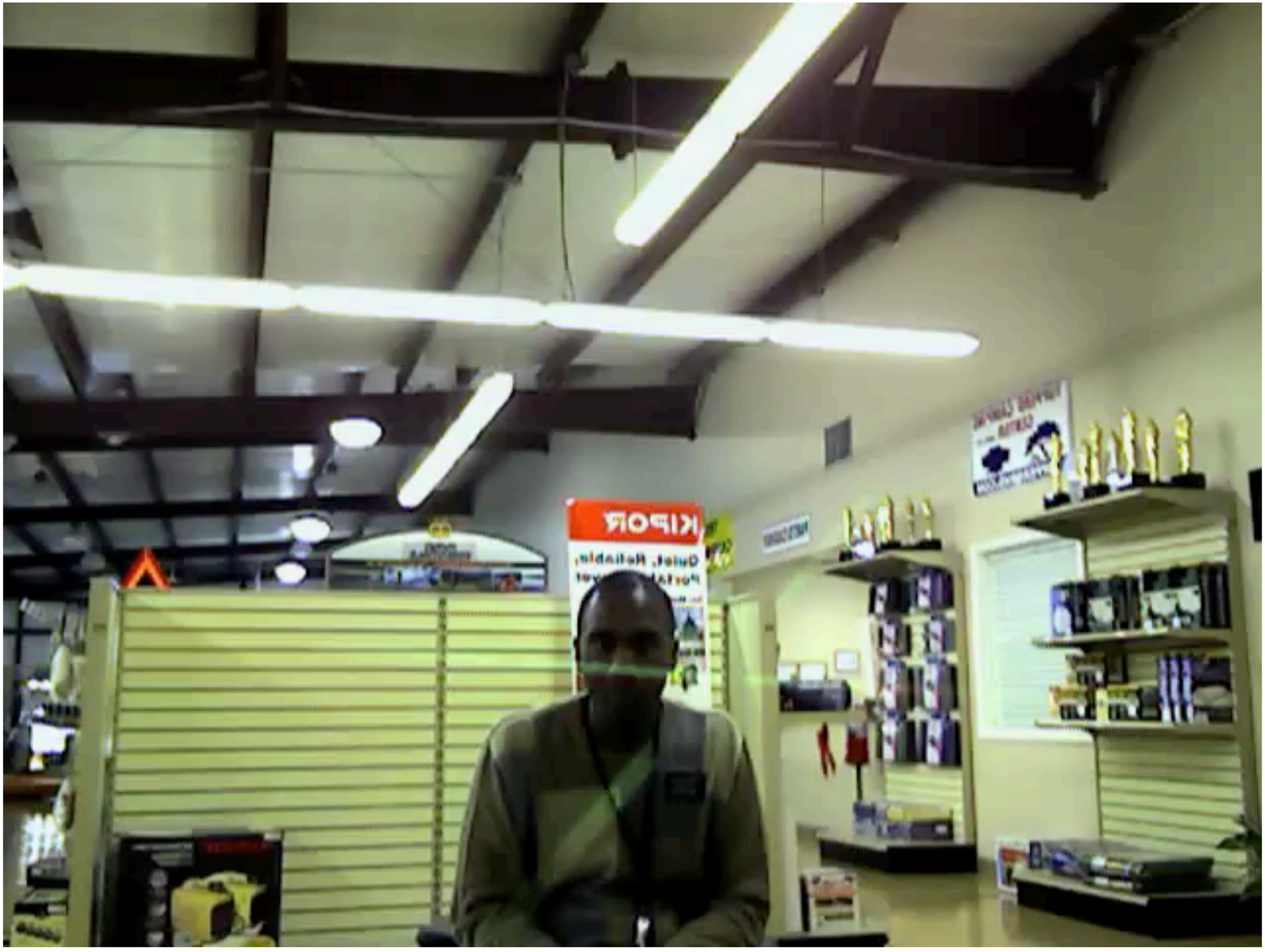
Bias and Ethics

Computer Vision
Fall 2018
Columbia University



HP laptop computer in 2008, video went viral

Why would computer vision be biased?

Is it just the training data?

Film crash course

Photons

Base

Light sensitive

Base

Light sensitive layer is made up of silver halides. When exposed to light, it become dark, depending on intensity of light

Film crash course

Photons

Base

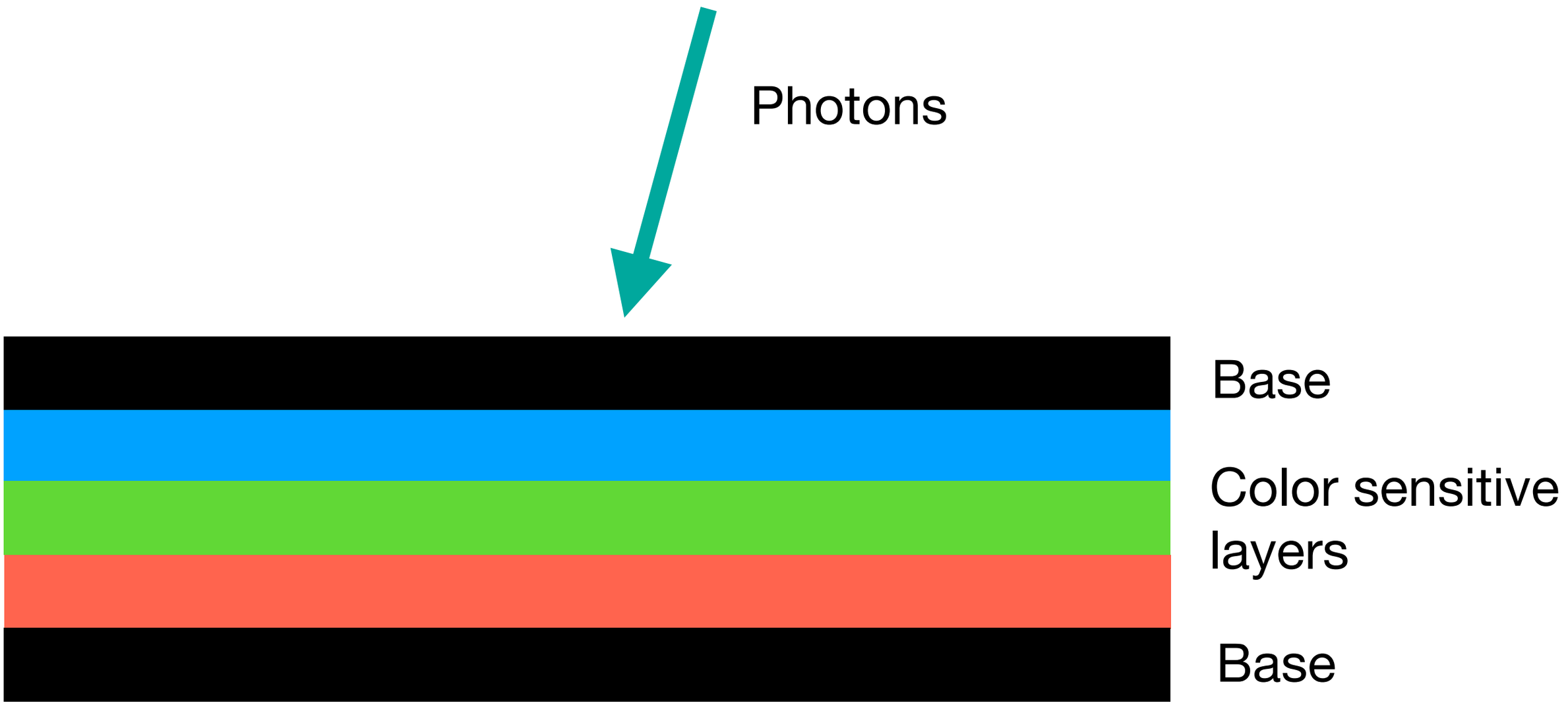
Light sensitive

Base

Film development wipes away undeveloped silver halides, resulting in the negative

