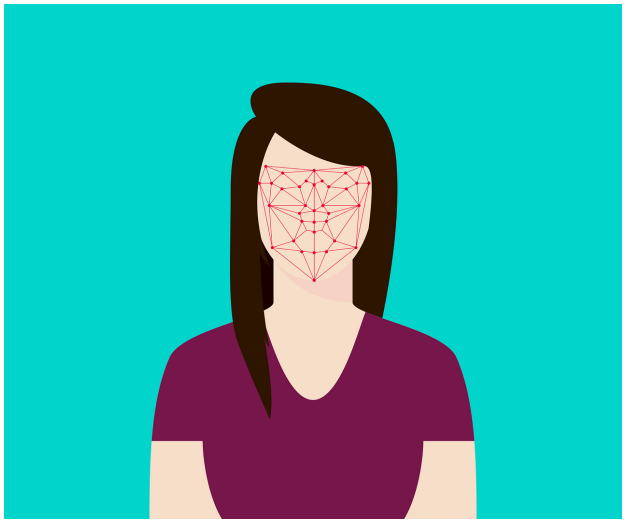
Walter J. Scheirer

Computer Vision Research Laboratory
Department of Computer Science and Engineering



Biometrics isn’t the only field studying faces…

Imagine the following scenario:

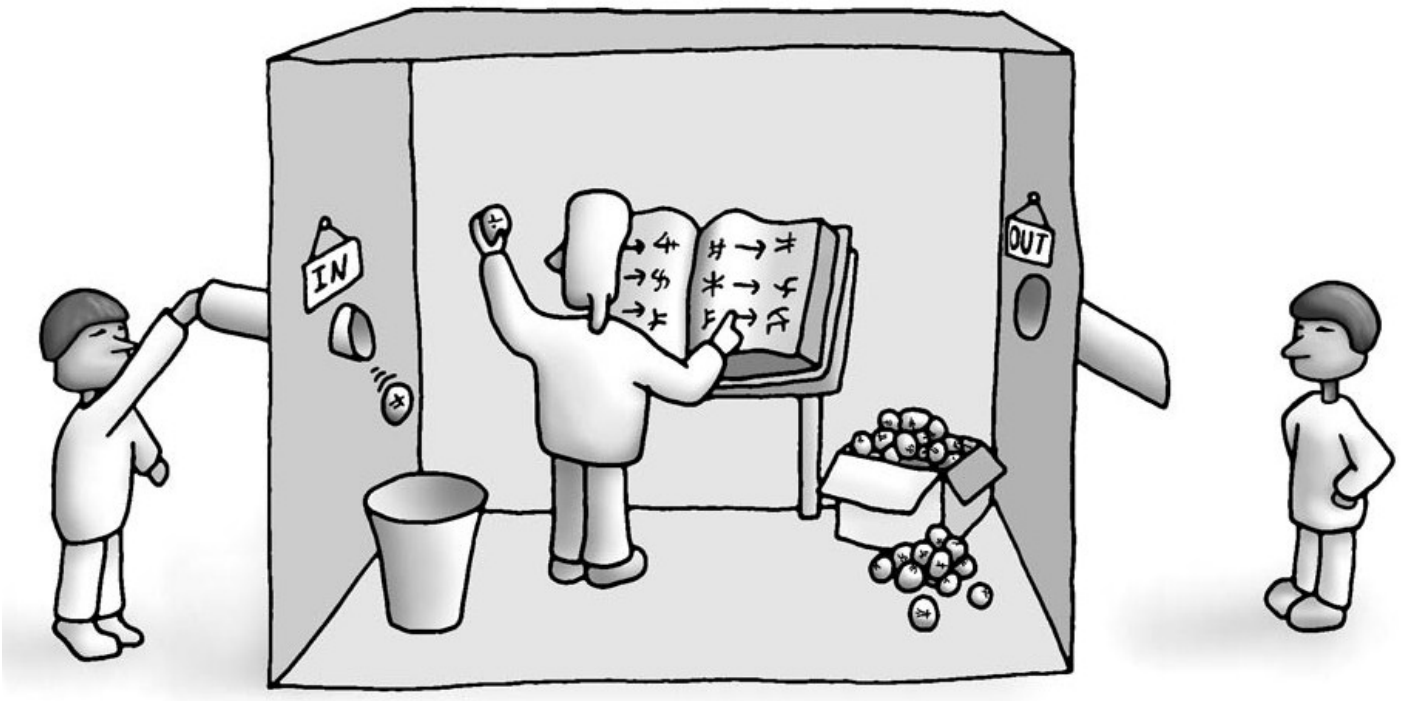


A proprietary face recognition system purportedly solves biometric identification with human-like ability.

By all accounts, the software achieves superior performance on established computer vision benchmarks

How would you go about falsifying the claim of human-like ability?

The Chinese Room

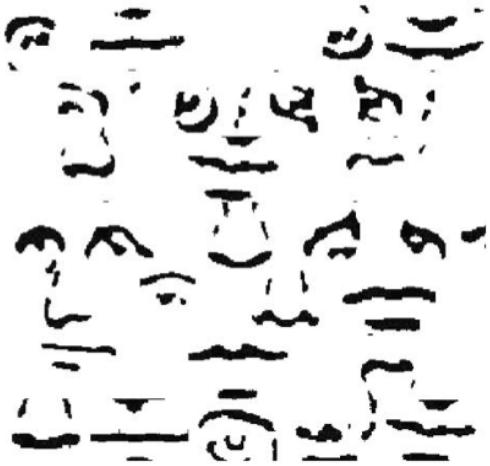


Searle. Behavioral and Brain Sciences 1980

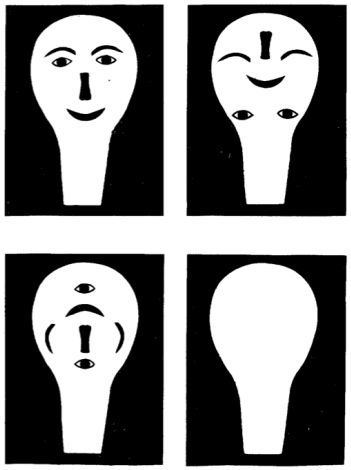
Visual Psychophysics

Probe psychological and perceptual thresholds through controlled manipulation of stimuli.

Careful management of stimulus construction, ordering and presentation allows for precise determination of perceptual thresholds.



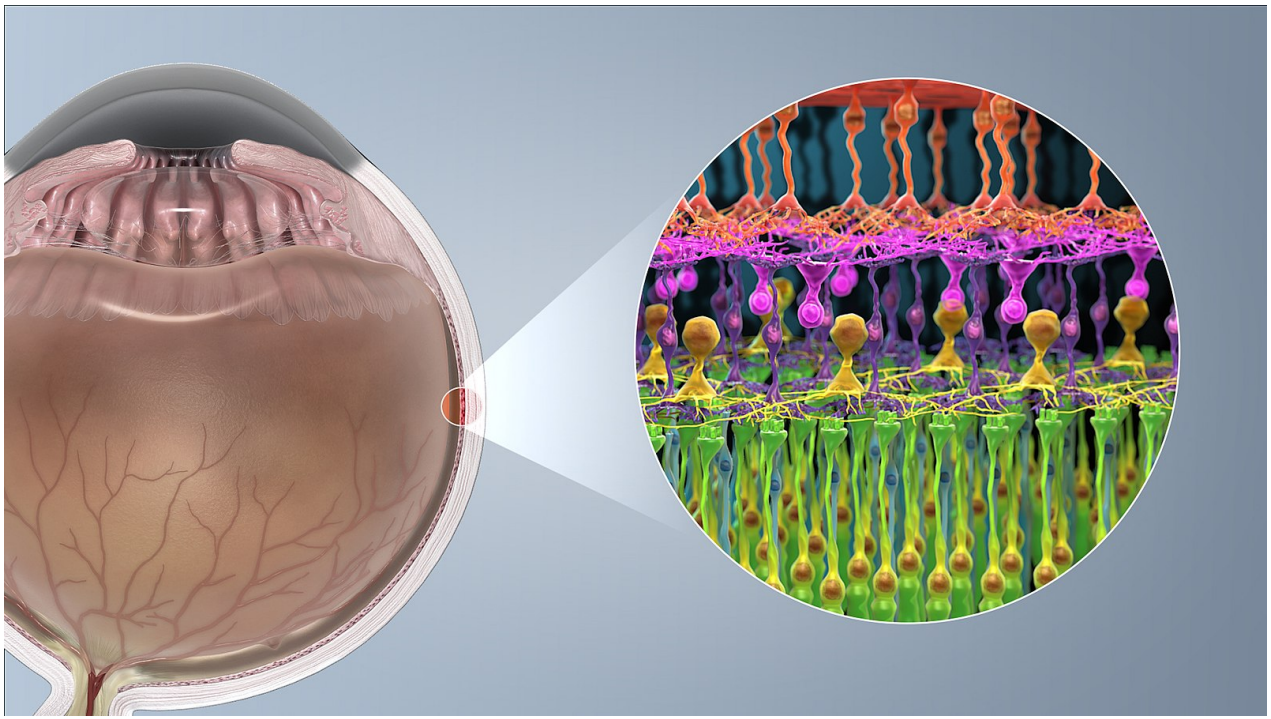
Garrido et al 2011



Goren et al 1975

Visual Psychophysics

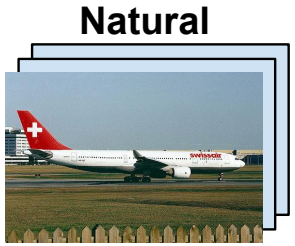
Canonical Early Example: minimum threshold for stimulation of an individual photoreceptor.



