The Impact of the Open Set Recognition Problem on Deep Learning

Walter J. Scheirer

Computer Vision Research Laboratory,
Department of Computer Science and Engineering

UNIVERSITY OF NOTRE DAME

Benchmarks in computer vision

Assume we have examples from all classes:

Places2 Data Set (part of ILSVRC 2016)

Out in the real world…

Detect the cars in this image

while rejecting the trees, signs, telephone poles…

M. Milford, W.J. Scheirer, E. Vig, A. Glover, O. Baumann, J. Mattingley, and D.D. Cox, “Condition Invariant Top-Down Visual Place Recognition,” ICRA 2014.

Open Set Recognition: incomplete knowledge of the world is present at training time, and unknown classes can be submitted to an algorithm during its operation.

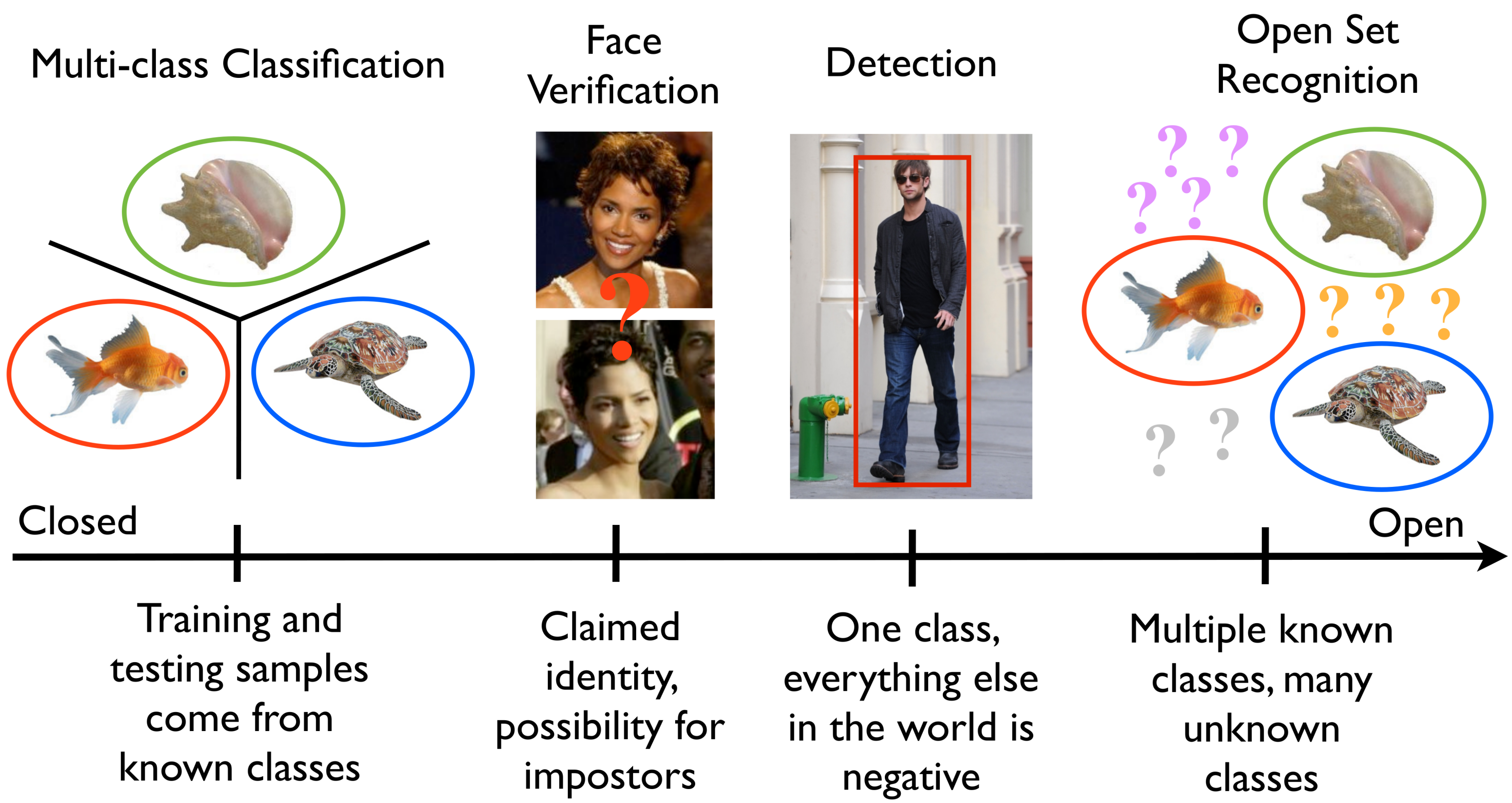
“There are known knowns…”

known classes: the classes with distinctly labeled positive training examples (also serving as negative examples for other known classes)

known unknown classes: labeled negative examples, not necessarily grouped into meaningful categories

unknown unknown classes: classes unseen in training

Vision problems in order of "openness"



W. J. Scheirer, A. Rocha, A. Sapkota, and T. Boult, "Towards Open Set Recognition," IEEE T-PAMI, 35(7) July 2013.

Fundamental multi-class recognition problem

$$\underset{f}{\operatorname{argmin}} \left\{ R_{\mathcal{I}}(f) := \int_{\mathbb{R}^d \times \mathbb{N}} L(x, y, f(x)) P(x, y) \right\}$$