Film crash course

Photons



Complex and delicate procedure to mix color negatives into final image

Shirley Card

- Film companies distributed reference cards so labs could test their color reproduction

- If lab developed a Shirley card, and photo looked different from reference, they re-calibrated

Some Shirley Cards

Historical bias in film

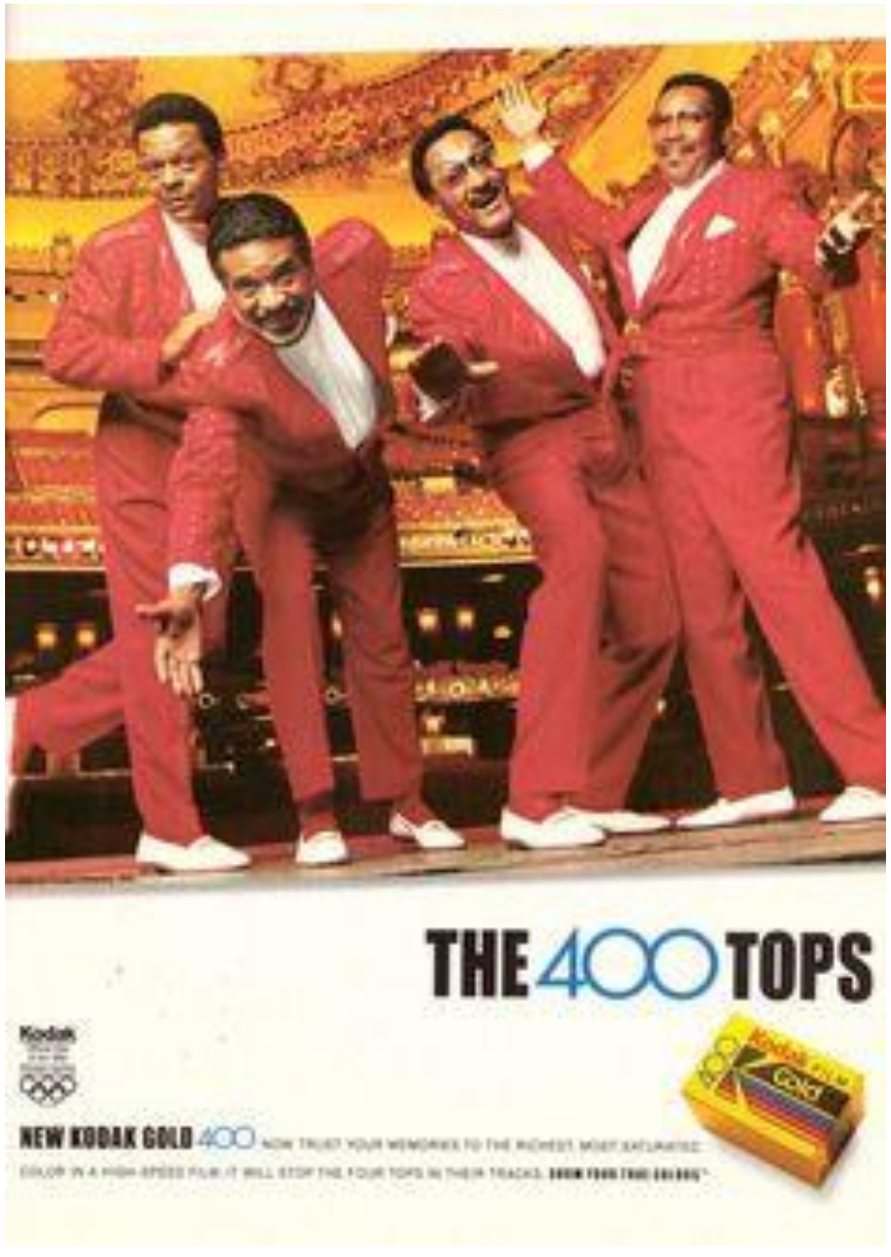
How to fix bias?

Create better reference image? But that did not happen until 1995!