Photoreceptor cell cc BY-SA 4.0 Manu5

S. Hecht, S. Shlaer, M. Pirenne, “Energy, quanta and vision,” J. Gen. Physiol., 25 (1942), pp. 819–840.

(a) Traditional

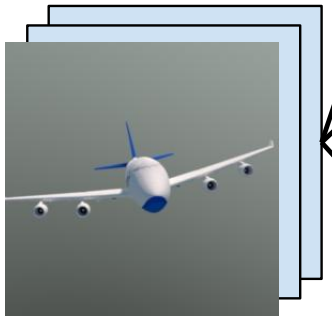


Evaluation

	Top-1	Top-5
A	89.20%	96.30%
B	91.10%	98.60%

(b) Visual Psychophysics

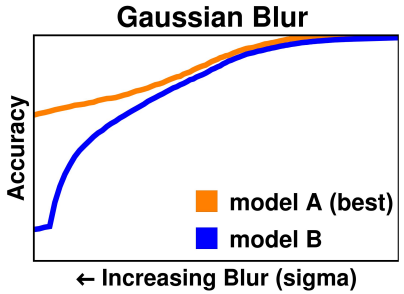
perturbation

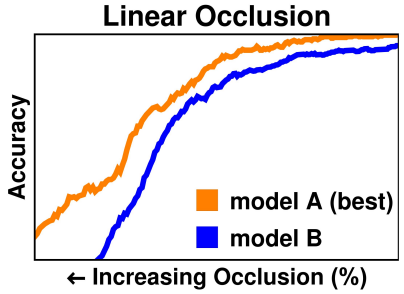


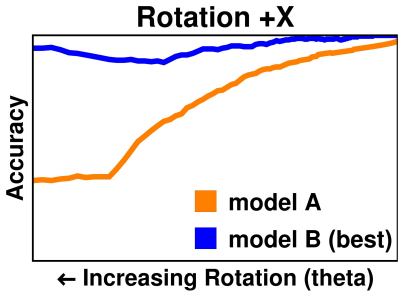
Gaussian blur

Linear Occlusion

Rotation







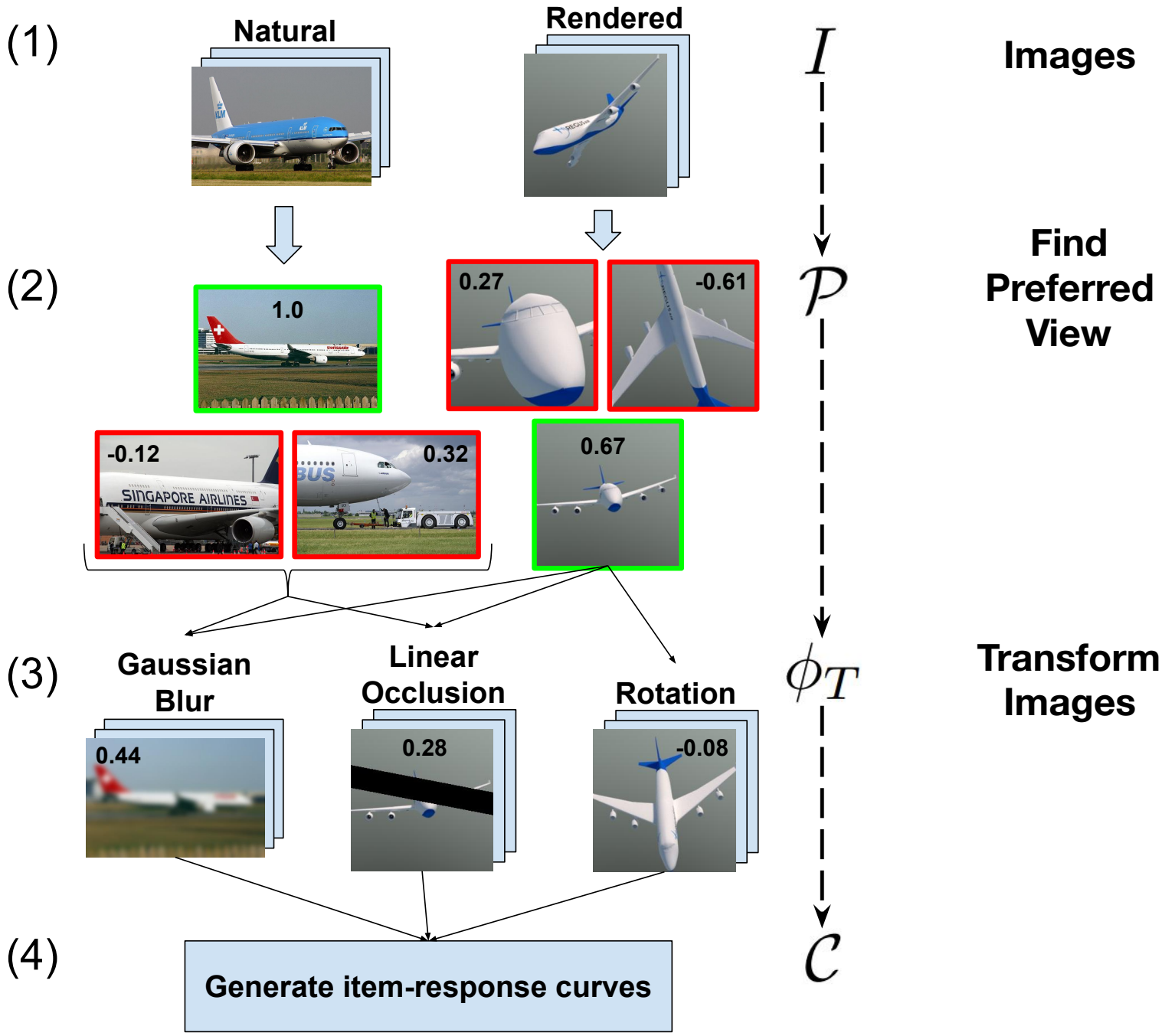
The limits of deep learning…

Brandon RichardWebster

Sam Anthony

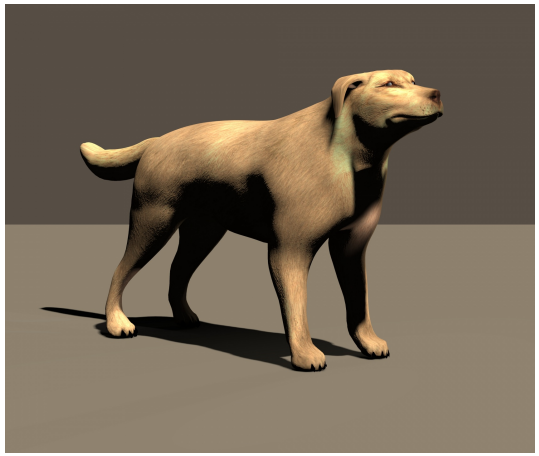
Visual Psychophysics for Object Recognition

https://arxiv.org/abs/1611.06448
IEEE T-PAMI, September 2019

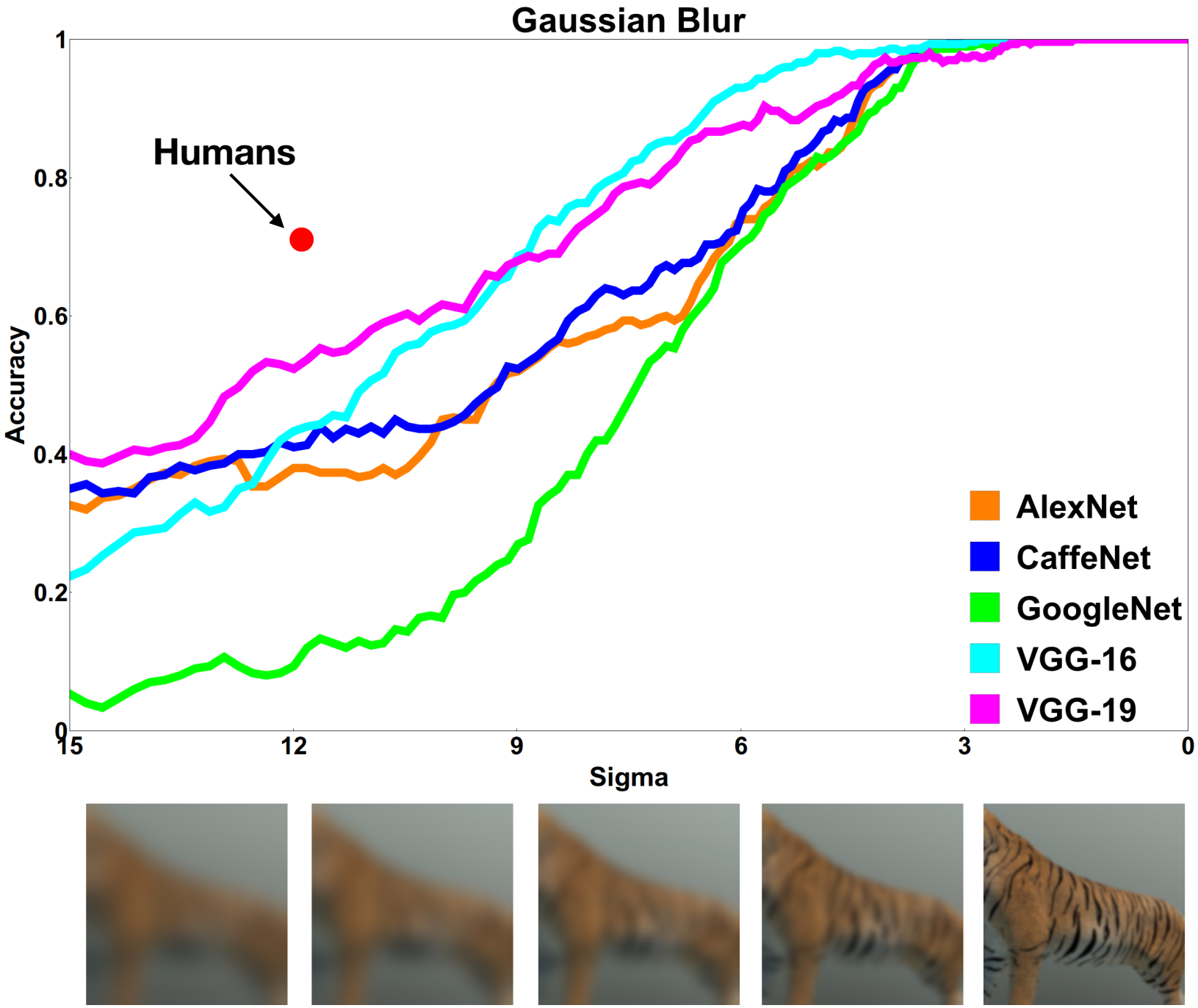

(1)
Natural
Rendered
I
Images
(2)
1.0
0.27
-0.61
$\mathcal{P}$
Find Preferred View
-0.12
0.32
0.67
(3)
Gaussian Blur
Linear Occlusion
Rotation
0.44
0.28
-0.08
ϕ_T
Transform Images
(4)
Generate item-response curves
$\mathcal{C}$

