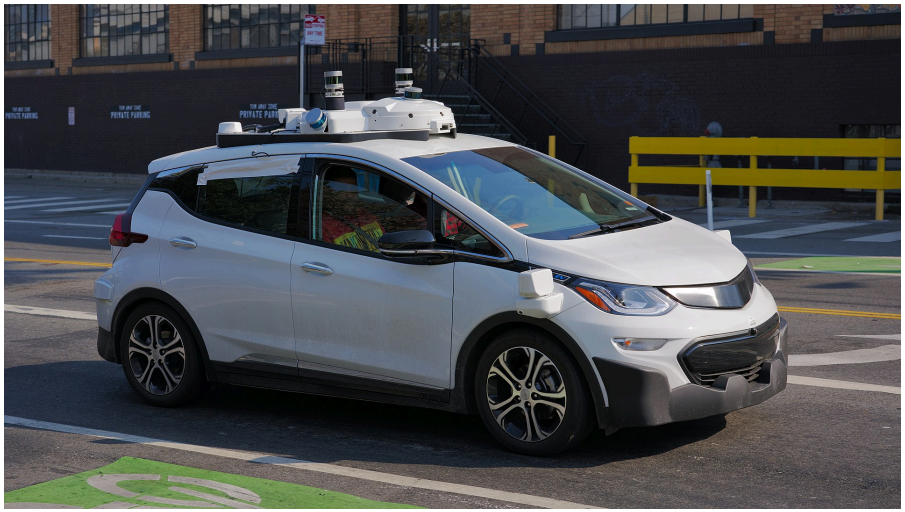
A Psychophysics Driven Evaluation Framework for Visual Recognition

Walter J. Scheirer

Computer Vision Research Laboratory
Department of Computer Science and Engineering
University of Notre Dame

Imagine the following scenario:



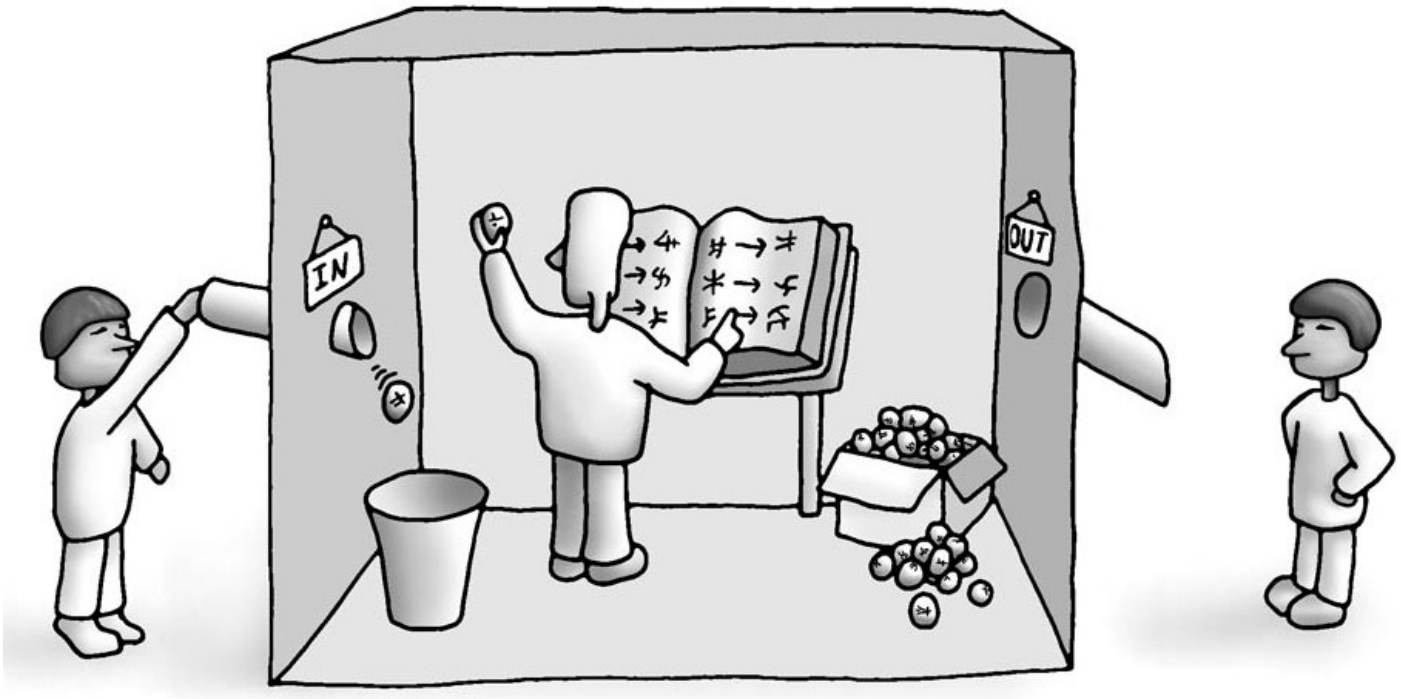
A Cruise Automation Chevrolet Bolt, third generation, seen in San Francisco.

A proprietary autonomous vehicle system purportedly solves driving with human-like ability.