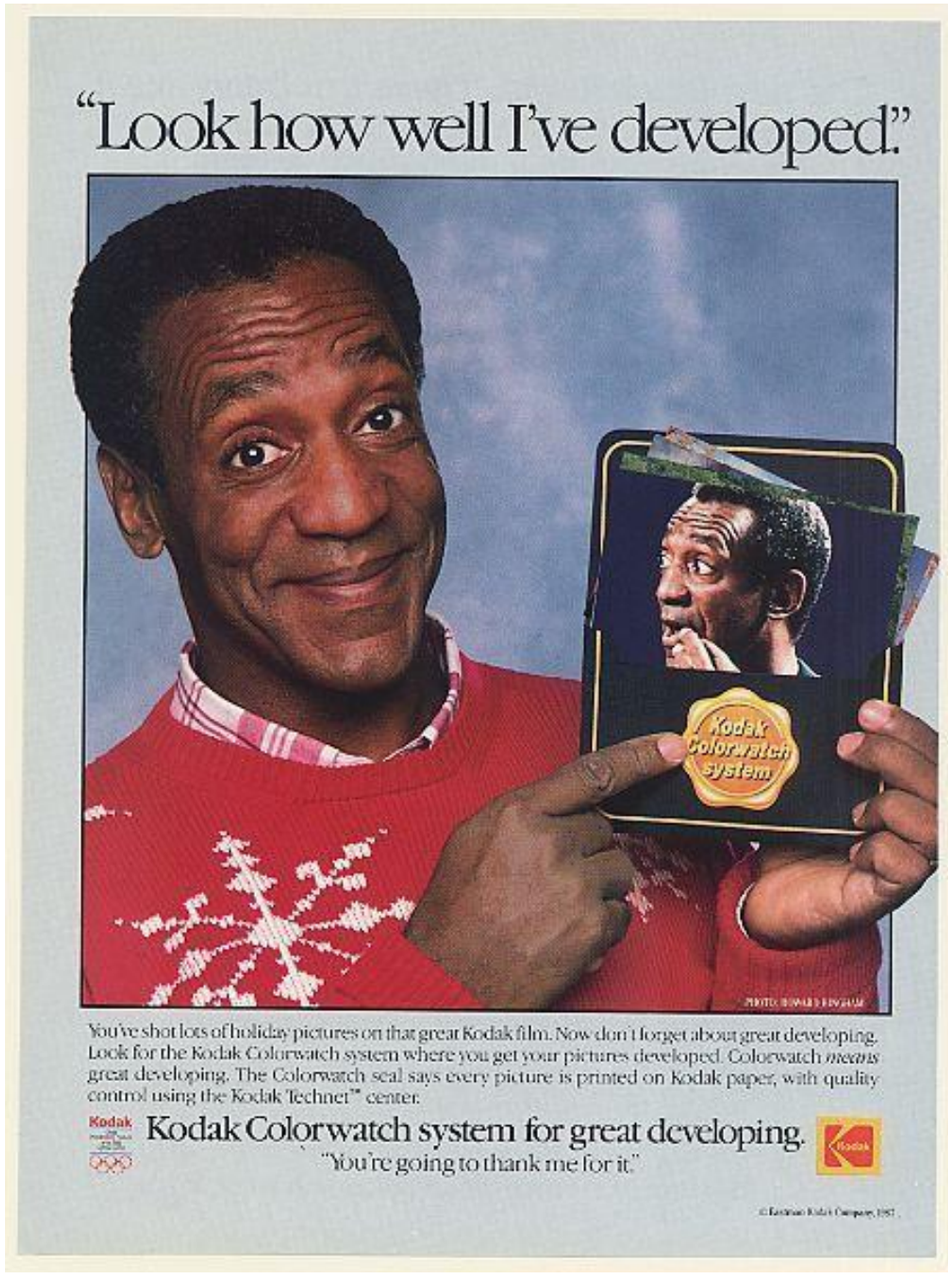
What caused change

- 1978: Jean-Luc Godard refuses to use Kodachrome film in Mozambique.
- 1980s: Chocolate, furniture industry wanted to sell advertisements

1980s – adverts



The Four Tops!



Bill Cosby!

Some other issues here too : (

Slide credit: James Tompkin

Evolution of Vision Datasets

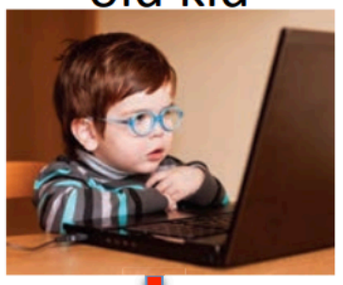
COIL-20

Caltech 101

MNIST (1998)

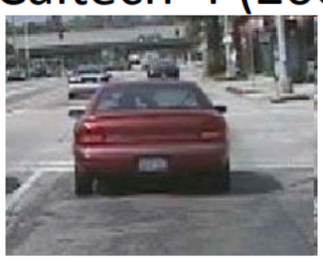
IMAGENET (2009)

2 year old kid



10^3 10^4 10^5 10^6 10^7 10^8 10^9 # images

Caltech-4 (2003)



PASCAL (2005)

Slide credit: Aude Oliva

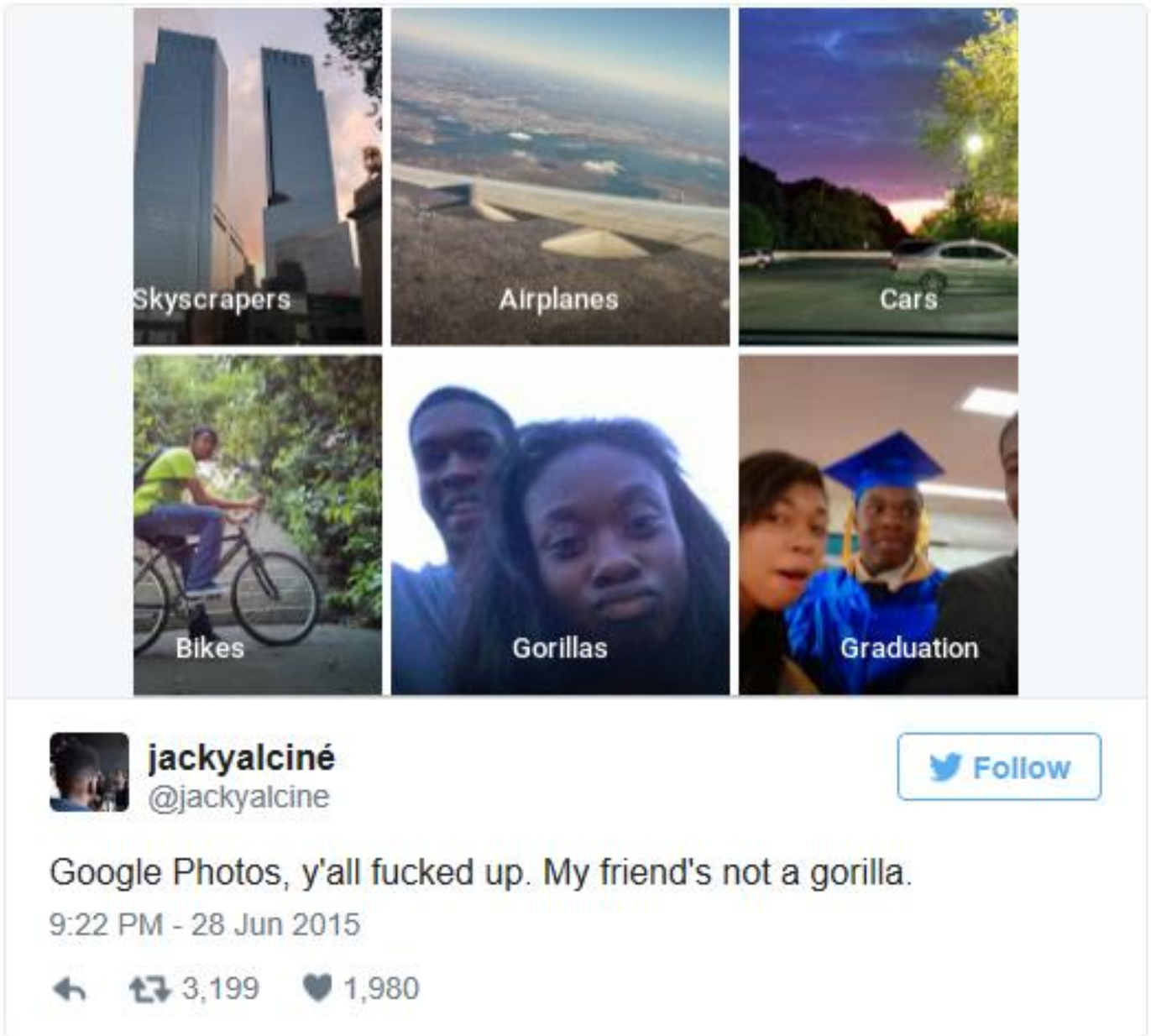
Lena

- Early reference image used in image processing
- To compare results, papers sometimes include their result on this same picture
- However, the image is a crop from a photo originally in an adult magazine

Microsoft Tay Chatbot

- Computer program trained to maximize likes on twitter
- Marketed as “The AI with zero chill”
- Sensitive topics black-listed with canned responses
- Learned to retweet offensive messages after 16 hours, shut down

Google Photos (2015)



[Jacky Alciné]

Slide credit: James Tompkin

Facial bias

The New York Times

Facial Recognition Is Accurate, if You're a White Guy

By Steve Lohr

Feb. 9, 2018

Facial recognition technology is improving by leaps and bounds. Some commercial software can now tell the gender of a person in a photograph.