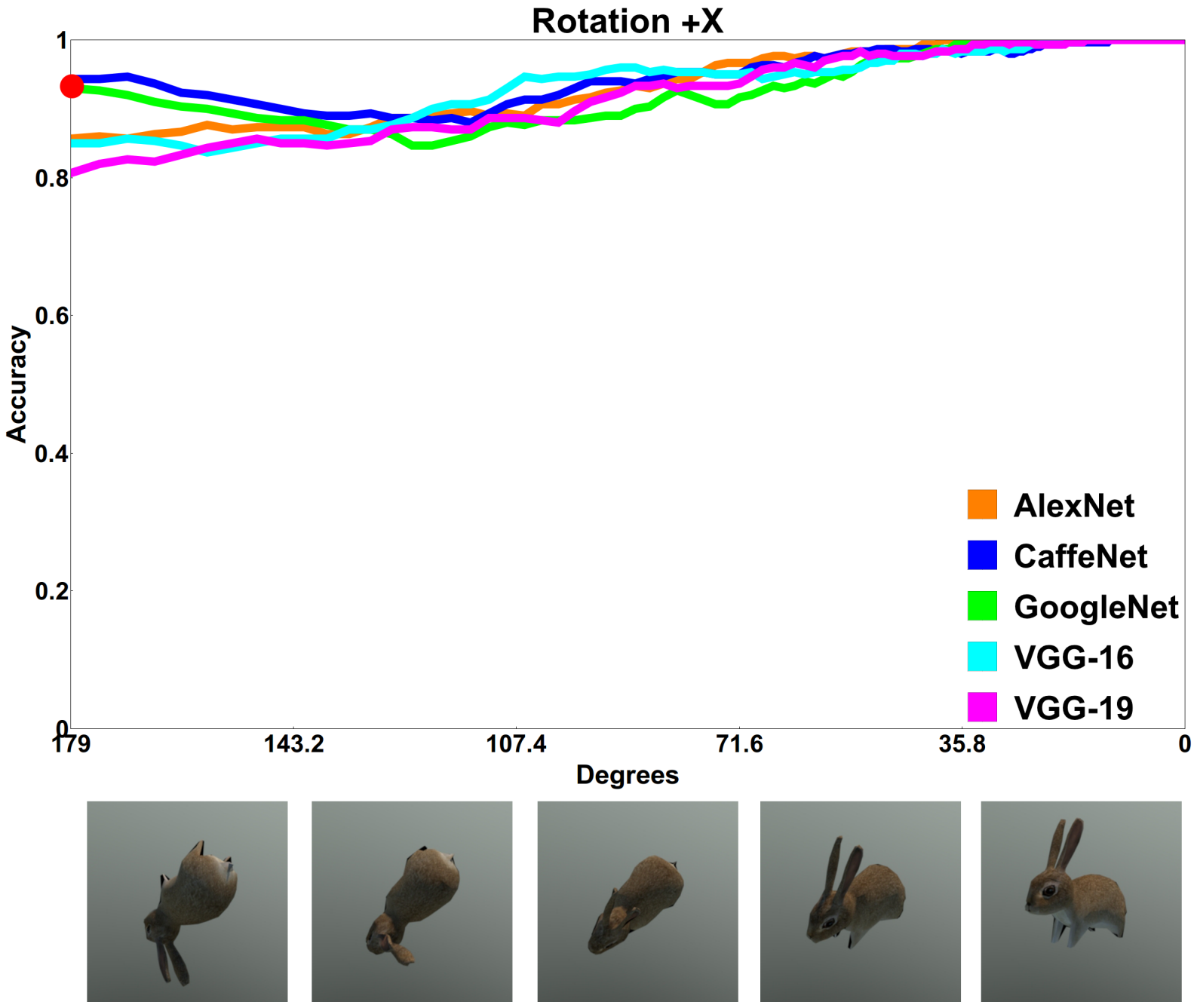
Two-Alternative Forced Choice (2AFC) Task

Matching Alternate Stimulus

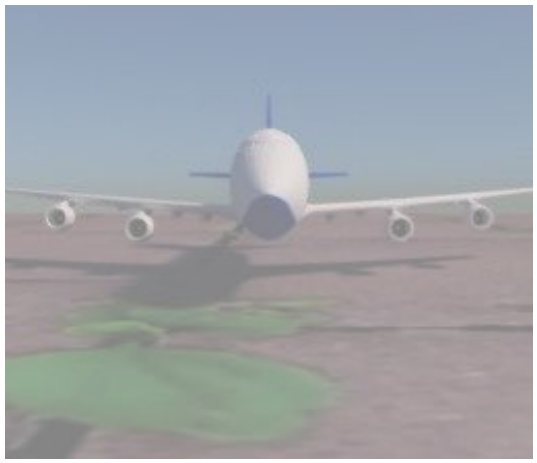


Non-Matching Alternate Stimulus


Gaussian Blur
Humans
Accuracy
1
0.8
0.6
0.4
0.2
0
15
12
9
6
3
0
Sigma
AlexNet
CaffeNet
GoogleNet
VGG-16
VGG-19

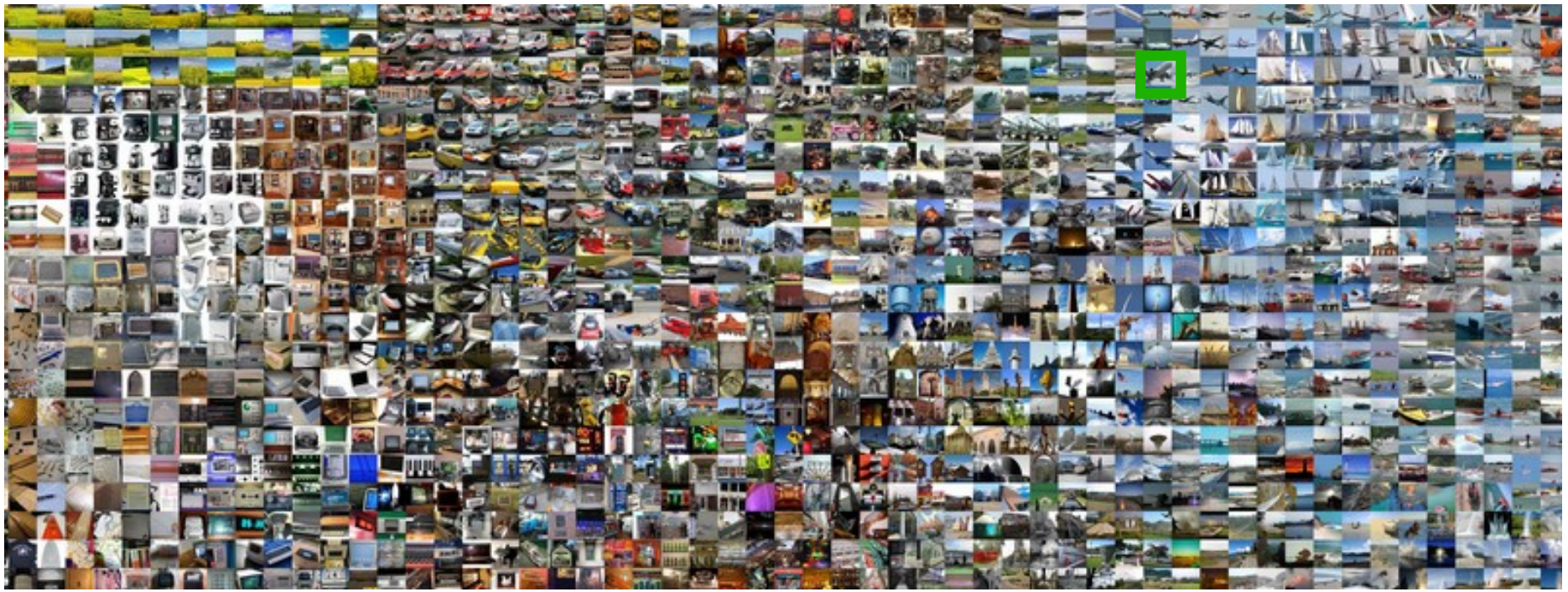

Rotation +X
Accuracy
1
0.8
0.6
0.4
0.2
0
179
143.2
107.4
71.6
35.8
0
Degrees
AlexNet
CaffeNet
GoogleNet
VGG-16
VGG-19