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

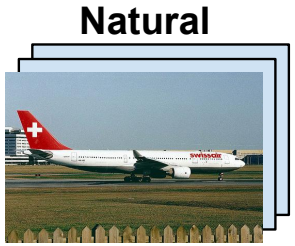
How would you go about falsifying the claim of human-like ability?

The Chinese Room



Searle. Behavioral and Brain Sciences 1980

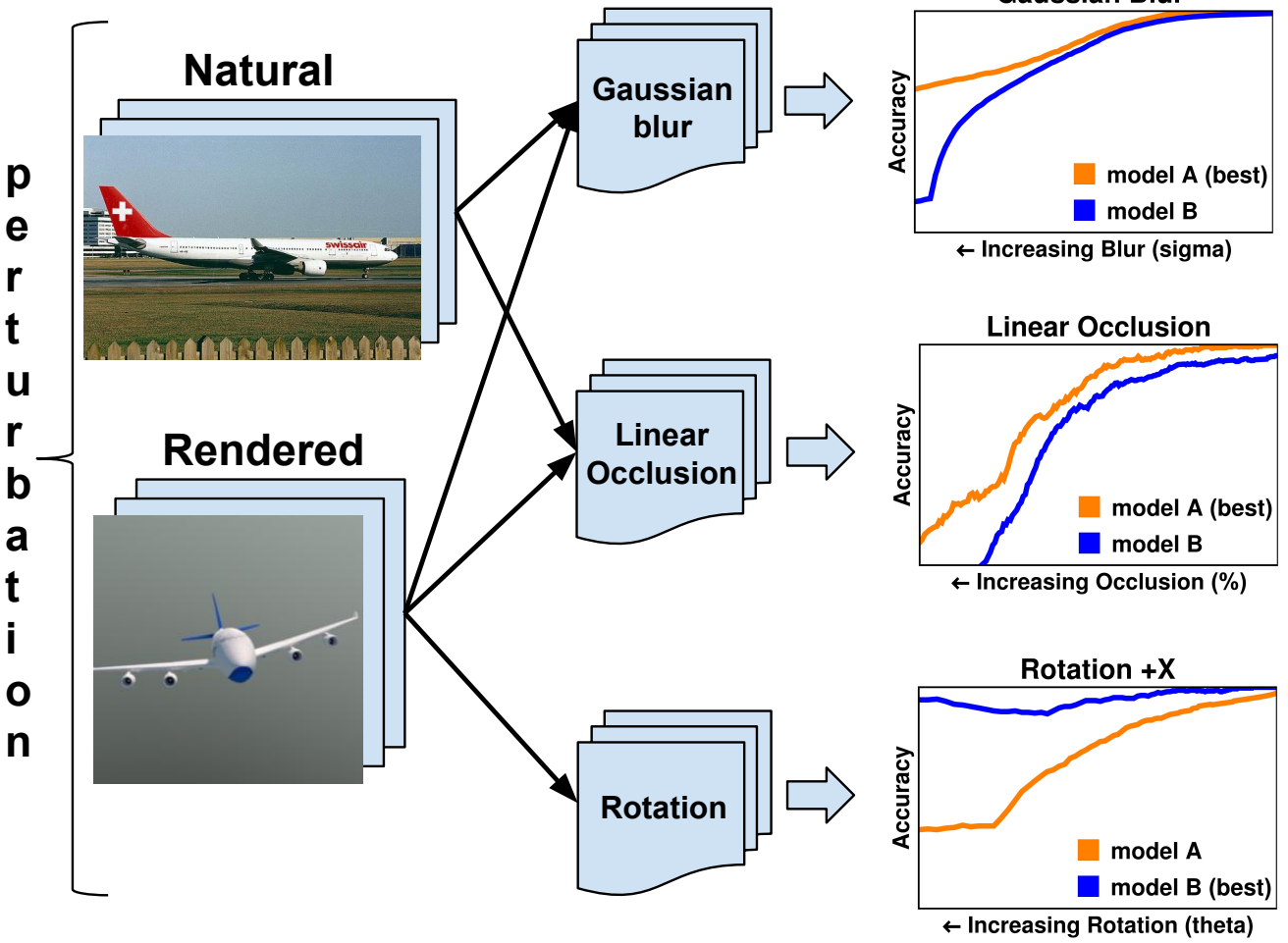
(a) Traditional



Evaluation