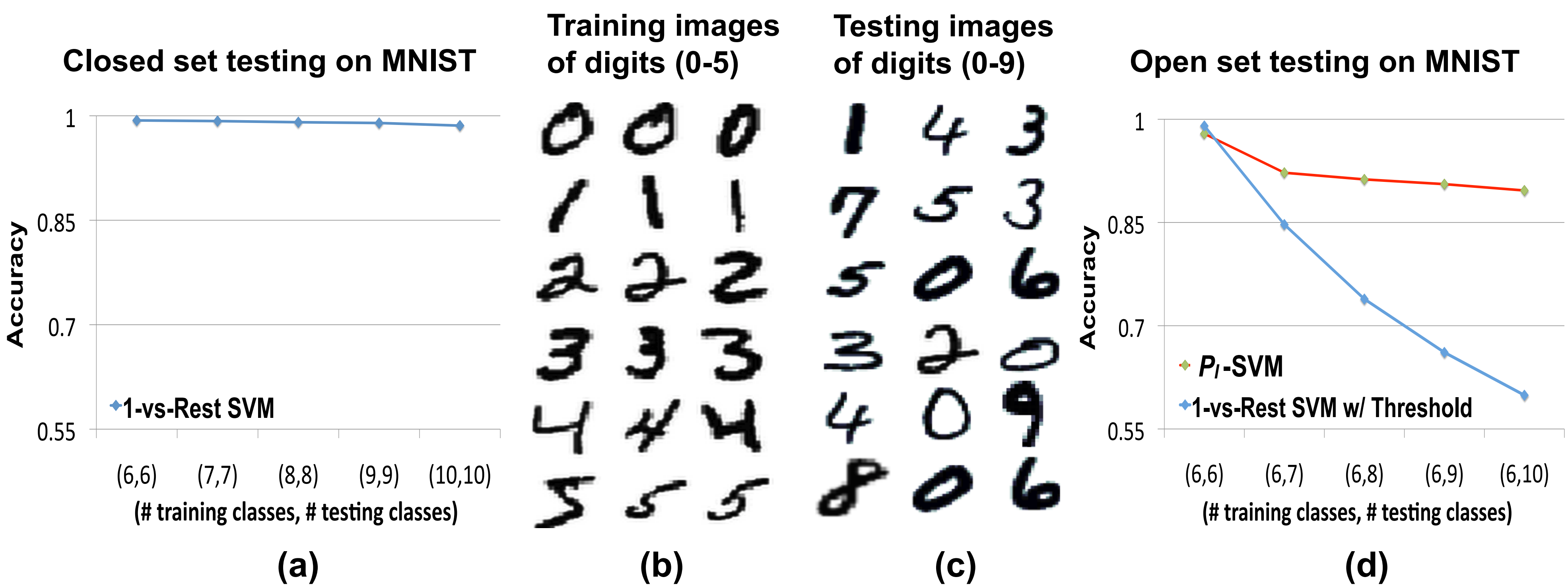
Ideal Risk Loss Function Joint Distribution

Undefined for open set recognition!