*M*AFC Task

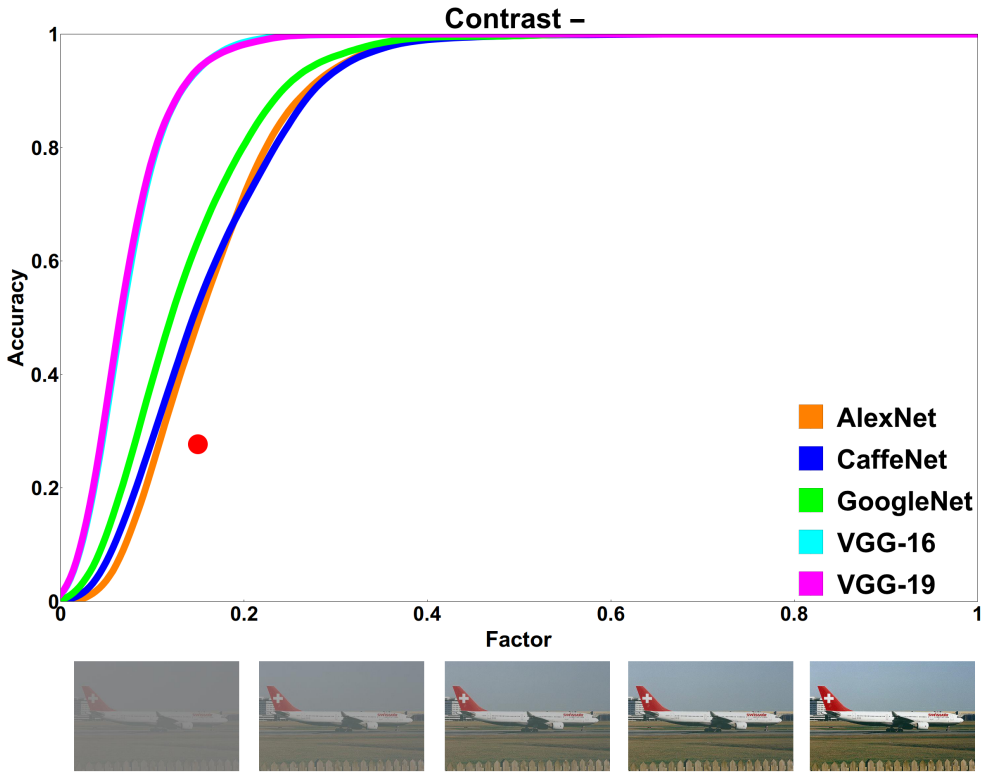


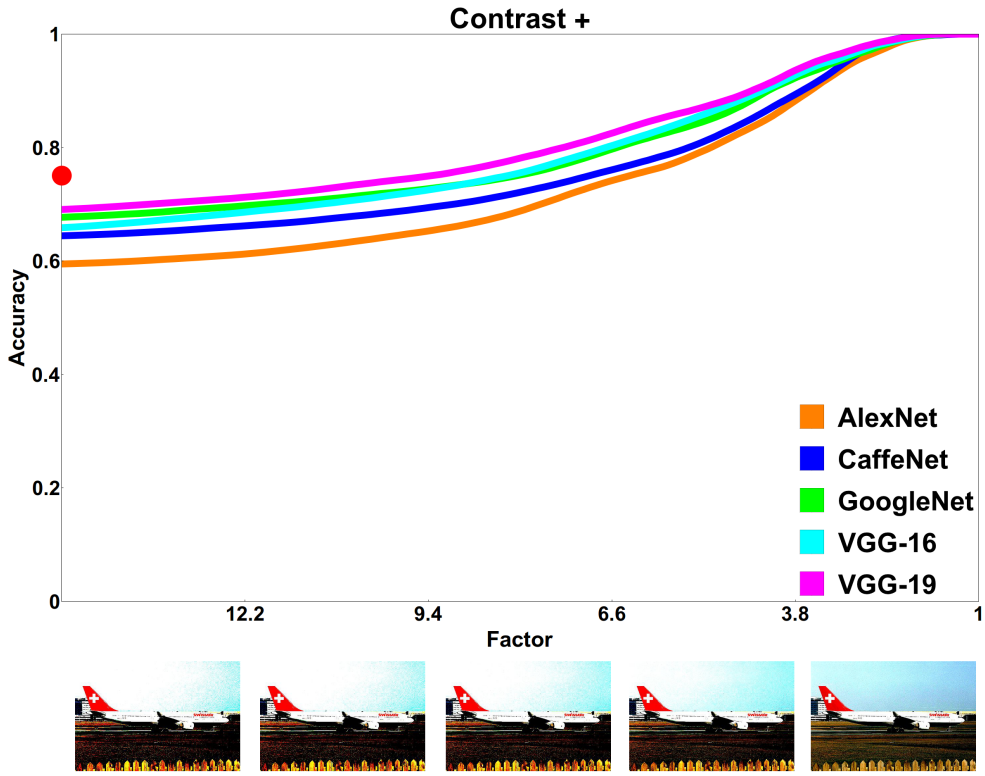
Probe Image

Classification Model



A Curious Contrast Deficit





Brandon RichardWebster

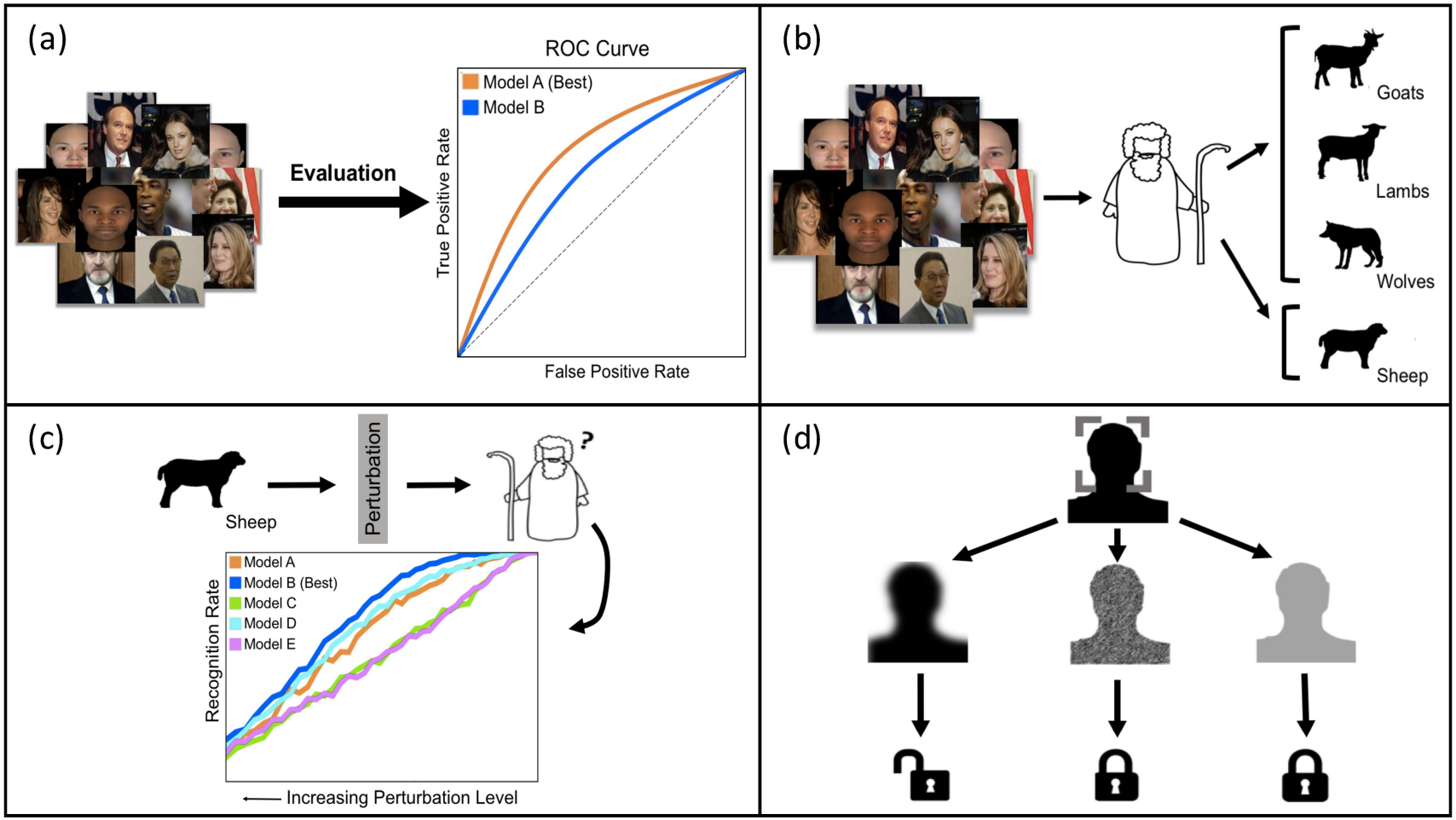
Alli Kwon

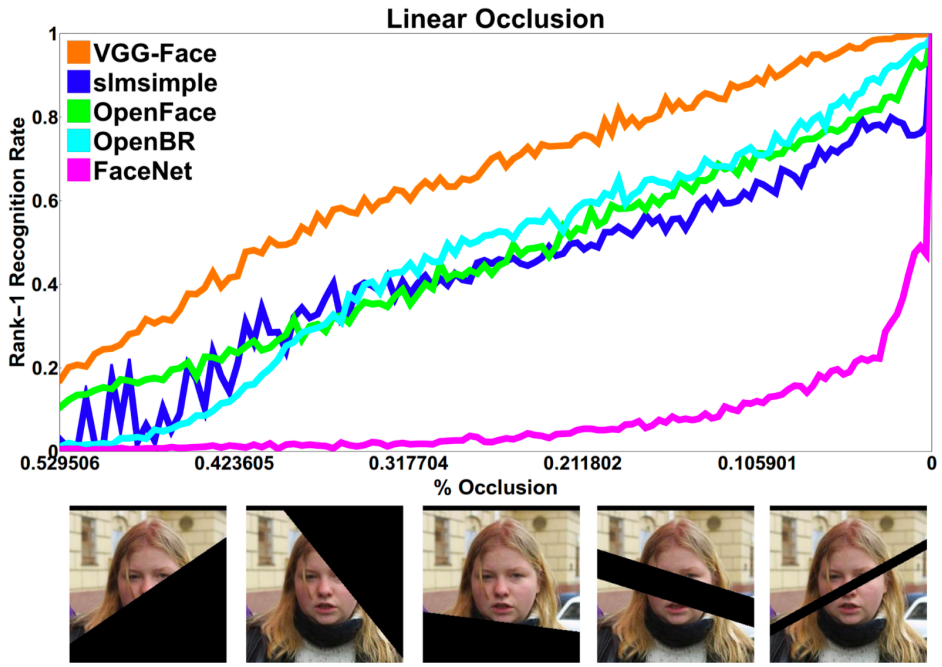
Sam Anthony

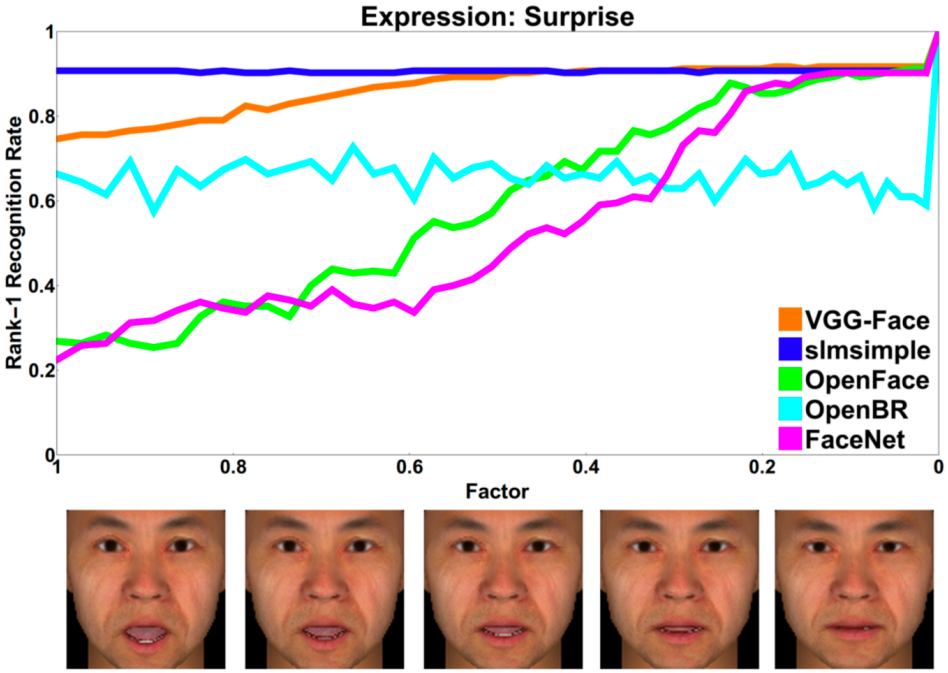
Visual Psychophysics for Face Recognition