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

(b) Visual Psychophysics



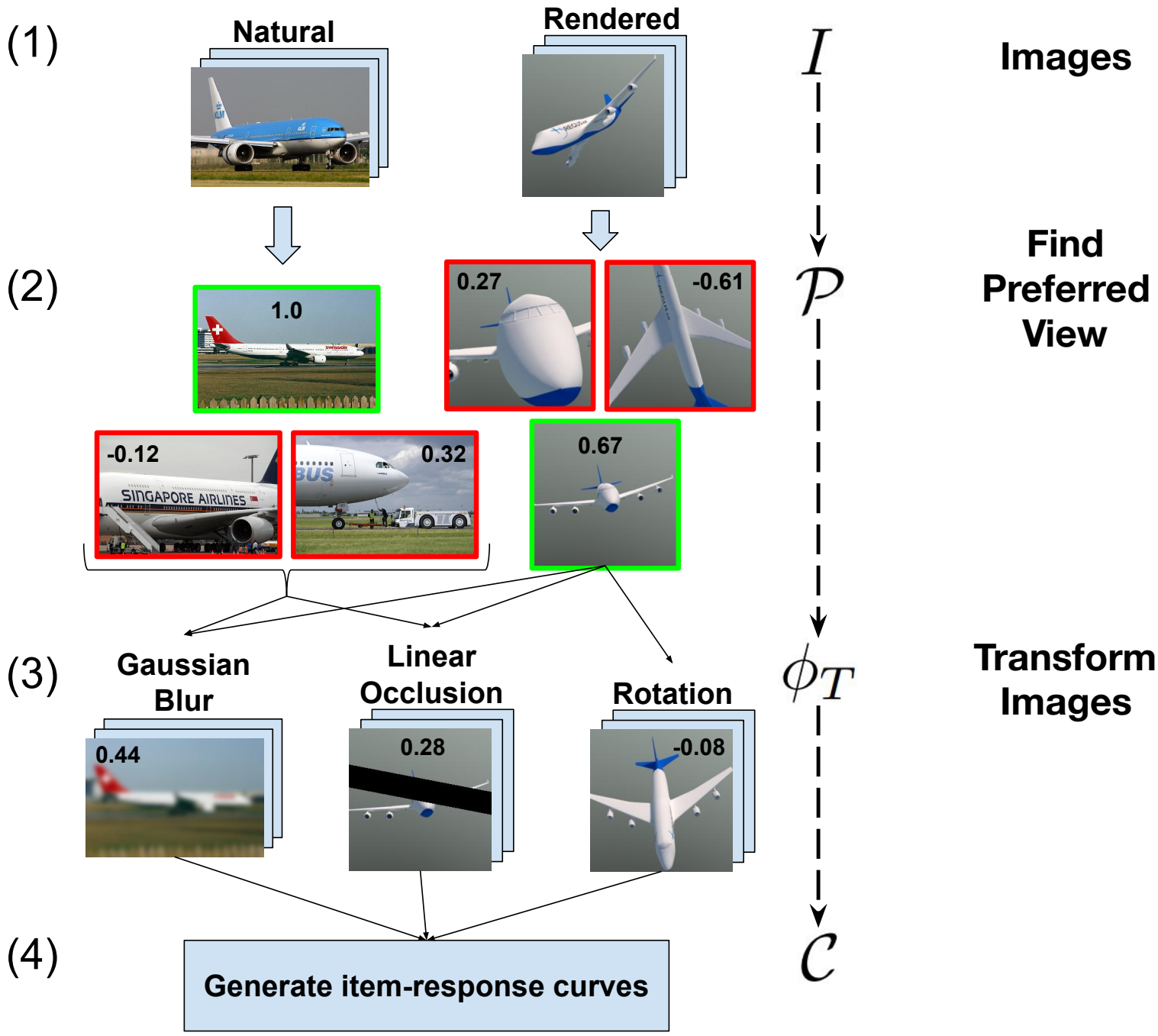
Brandon RichardWebster

Sam Anthony

Visual Psychophysics for Object Recognition

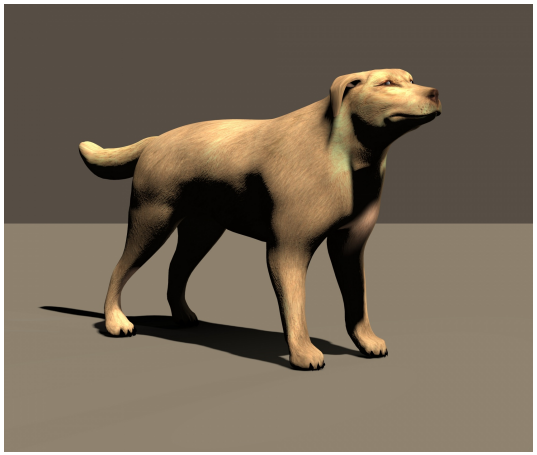
https://arxiv.org/abs/1611.06448

(To appear in T-PAMI)