Facial bias

Gender was misidentified in **up to 1 percent of lighter-skinned males** in a set of 385 photos.

Facial bias

Gender was misidentified in **up to 12 percent of darker-skinned males** in a set of 318 photos.

Facial bias

Gender was misidentified in **up to 7 percent of lighter-skinned females** in a set of 296 photos.

Facial bias

Gender was misidentified in **35 percent of darker-skinned females** in a set of 271 photos.

Facial bias

“One widely used facial-recognition data set was estimated to be more than 75 percent male and more than 80 percent white, according to another research study.”

Steve Lohr for *New York Times*

What are modern day Shirley cards?

Joy Buolamwini (lead author)

The Washington Post
Democracy Dies in Darkness

The Switch

Amazon is selling facial recognition to law enforcement — for a fistful of dollars

The Washington Post
Democracy Dies in Darkness

The Switch

Amazon is selling facial recognition to law enforcement — for a fistful of dollars

ACLU DONATE

Amazon's Face Recognition Falsely Matched 28 Members of Congress With Mugshots

By Jacob Snow, Technology & Civil Liberties Attorney, ACLU of Northern California

JULY 26, 2018 | 8:00 AM

Source: ACLU

The Telegraph

Chinese businesswoman accused of jaywalking after AI camera spots her face on an advert

Criminality

- Wu and Zhang, *Automated Inference on Criminality using Face Images*, on arXiv 2016
- https://arxiv.org/abs/1611.04135

(a) Three samples in criminal ID photo set S_c.

(b) Three samples in non-criminal ID photo set S_n

Figure 1. Sample ID photos in our data set.

Slide figures from paper

“Unlike a human examiner/judge, a computer vision algorithm or classifier has absolutely no subjective baggages, having no emotions, no biases whatsoever due to past experience, race, religion, political doctrine, gender, age, etc., no mental fatigue, no preconditioning of a bad sleep or meal. The automated inference on criminality eliminates the variable of meta-accuracy (the competence of the human judge/examiner) all together.”

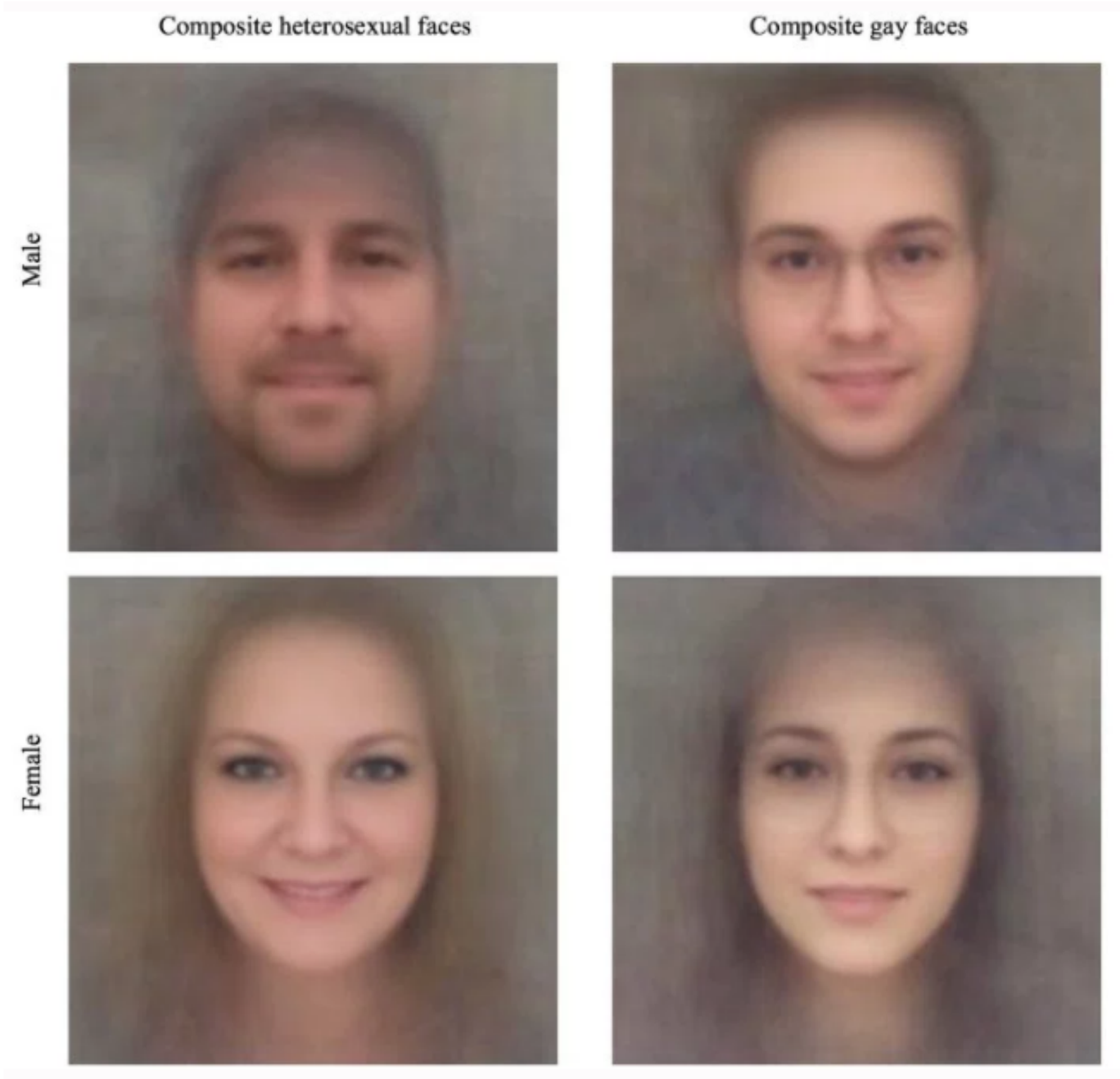
Slide credit: James Tompkin

Attempting to predict sexual orientation from facial photograph

- Yilun Wang, Michal Kosinski in 2017 tried to predict sexual orientation from a photograph