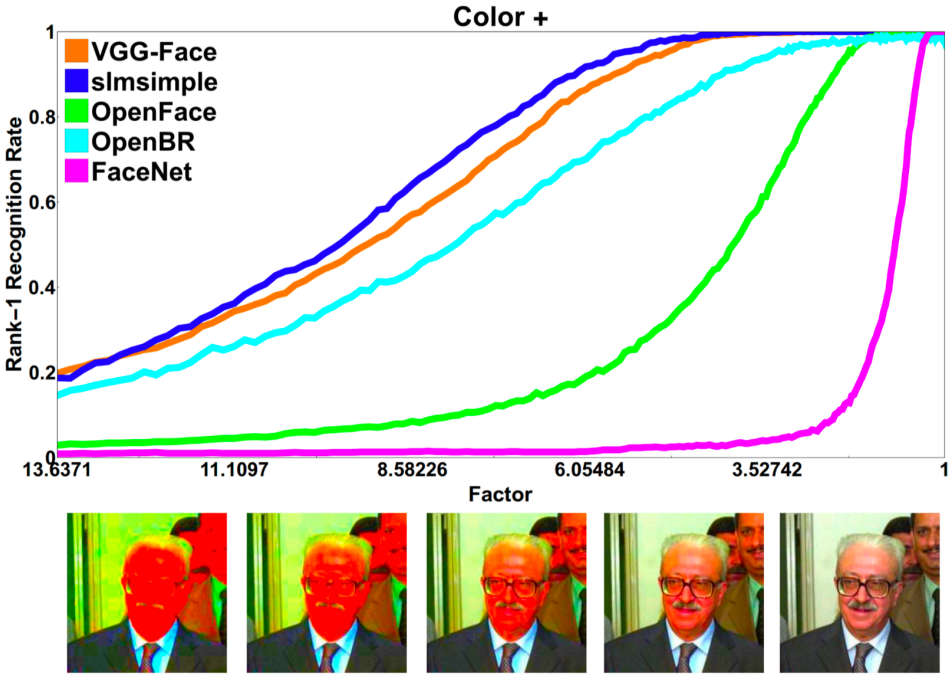
https://arxiv.org/abs/1803.07140
European Conference on Computer Vision, 2018

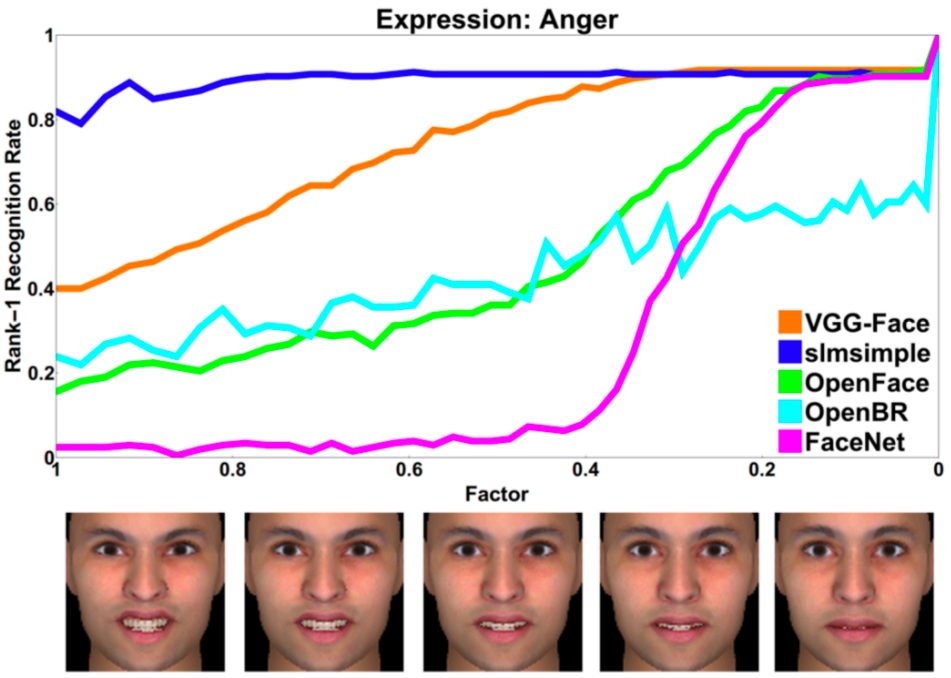
Instead of finding a preferred view, find “sheep” in a biometrics context

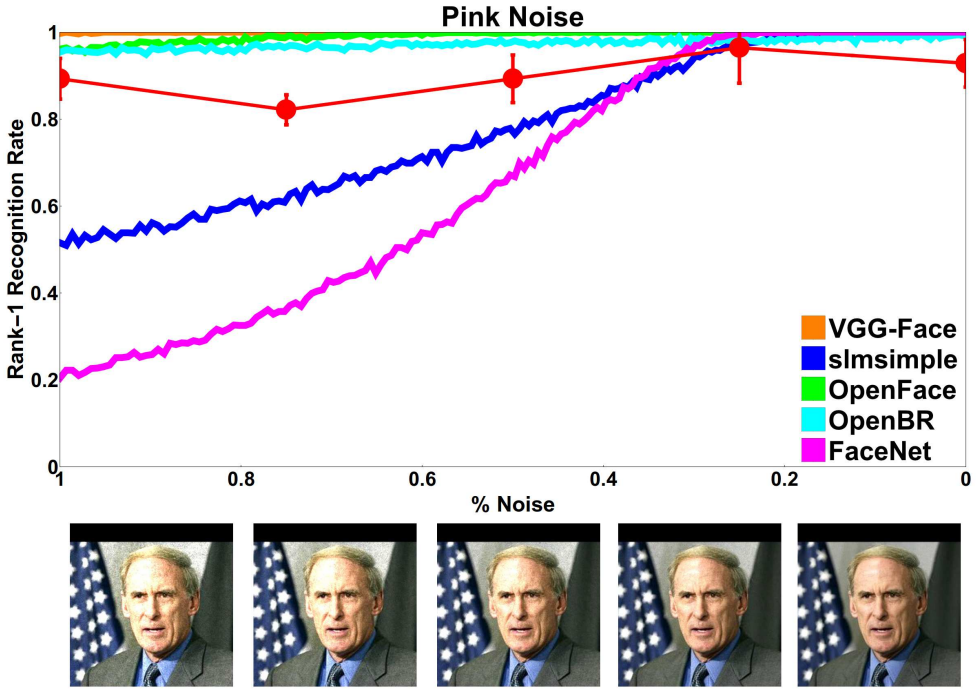


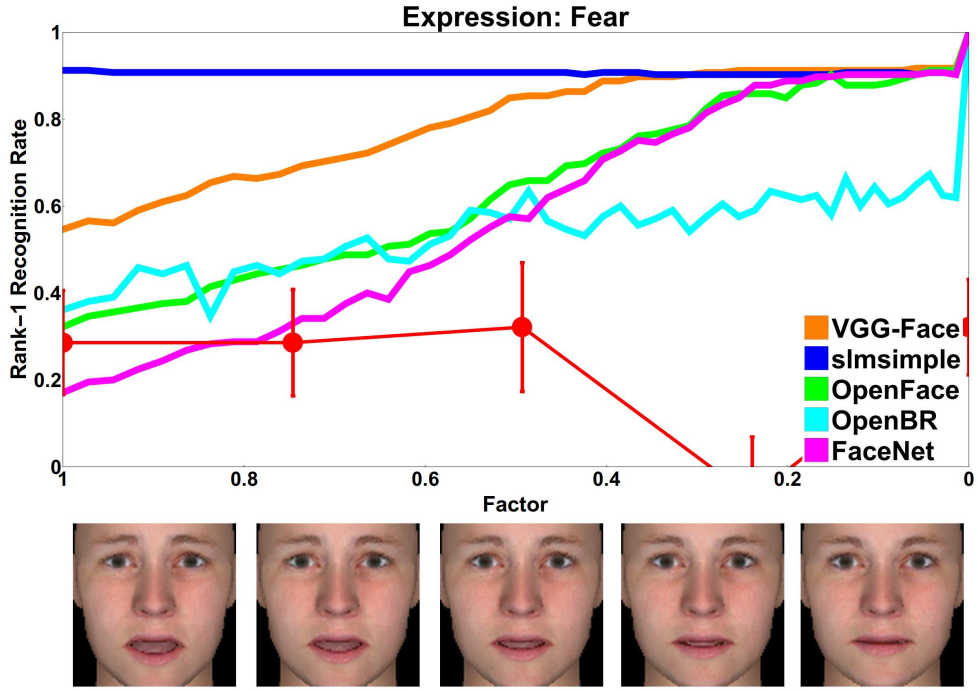

Linear Occlusion
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet
Rank−1 Recognition Rate
1
0.8
0.6
0.4
0.2
0
0.529506
0.423605
0.317704
0.211802
0.105901
0
% Occlusion