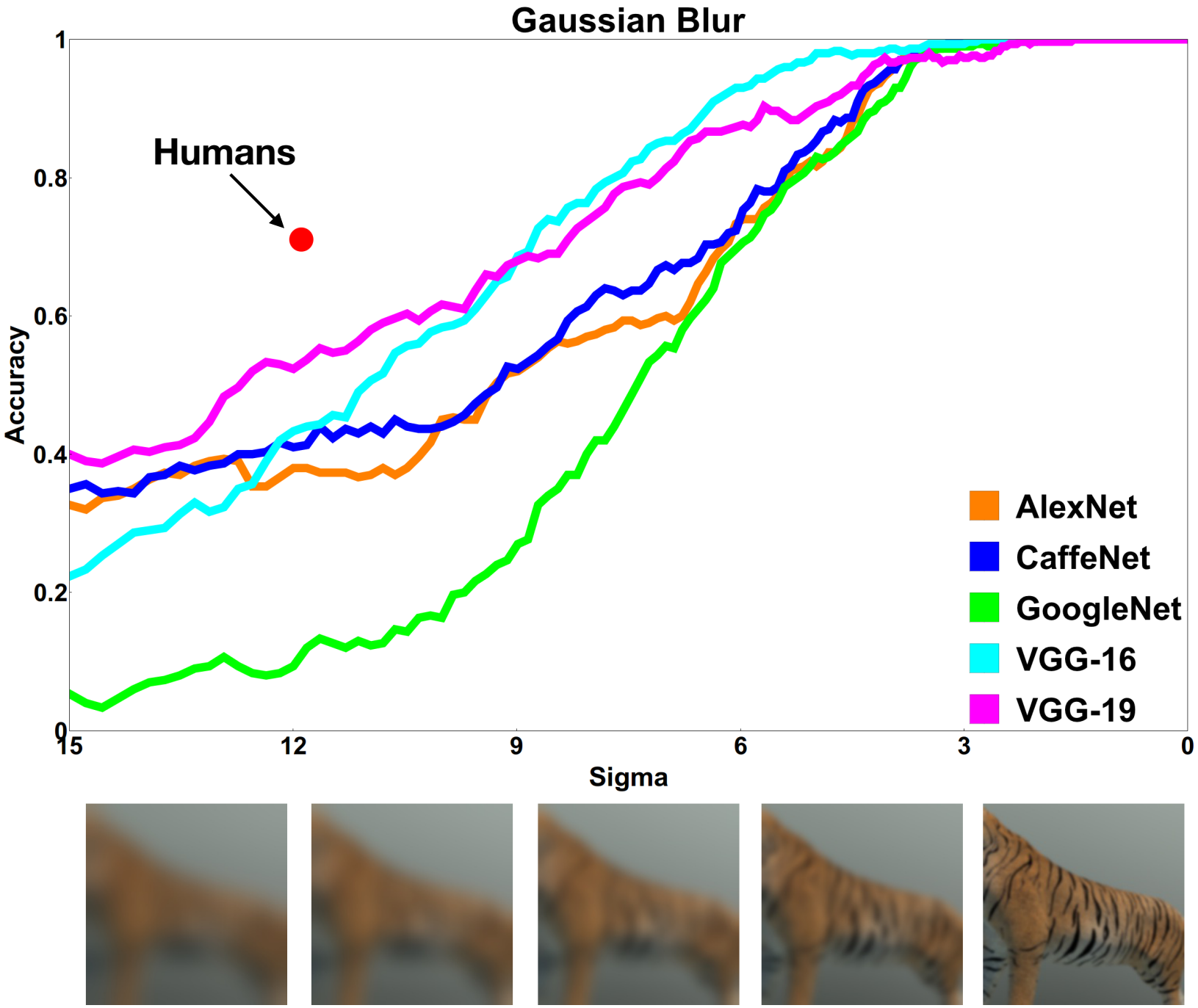

(1)
Natural
Rendered
I
Images
(2)
1.0
0.27
-0.61
$\mathcal{P}$
Find Preferred View
-0.12
0.32
0.67
(3)
Gaussian Blur
Linear Occlusion
Rotation
0.44
0.28
-0.08
ϕ_T
Transform Images
(4)
Generate item-response curves
$\mathcal{C}$

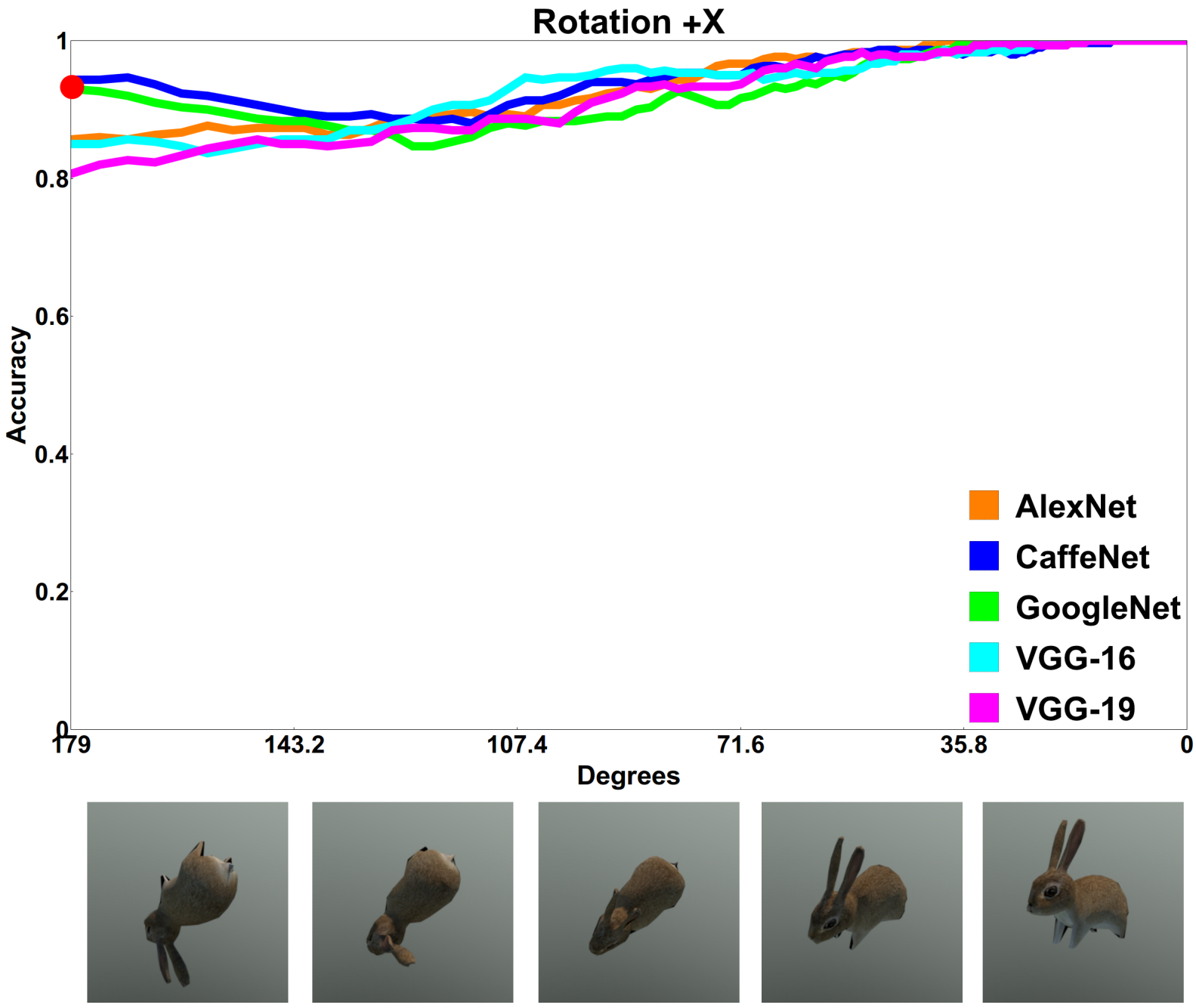
Two-Alternative Forced Choice (2AFC) Task