- They downloaded photographs from US dating websites

- Find some correlation, and claimed it was due to facial structure

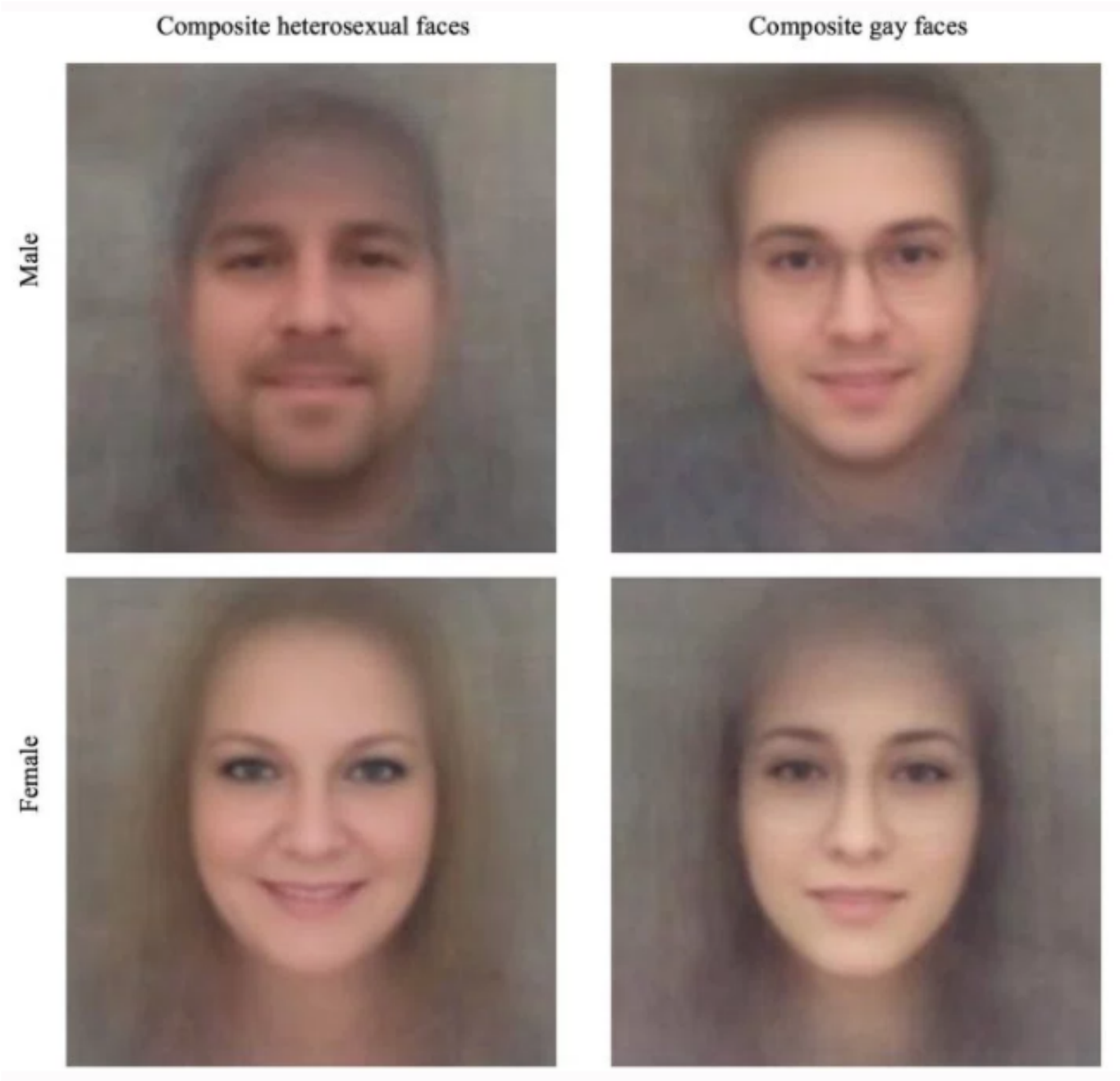


https://medium.com/@blaisea/do-algorithms-reveal-sexual-orientation-or-just-expose-our-stereotypes-d998fafdf477

Attempting to predict sexual orientation from facial photograph

- However, this was due to superficial biases from the dating website, such as angle of photograph, glasses, and makeup

- When you control for these variables, performance drops to chance



https://medium.com/@blaisea/do-algorithms-reveal-sexual-orientation-or-just-expose-our-stereotypes-d998fafdf477

Excellent critique on these issues: https://medium.com/@blaisea/do-algorithms-reveal-sexual-orientation-or-just-expose-our-stereotypes-d998fafdf477

Unbiased Look at Dataset Bias

Torralba and Efros, CVPR 2011

“The authors would like to thank the Eyjafjallajokull volcano as well as the wonderful kirs at the Buvette in Jardin du Luxembourg for the motivation (former) and the inspiration (latter) to write this paper.“

Next few slide contents are from the paper

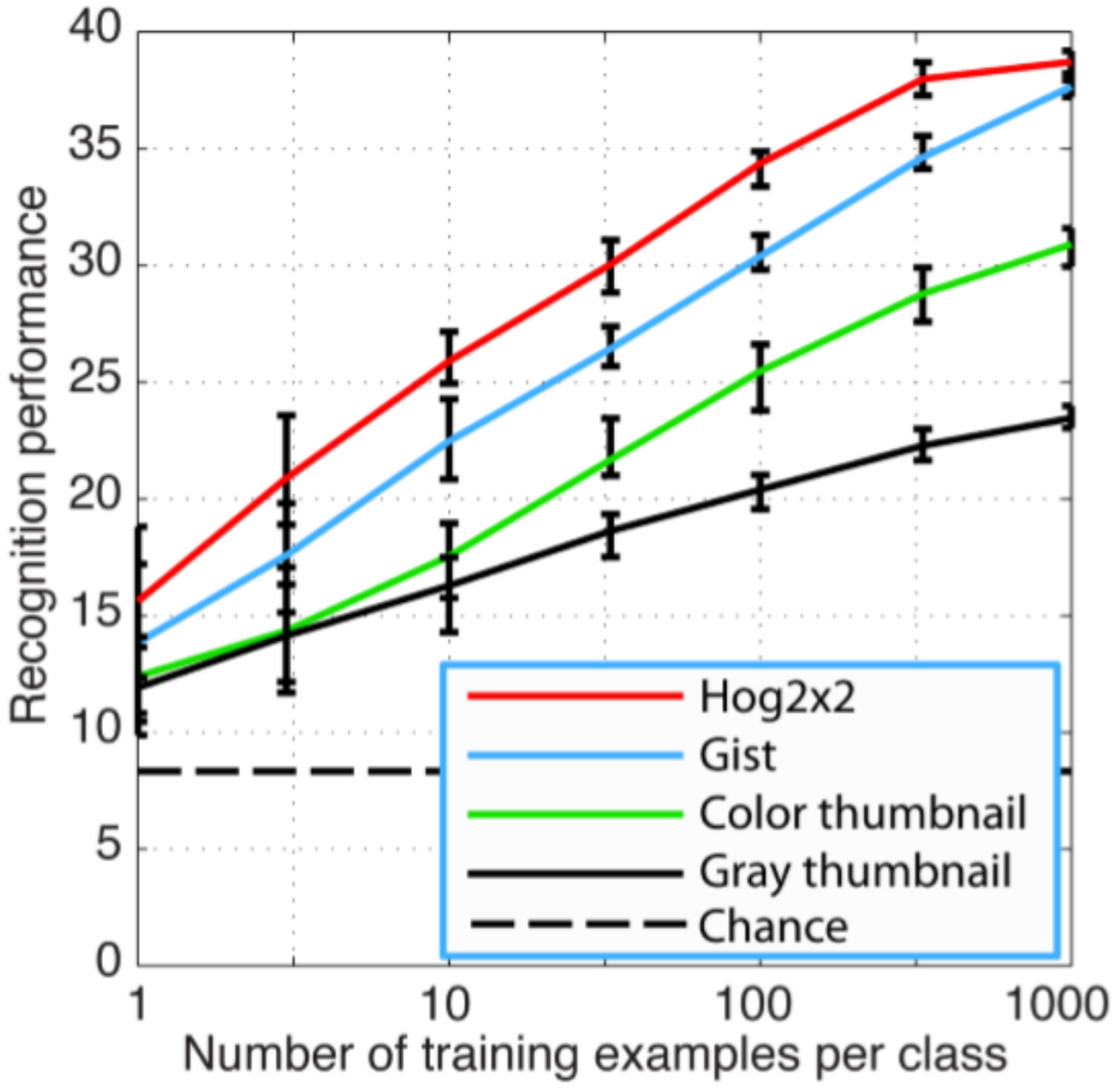
Caltech101 ____ Tiny ____ LabelMe ____ 15 Scenes ____