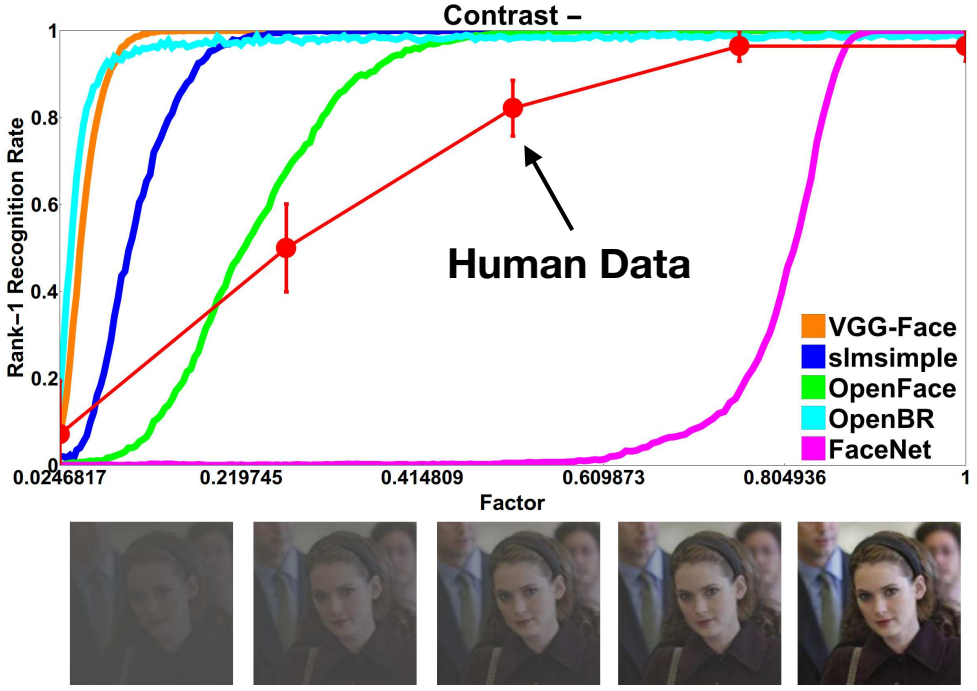

Expression: Surprise
Rank−1 Recognition Rate
1
0.8
0.6
0.4
0.2
0
1
0.8
0.6
0.4
0.2
0
Factor
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet

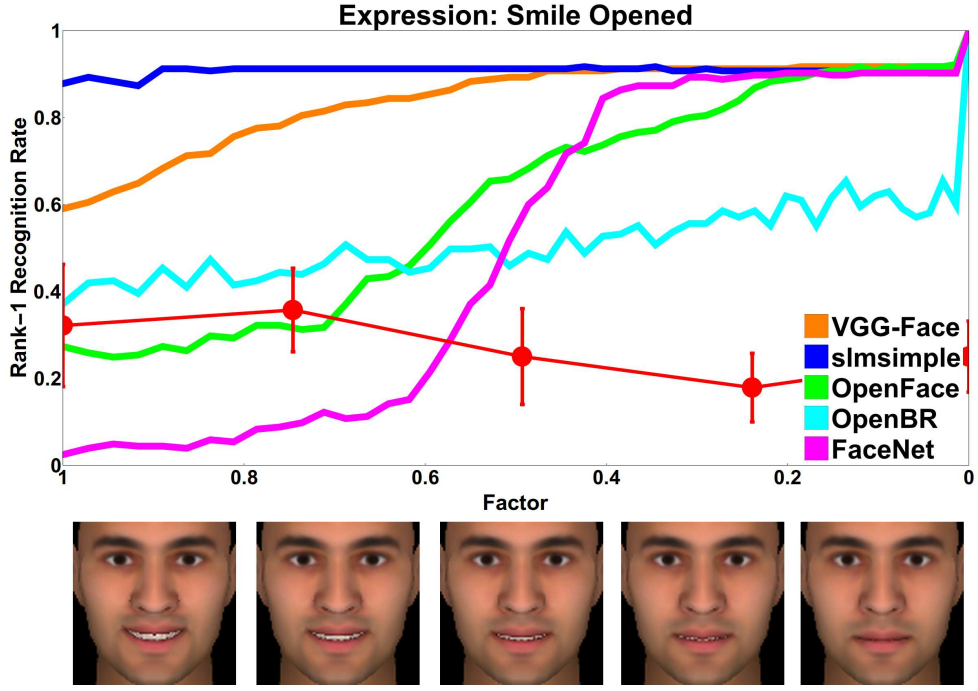

Color +
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet
Rank−1 Recognition Rate
1
0.8
0.6
0.4
0.2
0
13.6371
11.1097
8.58226
6.05484
3.52742
1
Factor


Expression: Anger
Rank−1 Recognition Rate
1
0.8
0.6
0.4
0.2
0
1
0.8
0.6
0.4
0.2
0
Factor
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet


Pink Noise
Rank−1 Recognition Rate
% Noise
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet


Expression: Fear
Rank−1 Recognition Rate
Factor
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet


Contrast −
Rank−1 Recognition Rate
Human Data
Factor
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet


Expression: Smile Opened
Rank−1 Recognition Rate
Factor
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet

The potentials of deep learning…

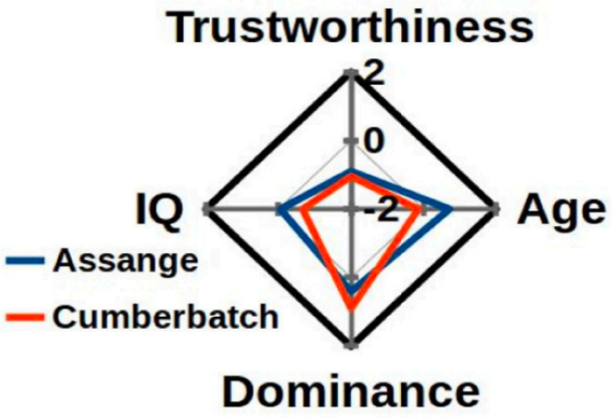
Mel McCurrie

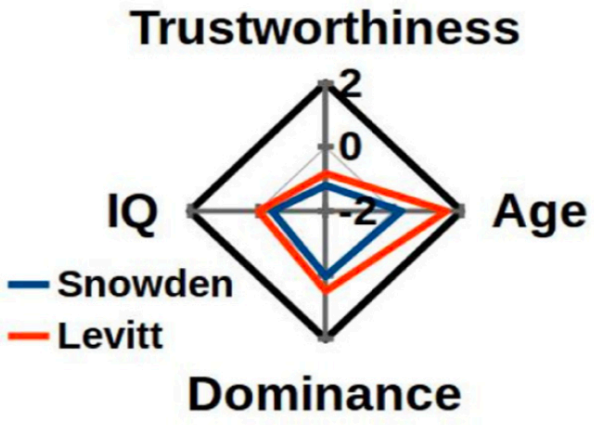
Sam Anthony

CNNs for Subjective Face Attributes

https://arxiv.org/abs/1610.08119
Image and Vision Computing, October 2018

Subjective Attribute Predictions

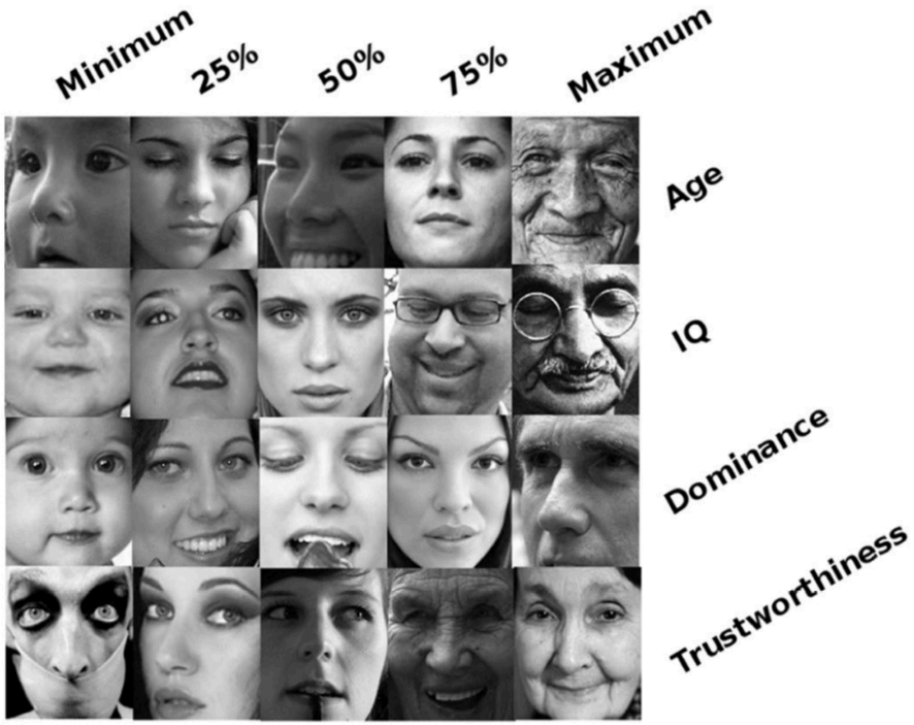


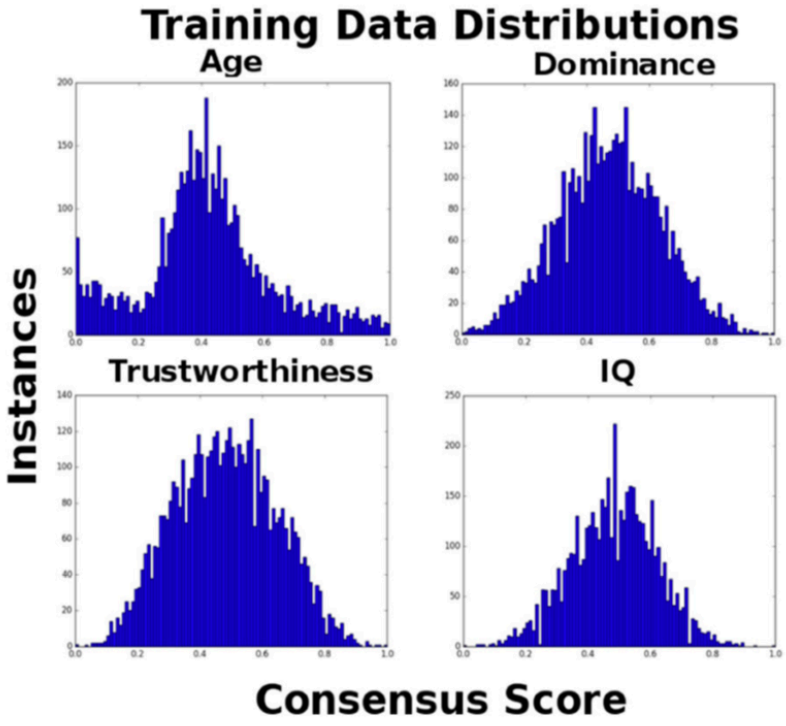


N.N. Oosterhof, A. Todorov, “The functional basis of face evaluation,” PNAS vol. 105, no. 32, 2008