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

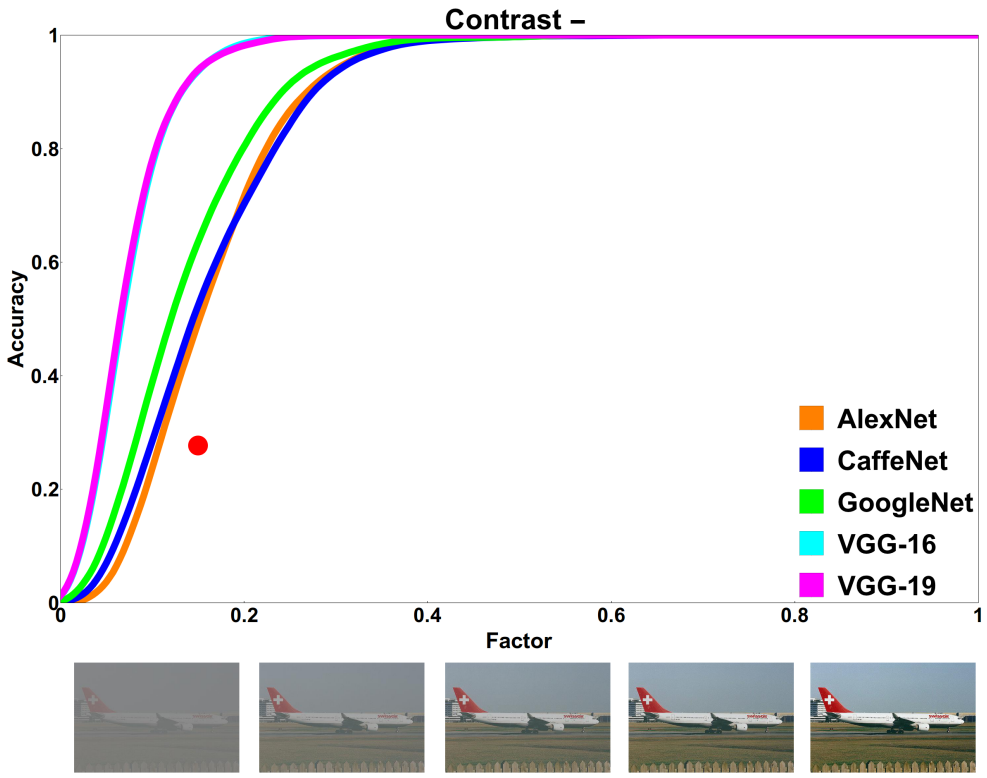

Rotation +X
Accuracy
1
0.8
0.6
0.4
0.2
0
179
143.2
107.4
71.6
35.8
0
Degrees
AlexNet
CaffeNet
GoogleNet
VGG-16
VGG-19