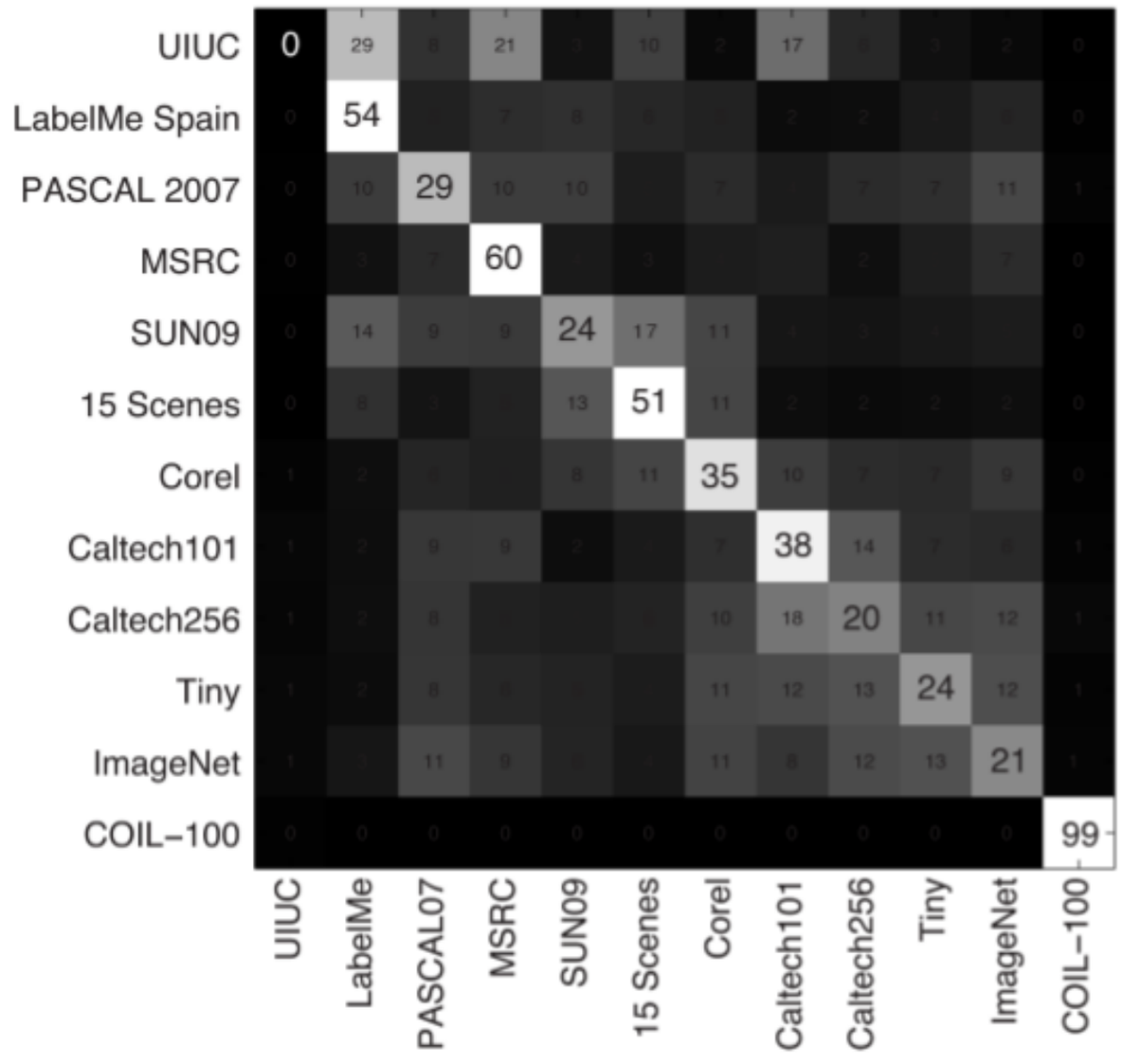
MSRC ____ Corel ____ COIL-100 ____ Caltech256 ____

UIUC ____ PASCAL 07 ____ ImageNet ____ SUN09 ____

Figure 1. Name That Dataset: Given three images from twelve popular object recognition datasets, can you match the images with the dataset? (answer key below)

CV plays name that dataset!