Data collection via TestMyBrain.org

Data Distributions



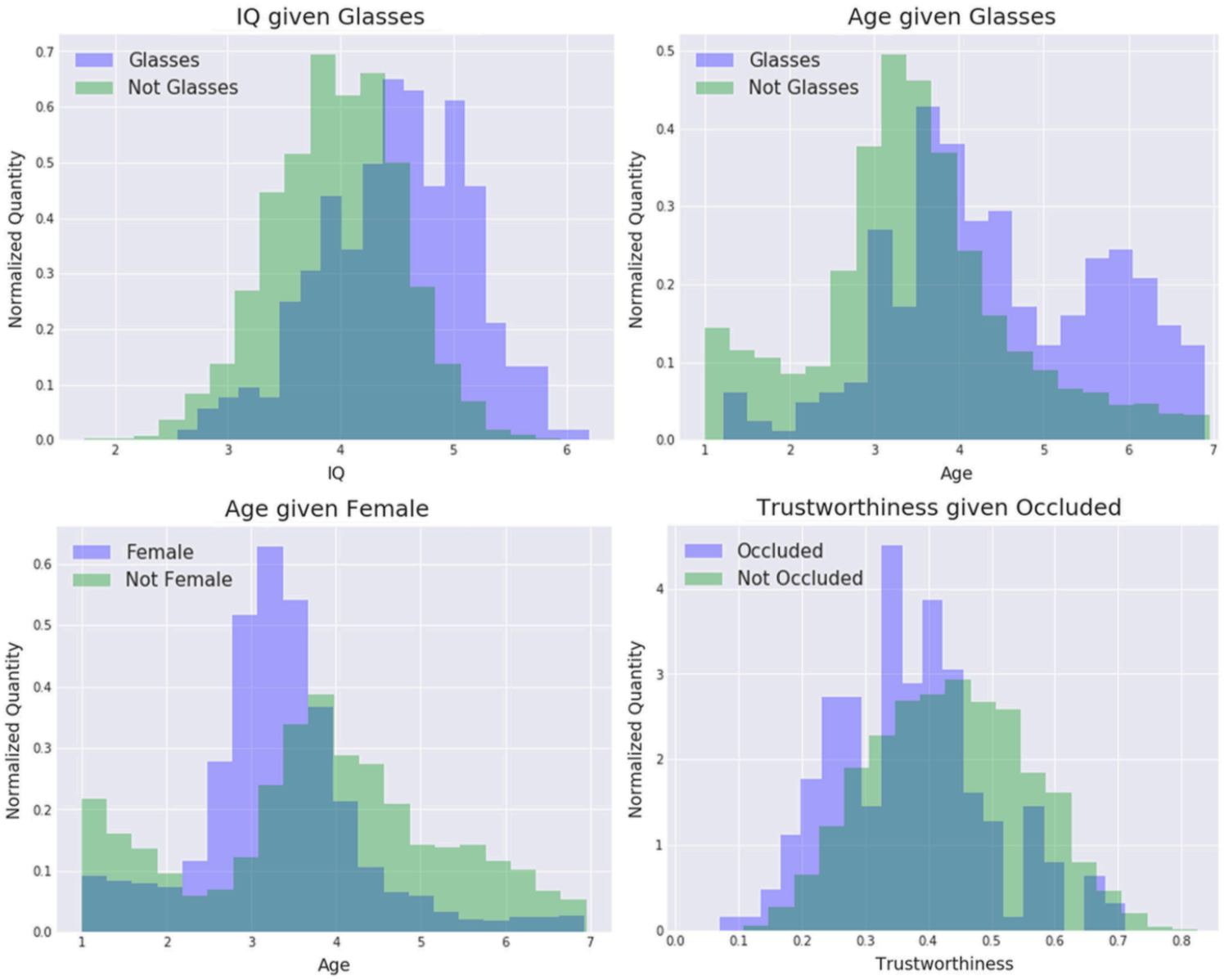


Dataset Statistics

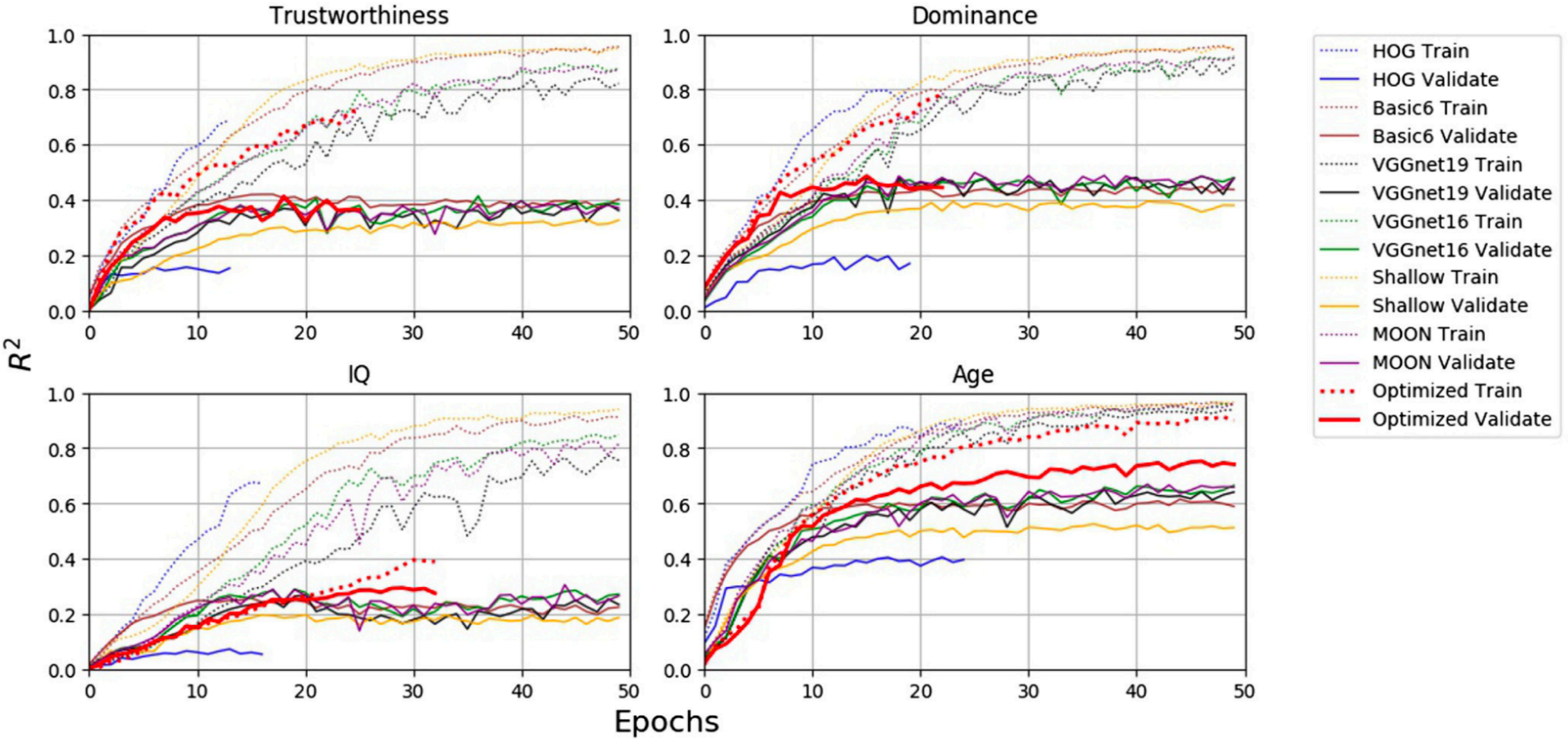
5040 Total Images

	Trust.	Dom.	Age	IQ
Mean of ratings	0.48	0.47	0.42	0.48
Std. of ratings	0.16	0.16	.20	0.14
Mean Std. of ratings	0.34	0.32	0.13	0.27
Mean Num. of ratings	32.47	32.19	15.80	15.79