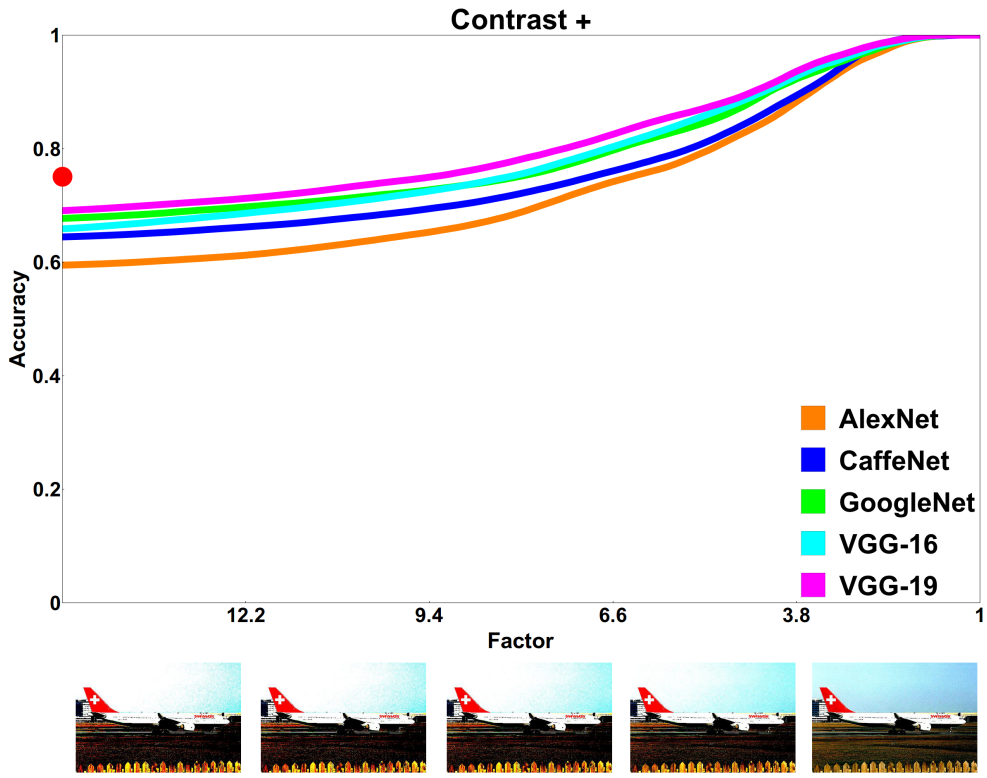
*M*AFC Task

Probe Image

Classification Model

A Curious Contrast Deficit





Brandon RichardWebster

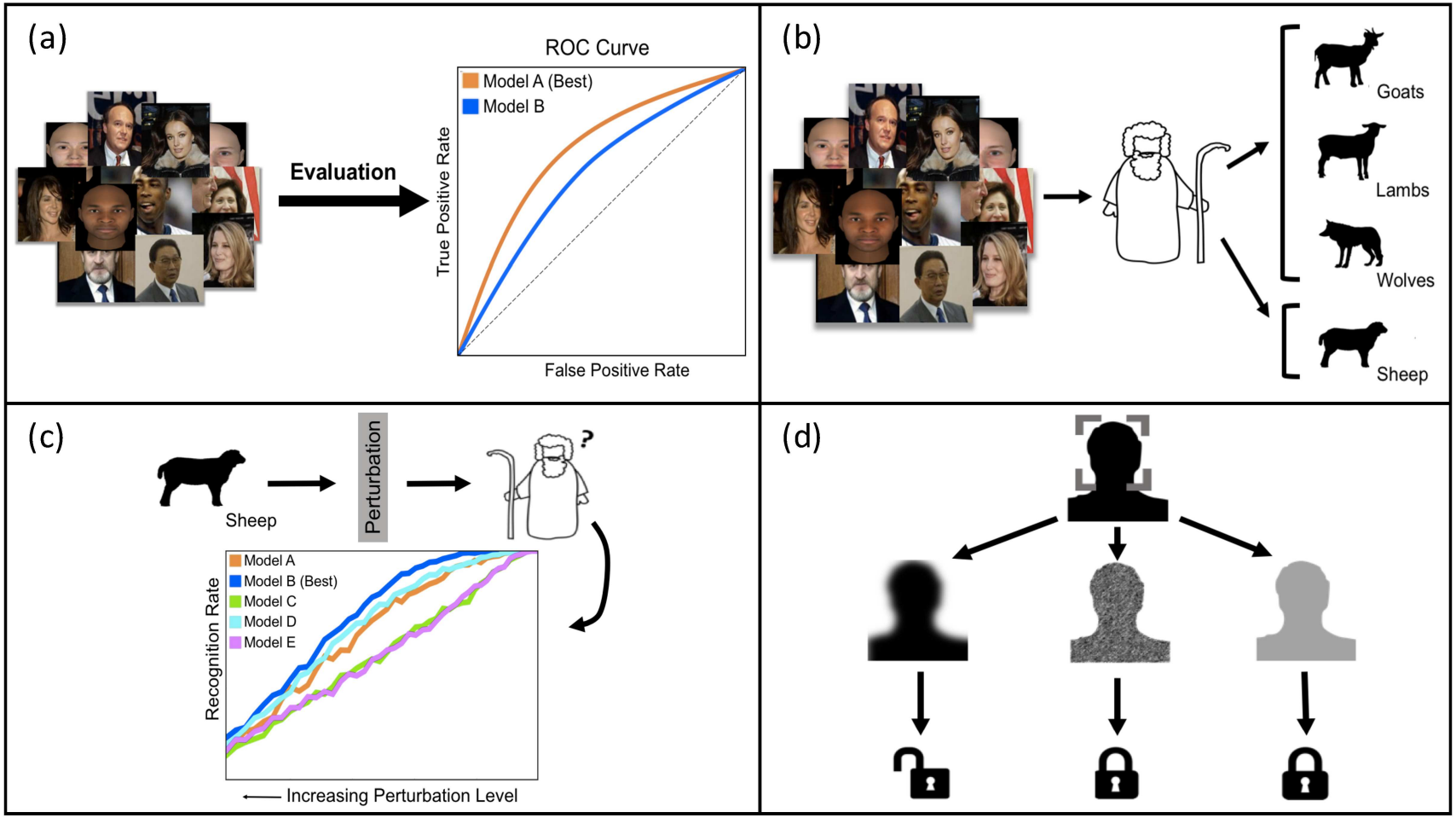