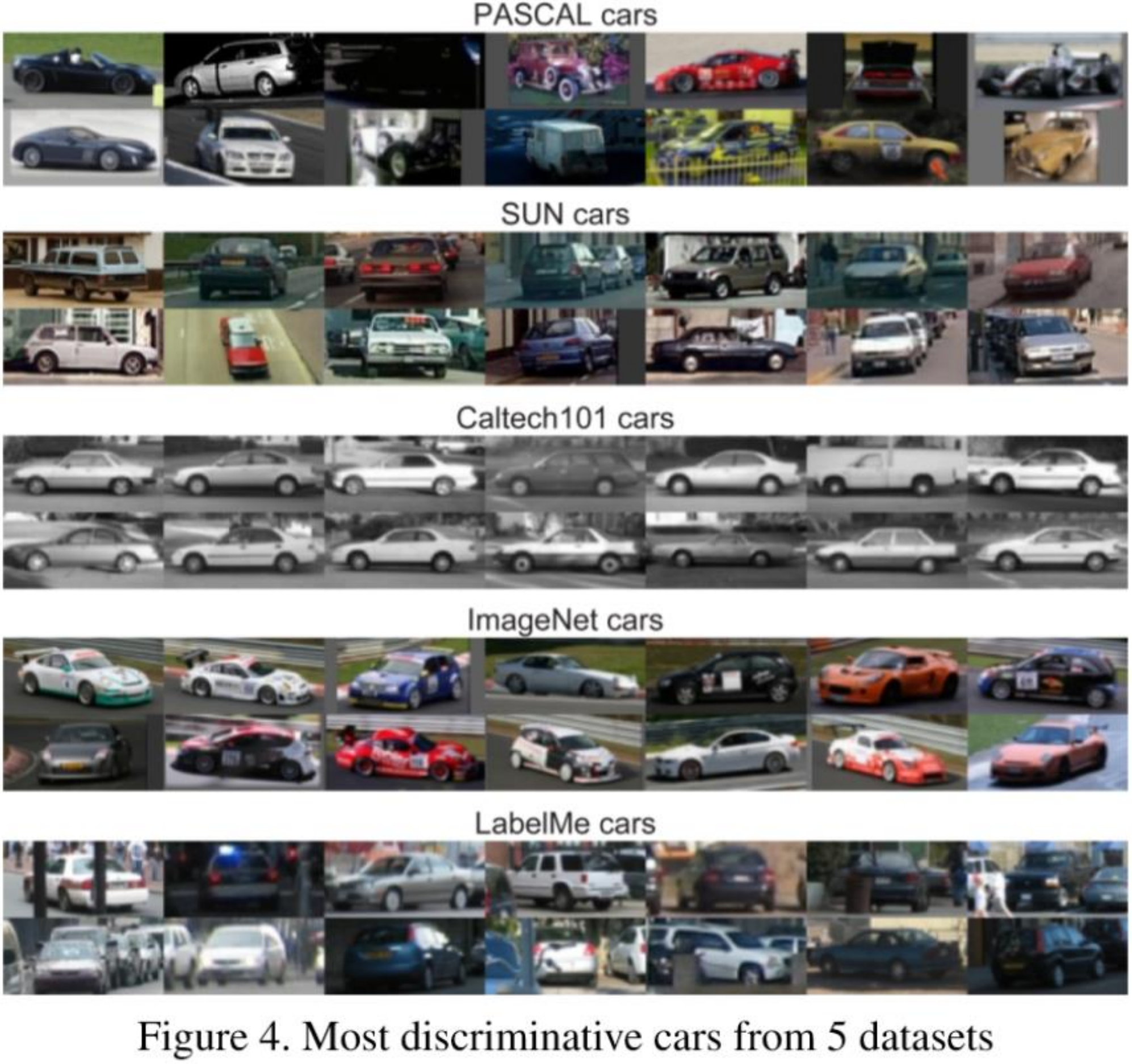
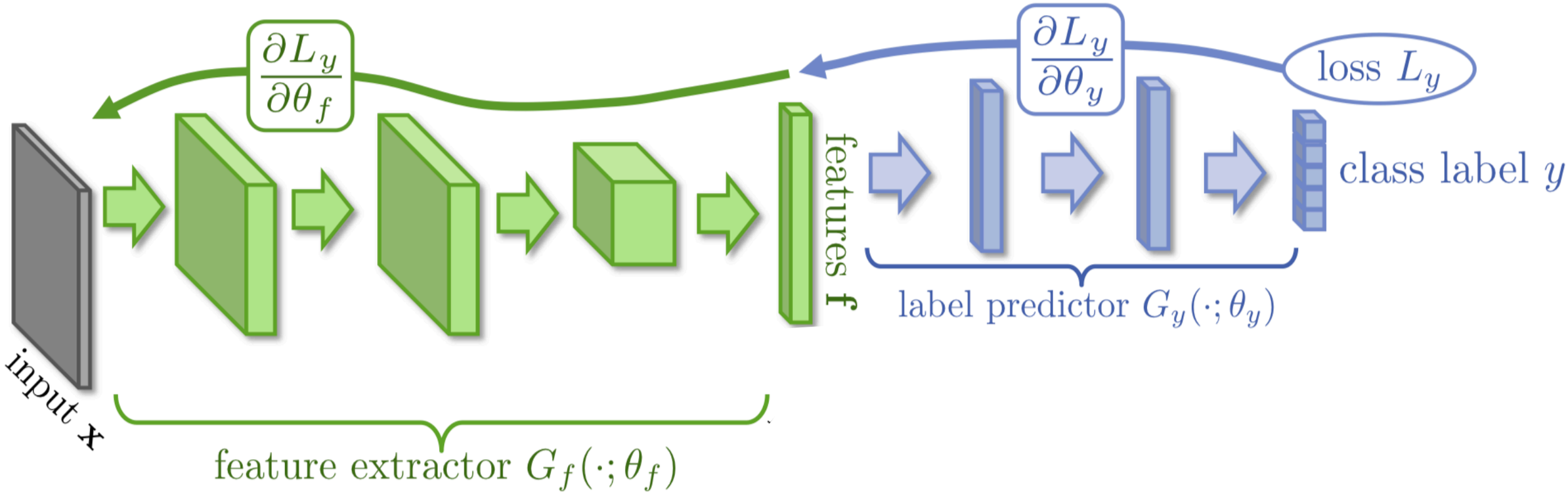


Figure 4. Most discriminative cars from 5 datasets

task	Train on: \ Test on:	SUN09	LabelMe	PASCAL	ImageNet	Caltech101	MSRC	Self	Mean others	Percent drop
"person" detection	SUN09	**69.6**	56.8	37.9	45.7	52.1	72.7	69.6	53.0	**24%**
	LabelMe	58.9	**66.6**	38.4	43.1	57.9	68.9	66.6	53.4	**20%**
	PASCAL	56.0	55.6	**56.3**	55.6	56.8	74.8	56.3	59.8	**-6%**
	ImageNet	48.8	39.0	40.1	**59.6**	53.2	70.7	59.6	50.4	**15%**
	Caltech101	24.6	18.1	12.4	26.6	**100**	31.6	100	22.7	**77%**
	MSRC	33.8	18.2	30.9	20.8	69.5	**74.7**	74.7	34.6	**54%**
	Mean others	44.4	37.5	31.9	38.4	57.9	63.7	71.1	45.6	**36%**
"person" classification	SUN09	**16.1**	11.8	14.0	7.9	6.8	23.5	16.1	12.8	**20%**
	LabelMe	11.0	**26.6**	7.5	6.3	8.4	24.3	26.6	11.5	**57%**
	PASCAL	11.9	11.1	**20.7**	13.6	48.3	50.5	20.7	27.1	**-31%**
	ImageNet	8.9	11.1	11.8	**20.7**	76.7	61.0	20.7	33.9	**-63%**
	Caltech101	7.6	11.8	17.3	22.5	**99.6**	65.8	99.6	25.0	**75%**
	MSRC	9.4	15.5	15.3	15.3	93.4	**78.4**	78.4	29.8	**62%**
	Mean others	9.8	12.3	13.2	13.1	46.7	45.0	43.7	23.4	**47%**