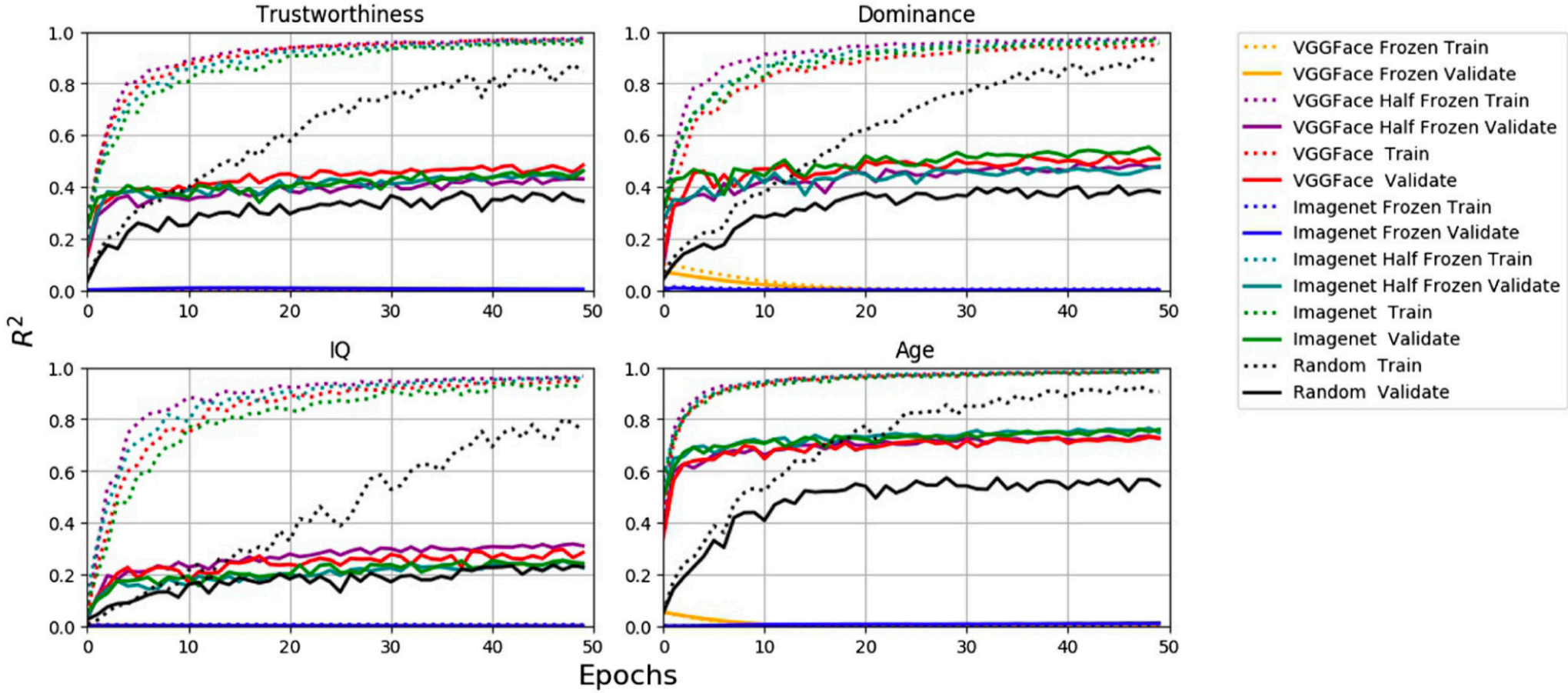
Dataset Patterns



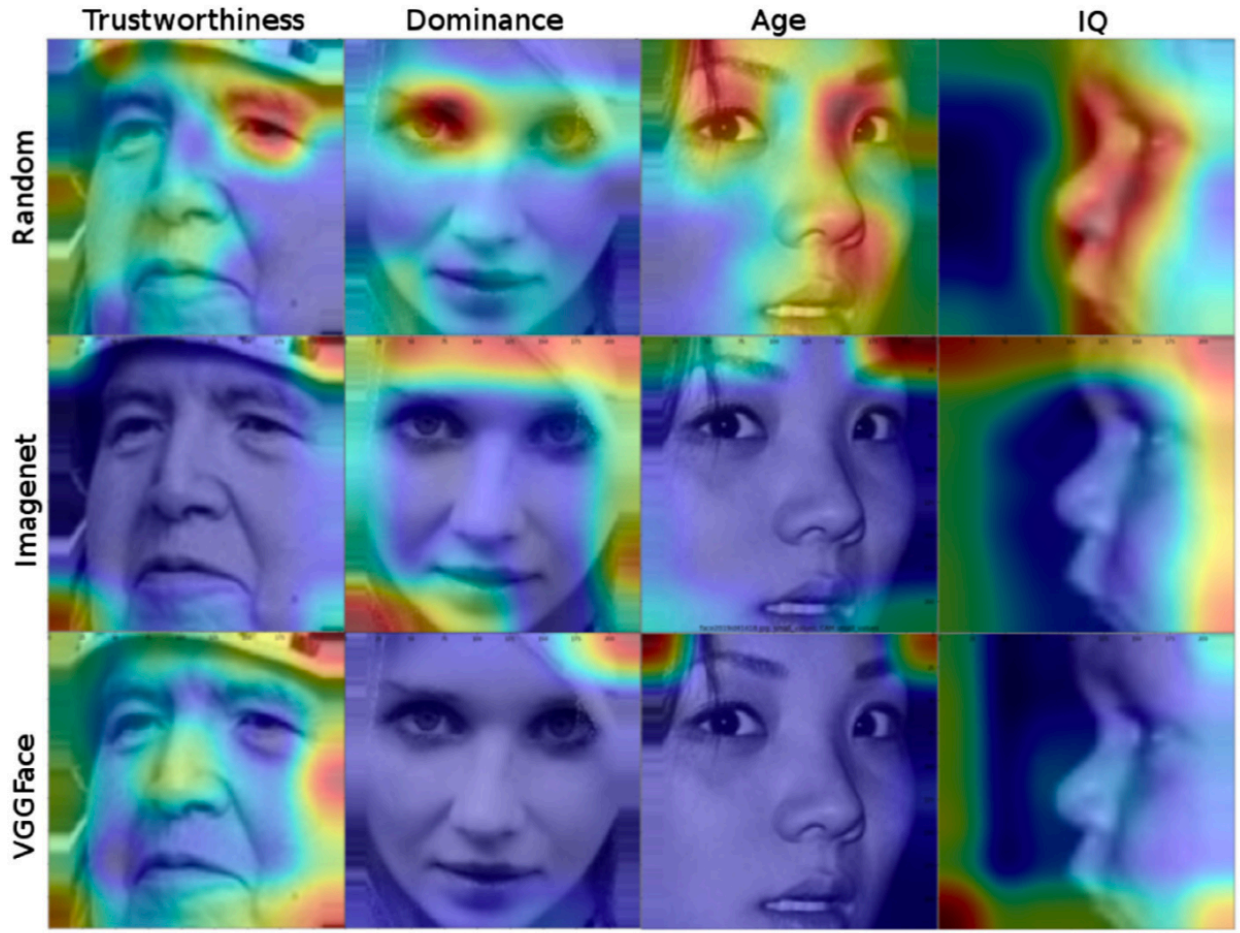
Which architecture is the best?



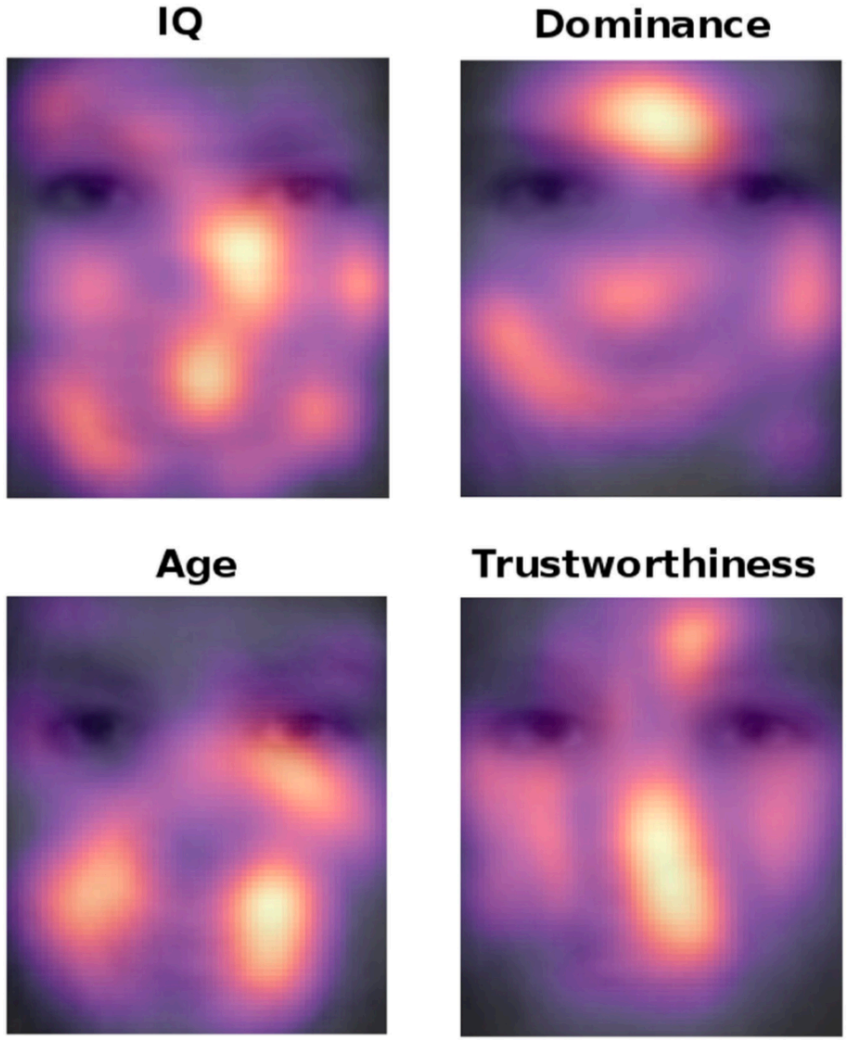
The dataset is small. Does transfer learning help?



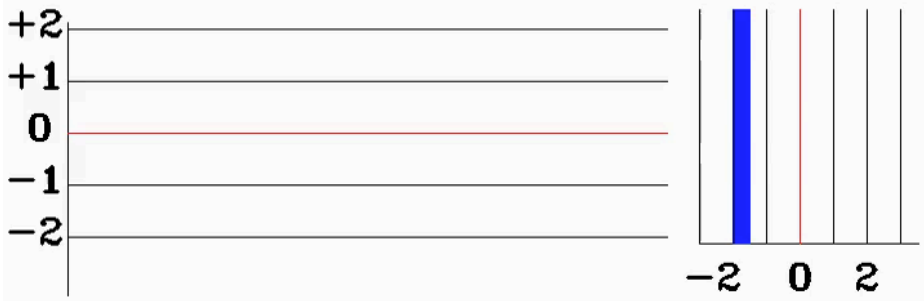
Regions of the face most important to the models (transfer learning)



Regions of the face most important to the models (saliency from occlusion)



Trustworthiness



What other attributes are interesting?

Thank you!