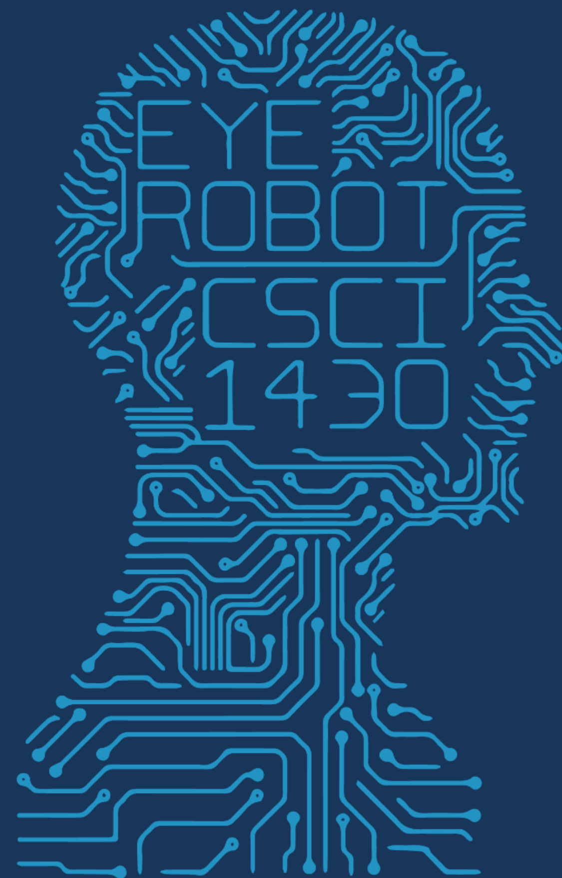
1950

Future Vision



2017 MWF 1pm 368

Computer Vision

CV as a social ~~good~~ bad?

CV / ML ‘human factors’

- Computer vision / machine learning is a tool.
- Tools are used under real world constraints.
  - Time, money.

- Like any tool, CVML can be used for good and for bad.
- What good/bad *is* sometimes depends on your point of view.

- Can also be used advertently or inadvertently.
- With awareness, or without awareness of ‘human factors’

Computer vision domain

In the beginning, there was light.

→

Eye
- Optics
- Dynamic range
- Color

Camera
- Optics
- Dynamic range
- Color

→

Digital processing

Low-level features

Calibration
- Geometric
- Photometric
- Radiometric

→

Data collection

Labeling

→

Model fitting

Recognition

Classification

Computer vision domain

In the beginning, there was light.