Domain Adaptation

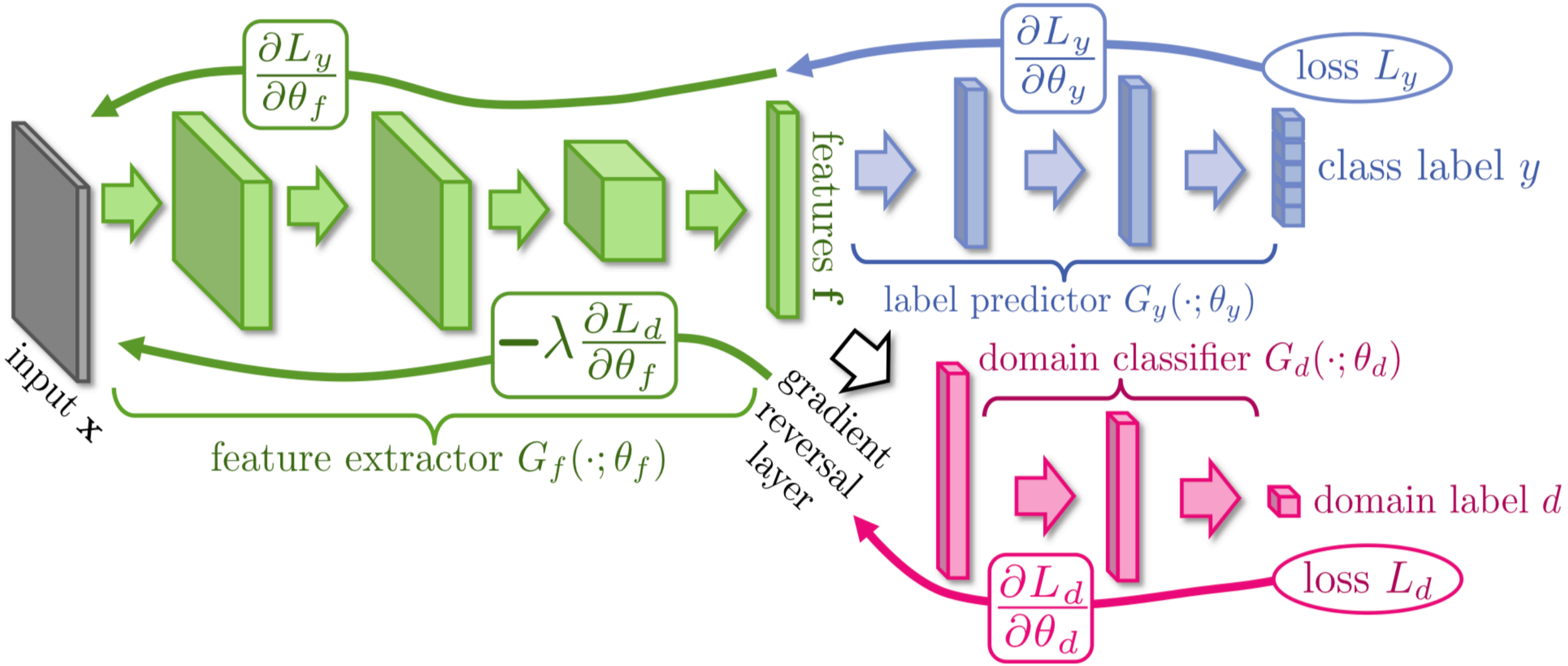
Make the target domain more like the source domain



Yaroslav Ganin and Victor Lempitsky.
Unsupervised Domain Adaptation by Backpropagation

Domain Adaptation

Make the target domain more like the source domain

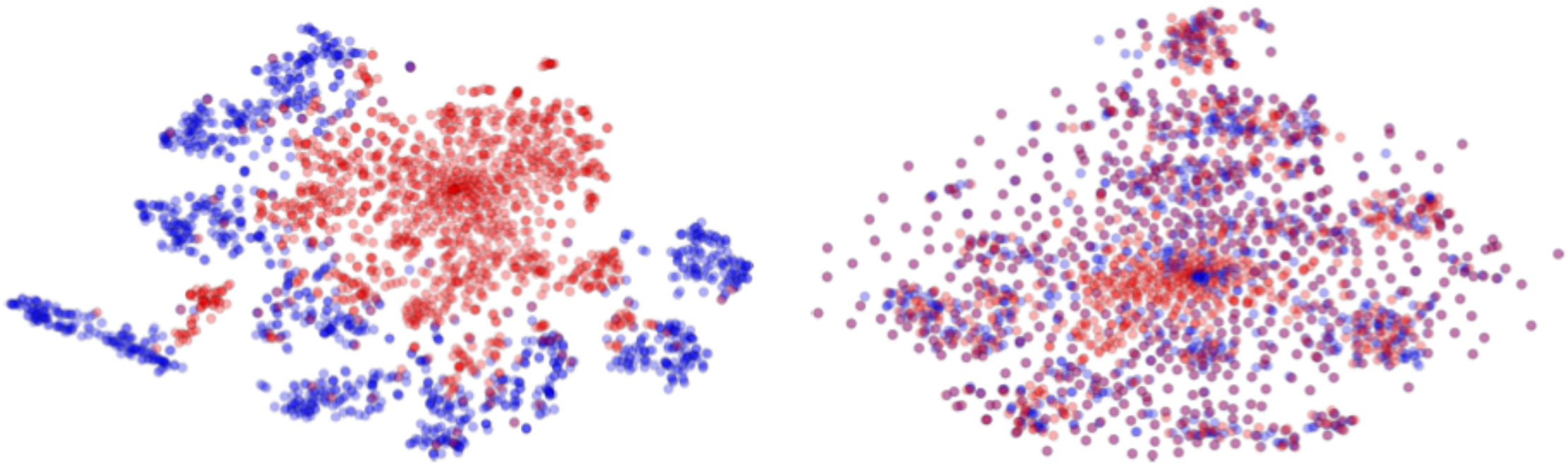


Yaroslav Ganin and Victor Lempitsky.
Unsupervised Domain Adaptation by Backpropagation

Domain Adaptation

Make the target domain more like the source domain

MNIST → MNIST-M: top feature extractor layer



(a) Non-adapted (b) Adapted

Yaroslav Ganin and Victor Lempitsky.
Unsupervised Domain Adaptation by Backpropagation

Technical fix is not enough

Where have we seen this before?

What can we do?

We need societal change.

Be aware of possible biases,
even implicit and unintentional ones.

Remain critical of “AI” applications.

AI4ALL's mission: increase diversity and inclusion in AI education, research, development and policy

"Until this program, I never thought that people who look like me could succeed in computer science and AI."

- AI4ALL 2016 student

Olga Russakovsky: The human side of computer vision

http://ai-4-all.org

“Like computers or the internal combustion engine, AI is a general-purpose technology that can be used to automate a great many tasks, including ones that should not be undertaken in the first place.”

- Blaise Agüera y Arcas, Alexander Todorov and Margaret Mitchell

Final Project Office Hours