A. Smola, "Learning with Kernels," Ph.D. dissertation, Technische Universität Berlin, Berlin, Germany, November 1998.

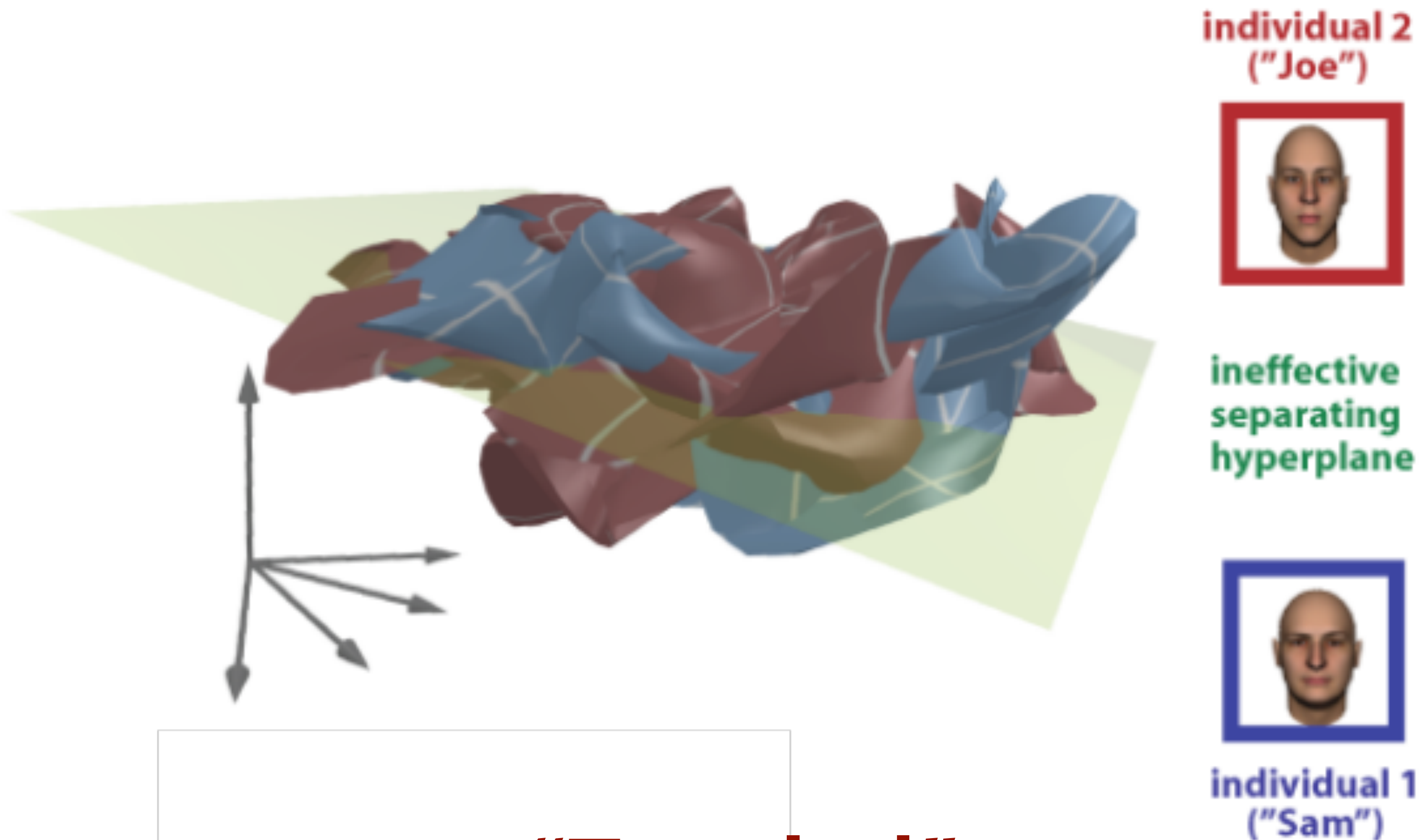
Open Space

Open Set MNIST Benchmark



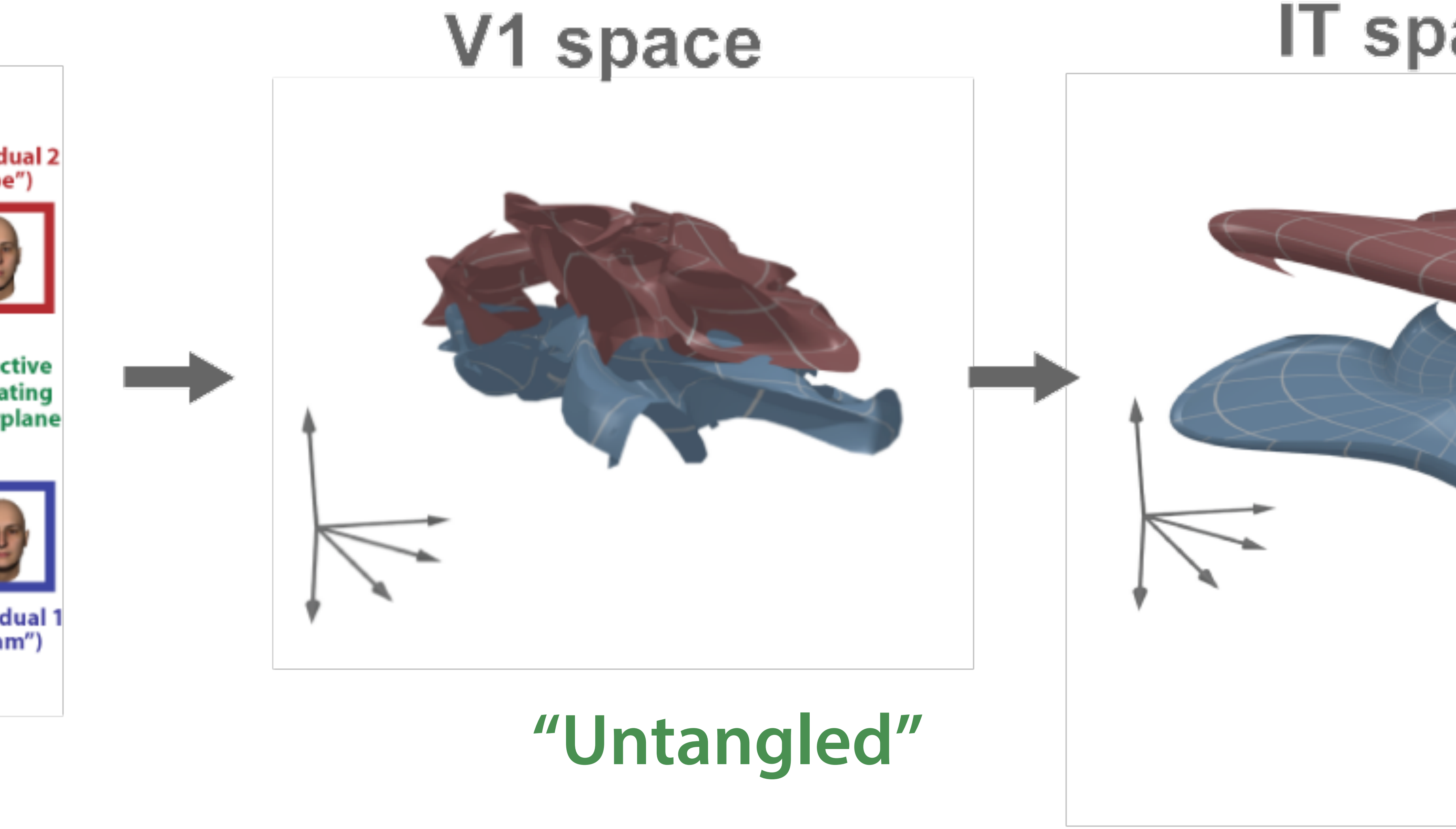
L. P. Jain, W. J. Scheirer, and T. Boult, "Multi-Class Open Set Recognition Using Probability of Inclusion," ECCV 2014.