Alli Kwon

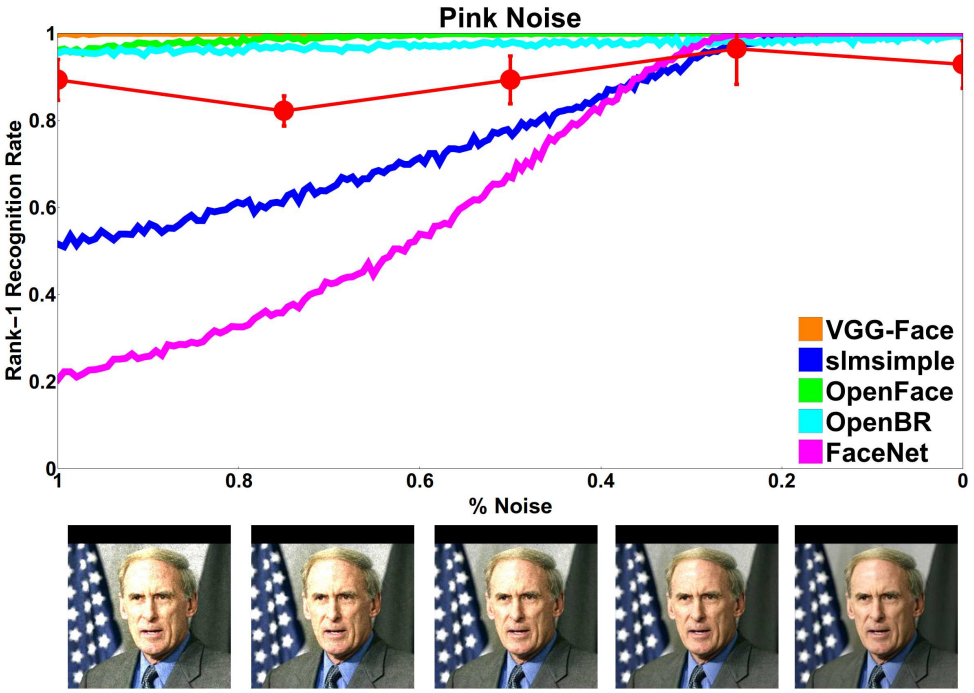
Sam Anthony

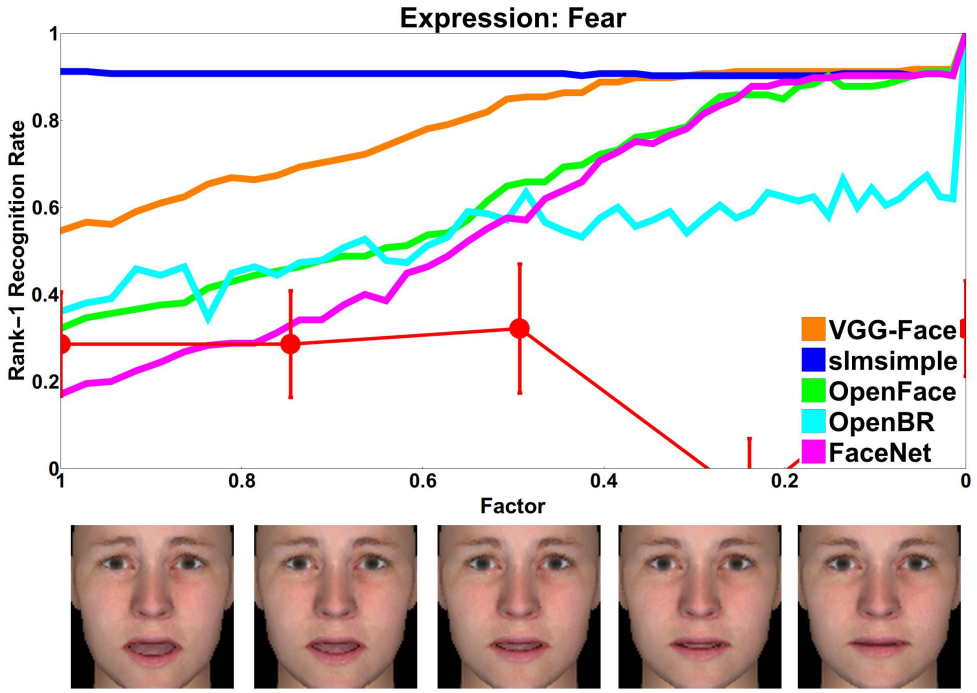
Visual Psychophysics for Face Recognition