pixel space

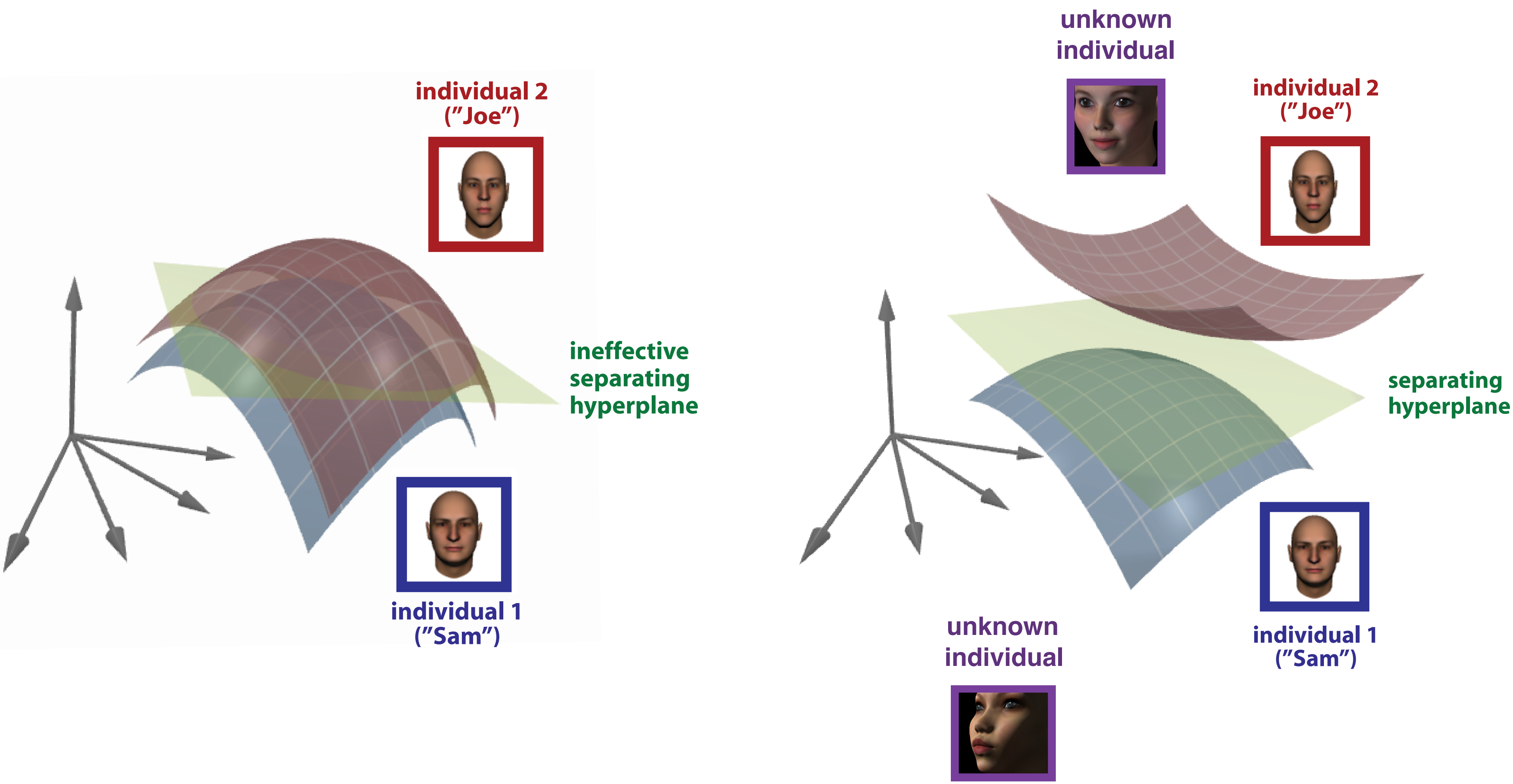


"Tangled"

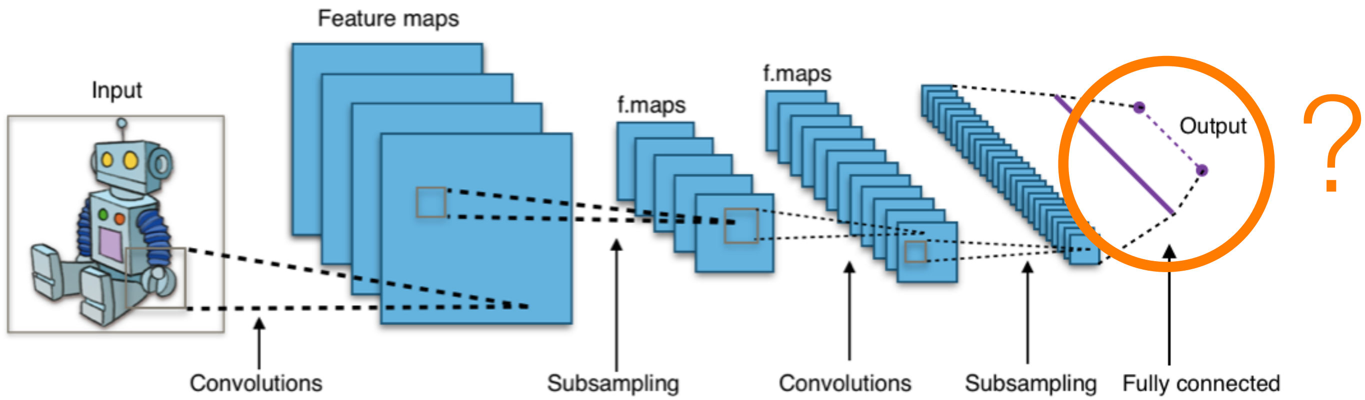
Adapted from an image by D. D. Cox


V1 space
“Untangled”

Linear separation of CNN feature representations



Read-out layer



Softmax

$$P(y=j|\mathbf{x}) = \frac{e^{\mathbf{v_j}(\mathbf{x})}}{\sum_{i=1}^{N} e^{\mathbf{v_i}(\mathbf{x})}}$$

Sum over all of the classes

Linear SVM

$$min\frac{1}{2}||w||^2$$

subject to

$$y_i(w * x_i + b) \geq 1, \forall_i$$

Known positive or negative sample

Cosine Similarity

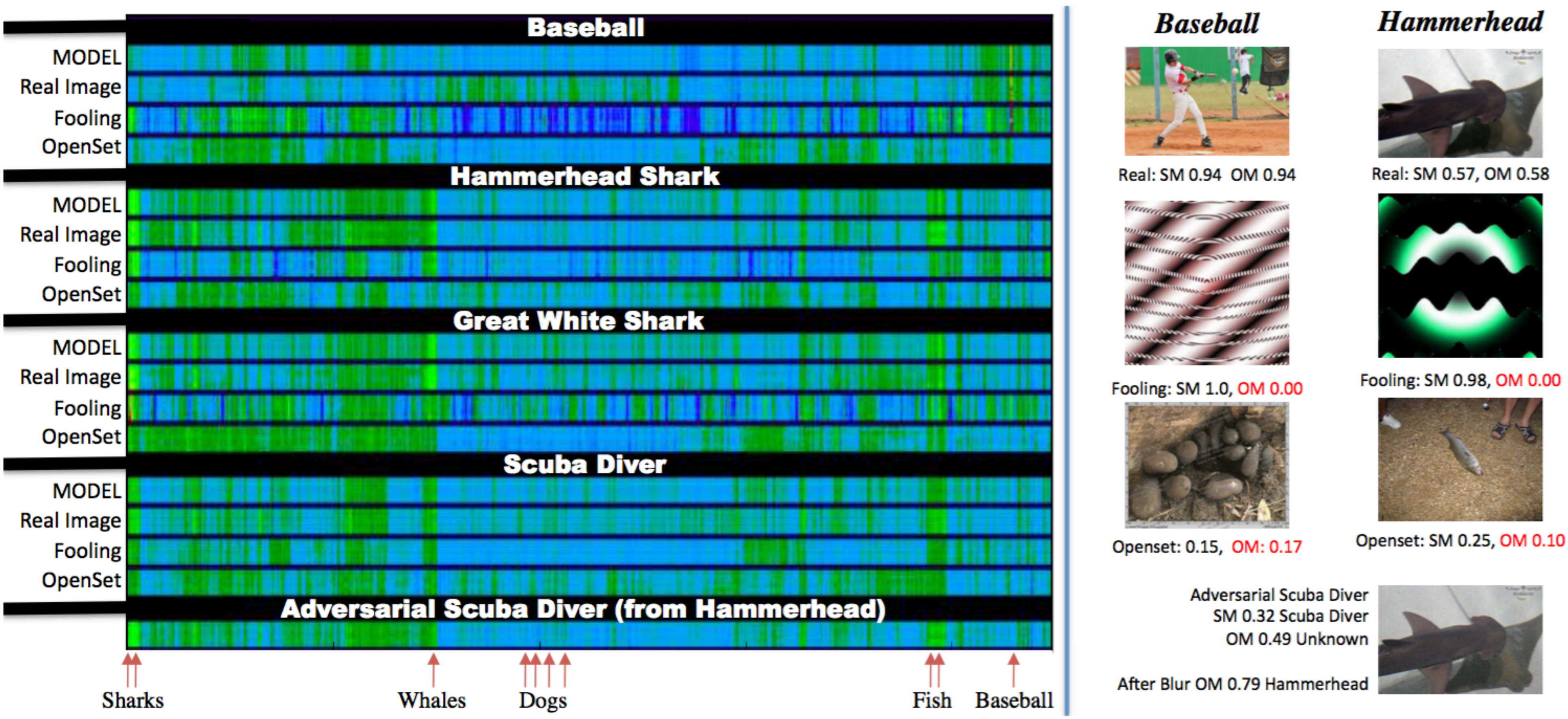
$$\frac{\mathrm{A} \cdot \mathrm{B}}{||\mathrm{A}||\ ||\mathrm{B}||} < \delta$$

Threshold determined empirically via known pairs

Evolving images to match CNN classes

A. Nguyen, J. Yosinski, and J. Clune, “Deep Neural Networks are Easily Fooled,” CVPR 2015.

A step towards a fix: OpenMax



A. Bendale and T. Boult, "Towards Open Set Deep Networks," CVPR 2016.

How does OpenMax work?

Require: Activation vector for $\mathbf{v}(\mathbf{x}) = v_1(x), \ldots, v_N(x)$
Require: **means** μ_j and libMR models $\rho_j = (\tau_i, \lambda_i, \kappa_i)$
Require: α, the numer of "top" classes to revise

Let $s(i) = \text{argsort}(v_j(x))$; Let $\omega_j = 1$
for $i = 1, \ldots, \alpha$ **do**
$$\omega_{s(i)}(x) = 1 - \frac{\alpha - i}{\alpha} e^{-\left(\frac{\|x - \tau_{s(i)}\|}{\lambda_{s(i)}}\right)^{\kappa_{s(i)}}}$$
end for
Revise activation vector $\hat{v}(x) = \mathbf{v}(\mathbf{x}) \circ \omega(\mathbf{x})$
Define $\hat{v}_0(x) = \sum_i v_i(x)(1 - \omega_i(x))$.
$$\hat{P}(y = j|\mathbf{x}) = \frac{e^{\hat{\mathbf{v}}_\mathbf{j}(\mathbf{x})}}{\sum_{i=0}^{N} e^{\hat{\mathbf{v}}_\mathbf{i}(\mathbf{x})}}$$
Let $y^* = \text{argmax}_j\, P(y = j|\mathbf{x})$
Reject input if $y^* == 0$ or $P(y = y^*|\mathbf{x}) < \epsilon$

Apply probability models derived from statistical extreme value theory to calculate class weights

Use weights to adjust activation

Apply rejection threshold

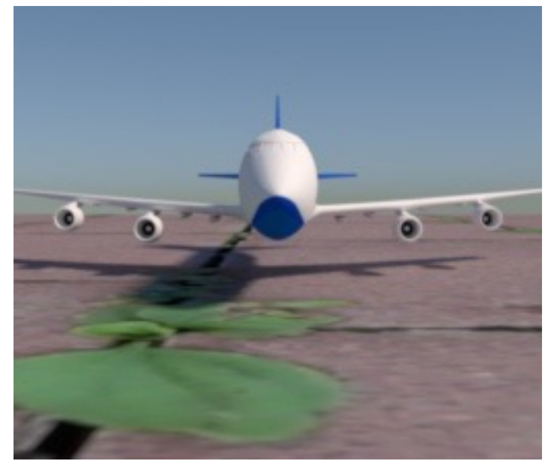
But you don’t have to use tricky manipulations