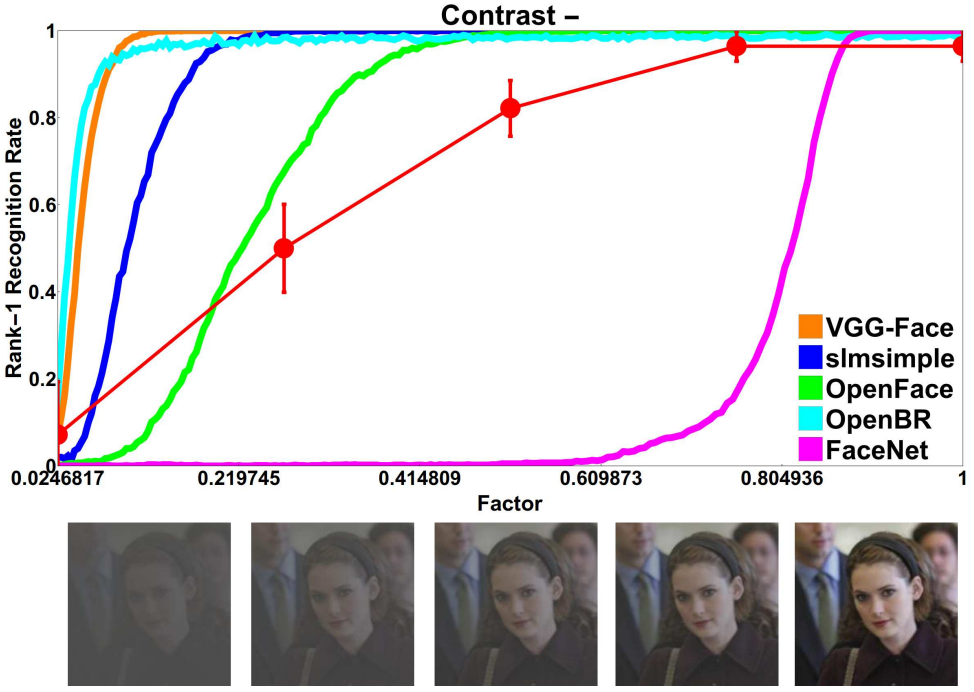
https://arxiv.org/abs/1803.07140

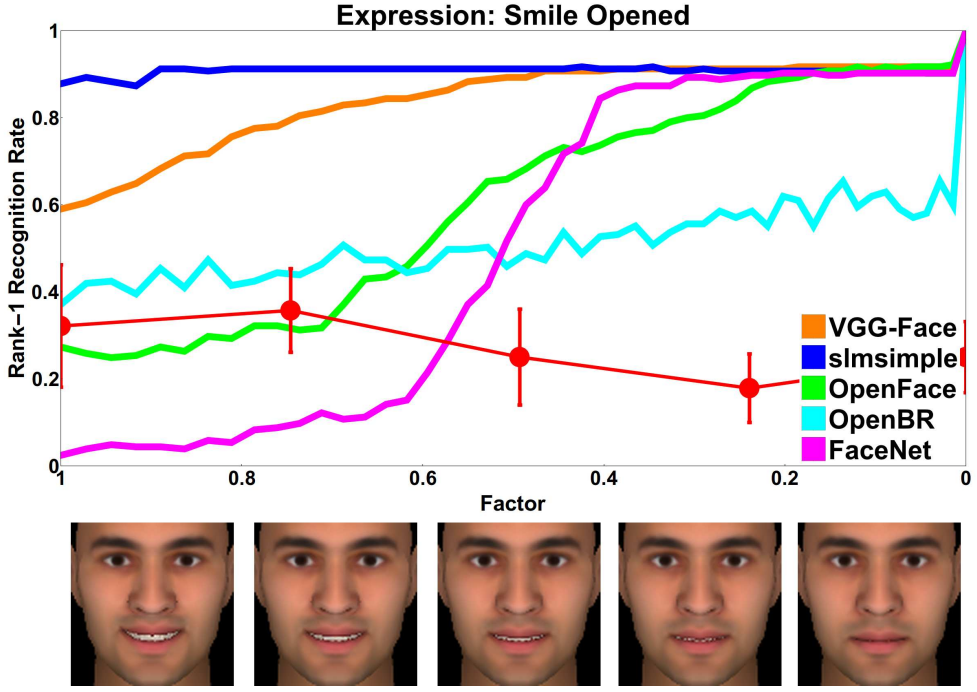
Instead of finding a preferred view, find “sheep” in a biometrics context




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
0.0246817
0.219745
0.414809
0.609873
0.804936
Factor
VGG-Face
slmsimple
OpenFace
OpenBR
FaceNet


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

A few parting thoughts for robotics

- There is a large disparity between dataset performance and real world performance
  - Think self driving cars in 2018
- Visual psychophysics is a primary tool used to study vision in psychology and neuroscience
  - Why aren't we using it in computer vision?
- Datasets are still important — they give us information to use for training and evaluations
  - Combine datasets with psychophysics-based evaluation

Thank you!