GoogleNet Output

Label: Hammerhead Shark

Label: Blow Dryer

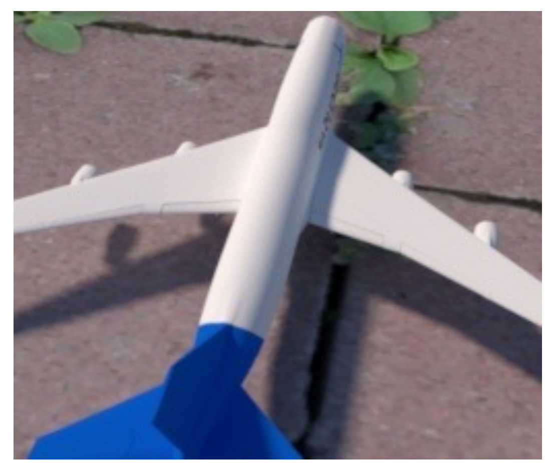
Label: Mosque



Label: Syringe

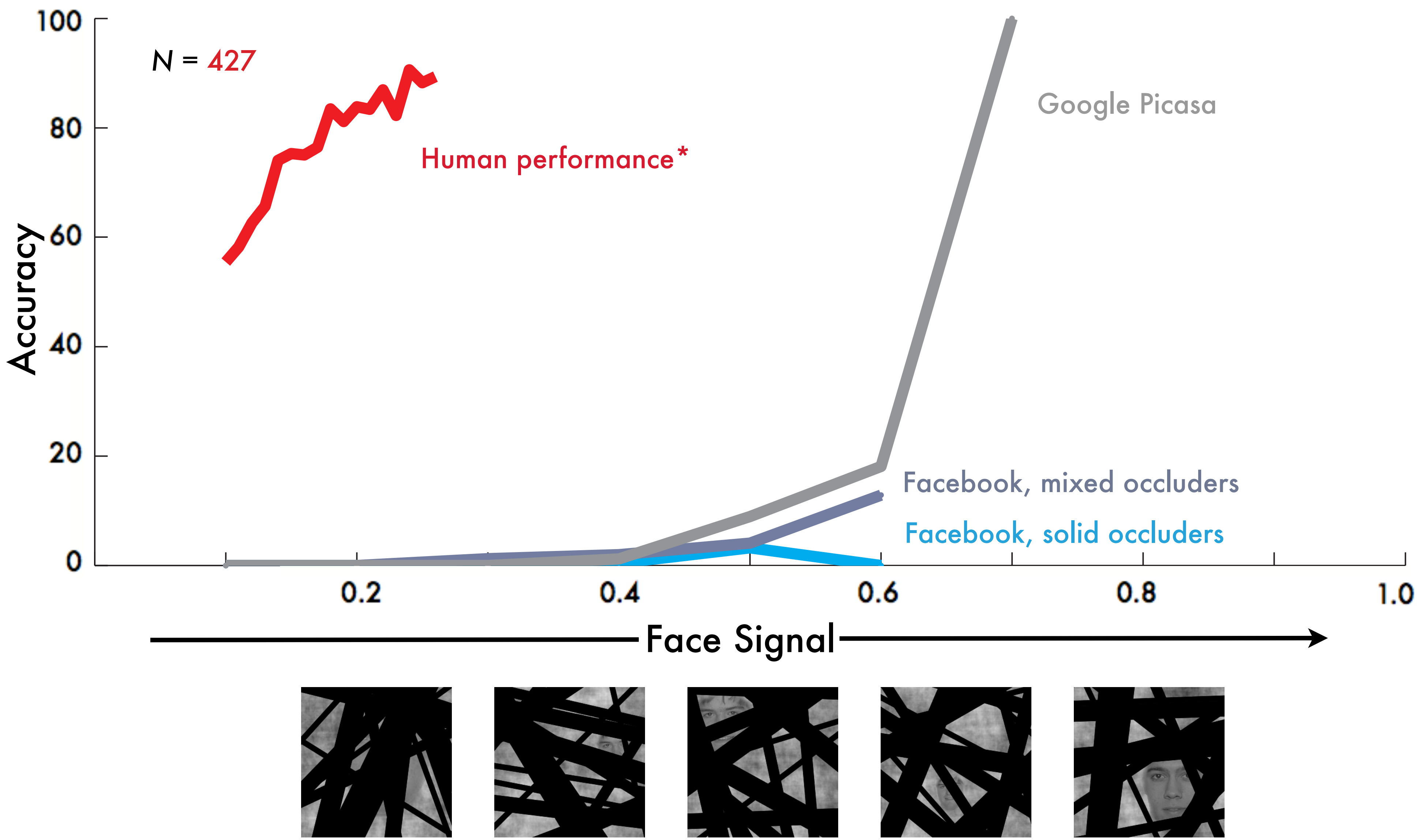
Label: Trimaran

Label: Missile



Are performance measures misleading us?

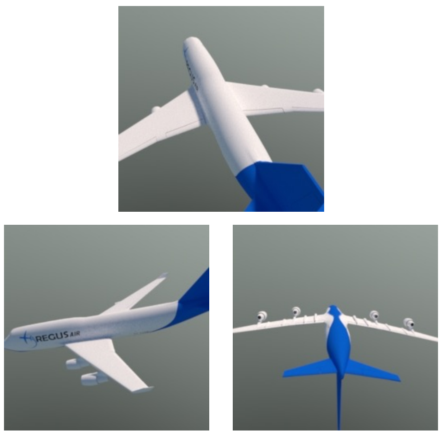
Psychophysics on the Model



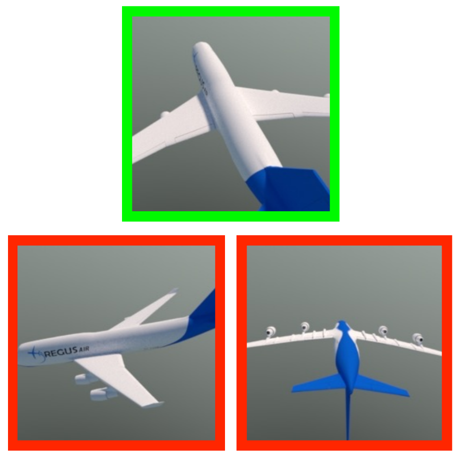
W.J. Scheirer, S. Anthony, K. Nakayama, and D. D. Cox, “Perceptual Annotation: Measuring Human Vision to Improve Computer Vision,” IEEE T-PAMI, 36(8) August 2014.

Psychophysics pipeline

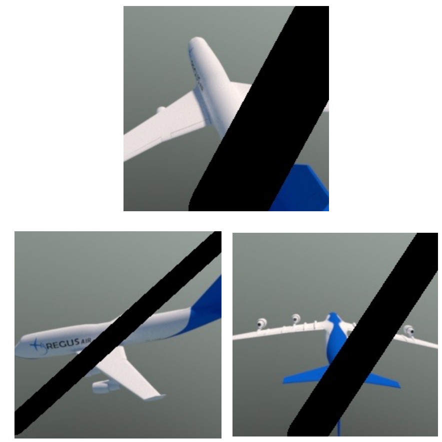
1. Render Class Canonical View (CCV) Candidates



2. CCV Classifier



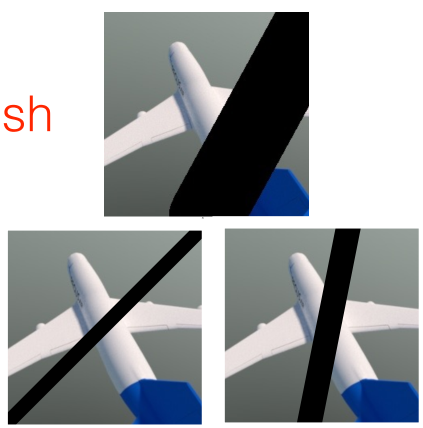
3. Manipulate Chosen Variable



4. Classify Images

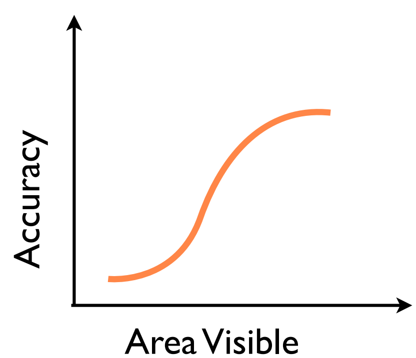
Fish

Plane

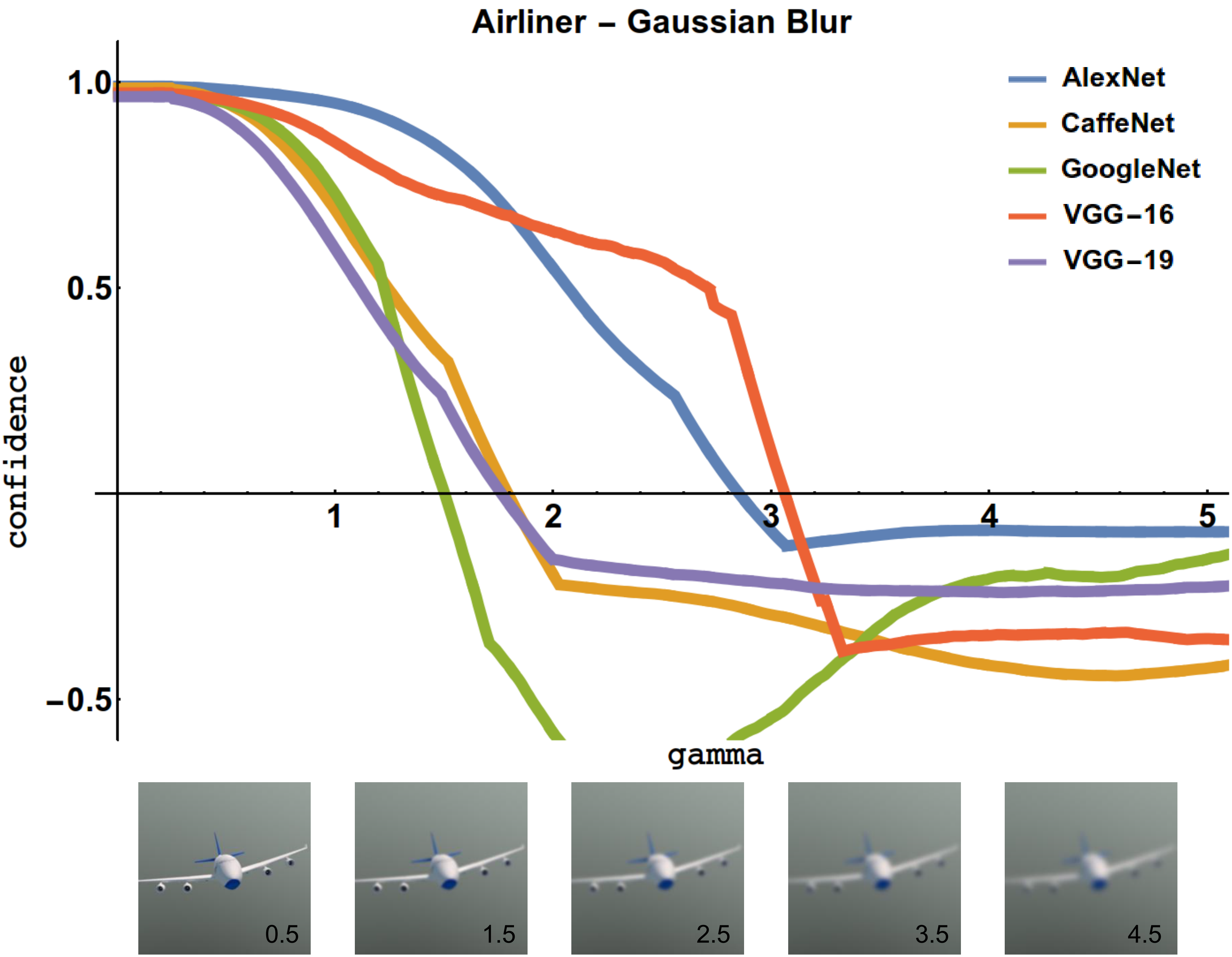


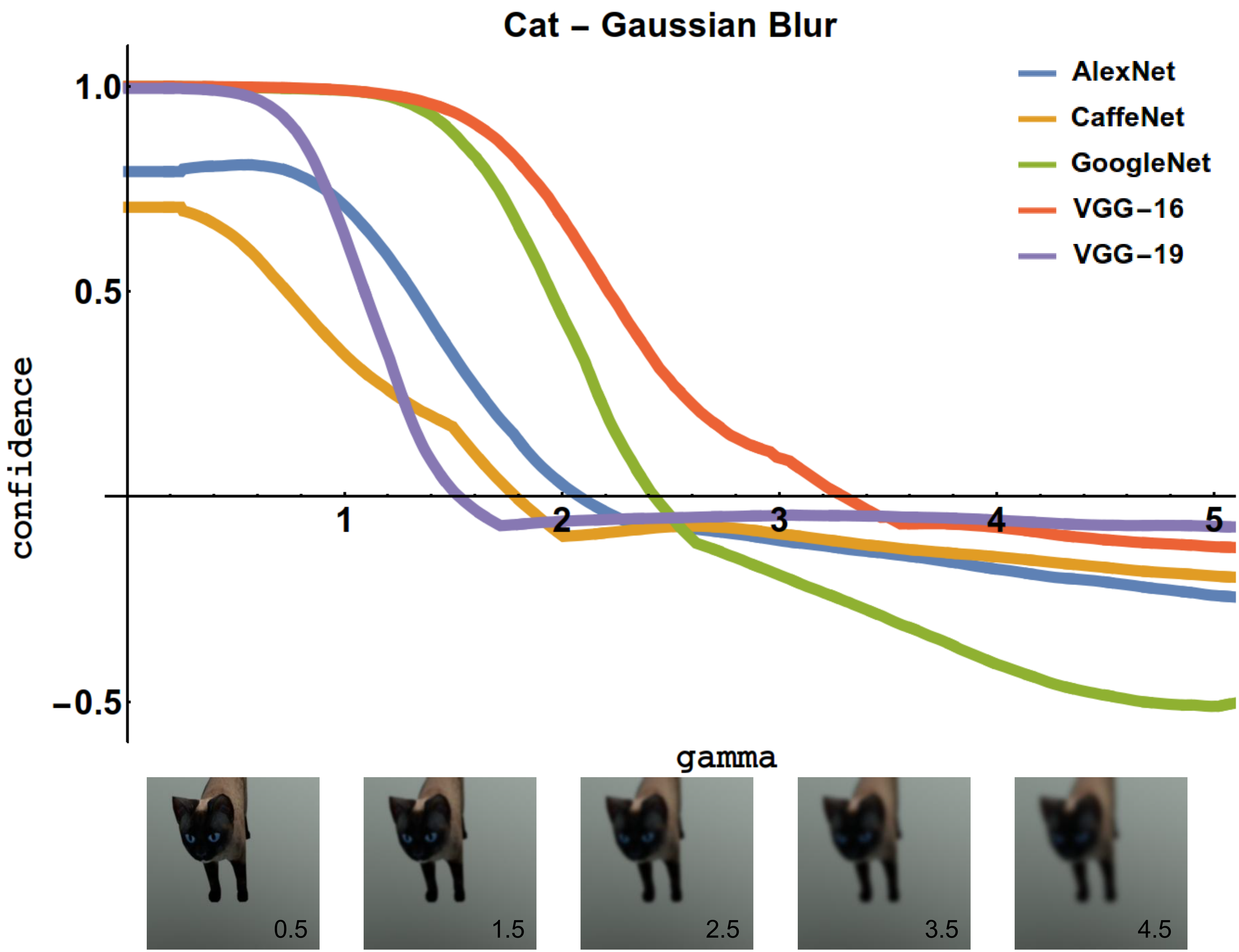
Skyscraper

5. Generate Psychometric Curve

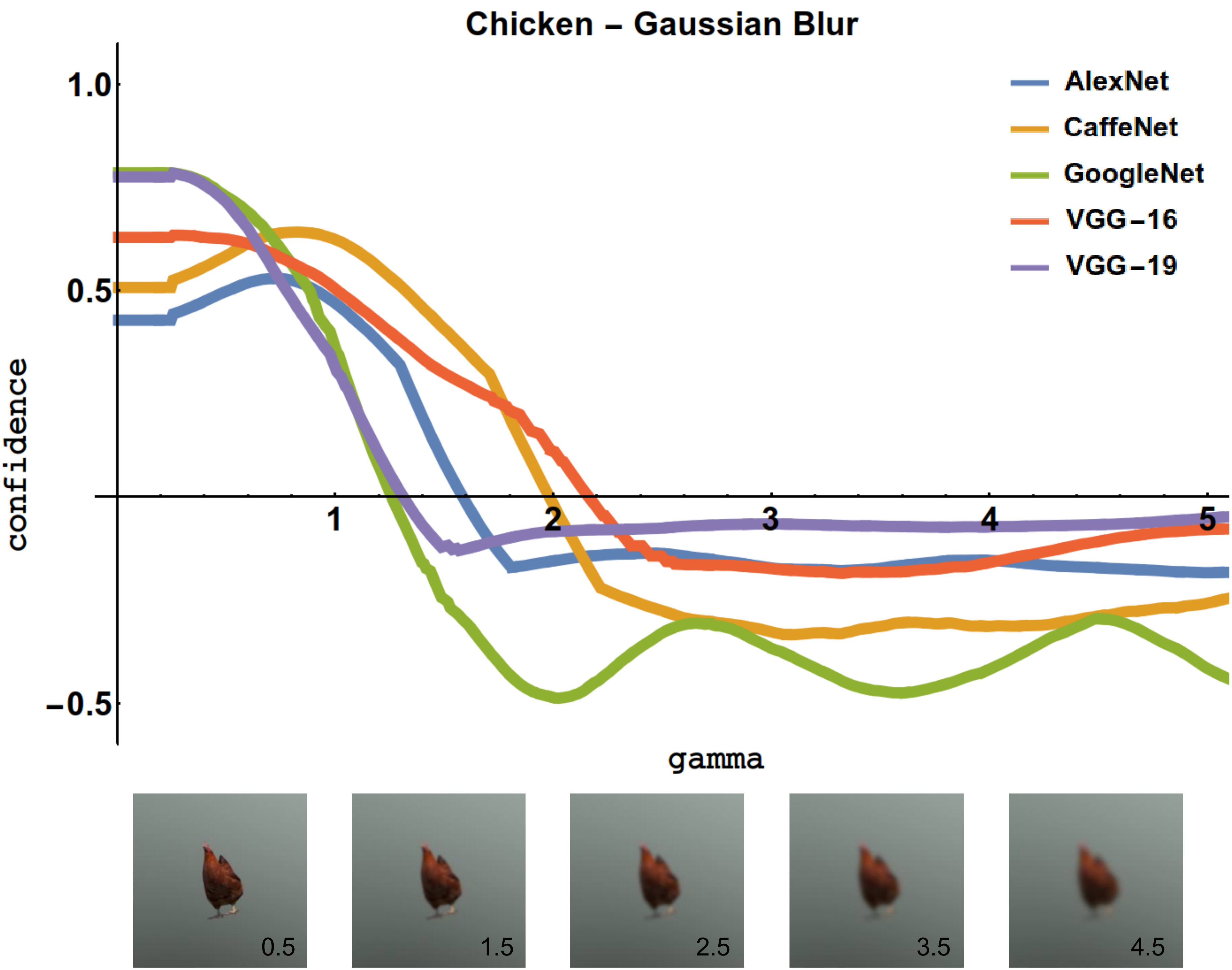


Brandon RichardWebster


Airliner – Gaussian Blur
AlexNet
CaffeNet
GoogleNet
VGG–16
VGG–19
confidence
1.0
0.5
–0.5
1
2
3
4
5
gamma
0.5
1.5
2.5
3.5
4.5



Cat – Gaussian Blur
AlexNet
CaffeNet
GoogleNet
VGG-16
VGG-19
1.0
0.5
−0.5
confidence
1
2
3
4
5
gamma
0.5
1.5
2.5
3.5
4.5



Chicken – Gaussian Blur
1.0
0.5
−0.5
confidence
1
2
3
4
5
gamma
AlexNet
CaffeNet
GoogleNet
VGG−16
VGG−19
0.5
1.5
2.5
3.5
4.5

Practical implications

FUTURE TENSE THE CITIZEN'S GUIDE TO THE FUTURE. MARCH 1 2016 12:23 PM

FROM SLATE, NEW AMERICA, AND ASU

The Trollable Self-Driving Car

Humans are pretty good at guessing what others on the road will do. Driverless cars are not—and that can be exploited.

770 234 155

By Samuel English Anthony

http://goo.gl/78fglb

Thank you!

Read more: www.wjscheirer.com