→

Eye
- Optics
- Dynamic range
- Color

Camera
- Optics
- Dynamic range
- Color

→

Digital processing

Low-level features

Calibration
- Geometric
- Photometric
- Radiometric

→

Data collection

Labeling

→

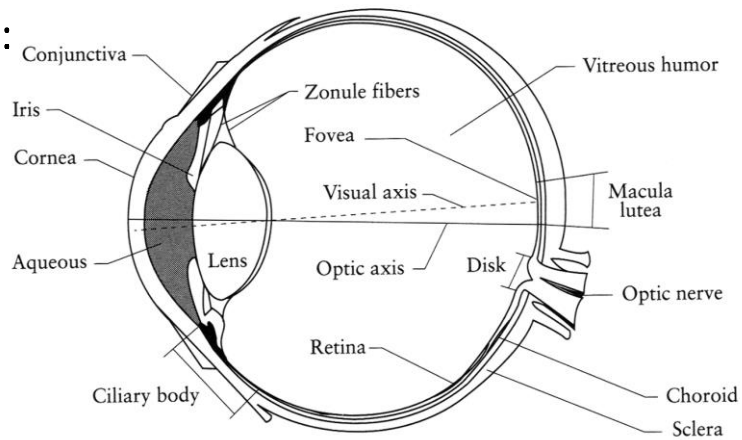
Model fitting

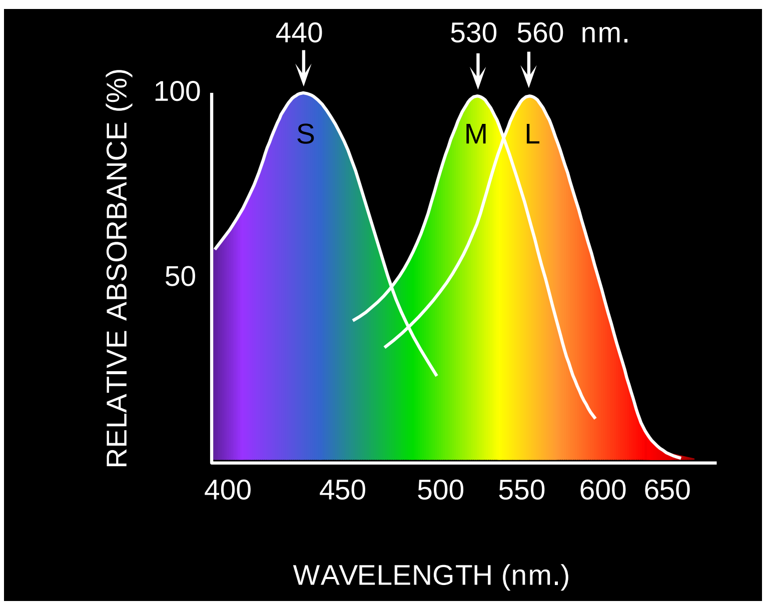
Recognition

Classification

Light response curves

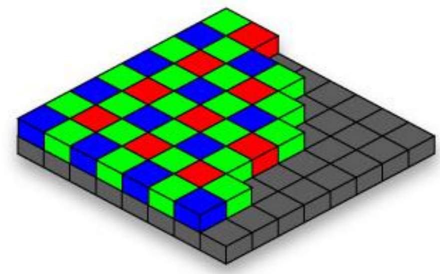
Eye Sensor:

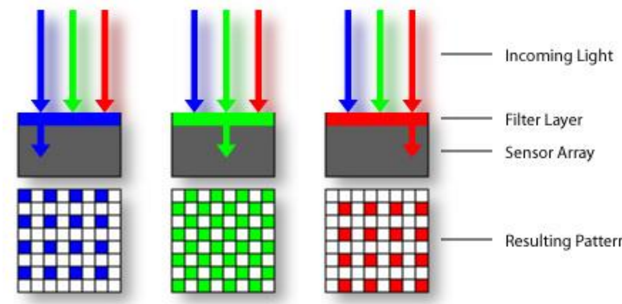


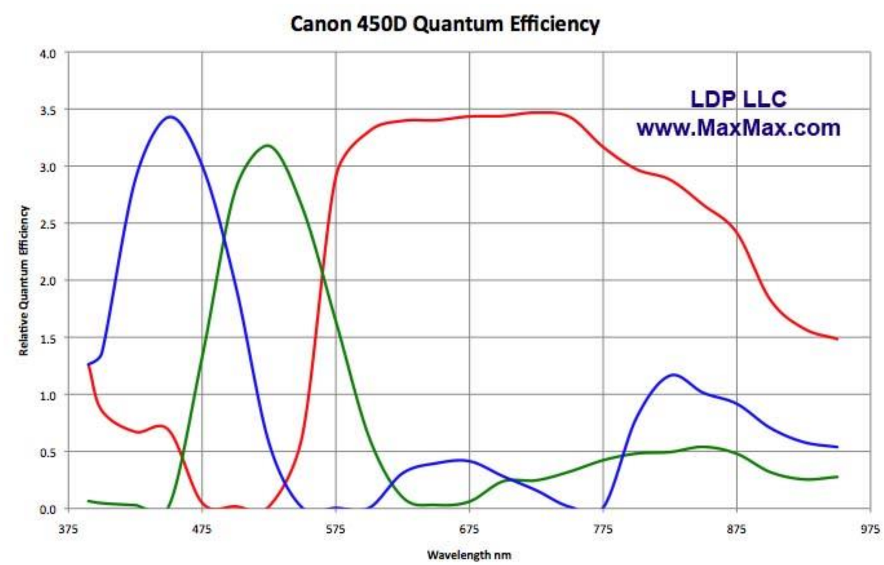


© Stephen E. Palmer, 2002

Camera Sensor:

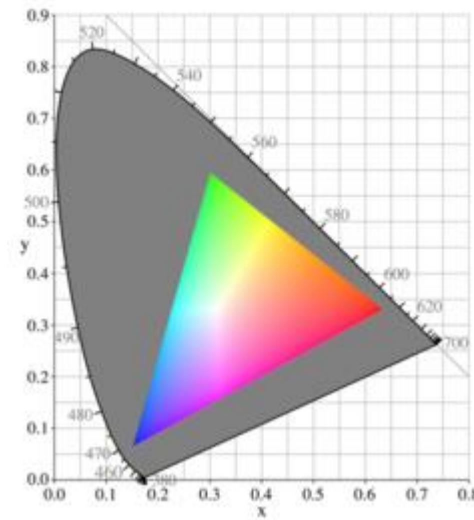






Light/reflectance output curves

250-500:1 contrast ratio (OLED = inf.)
6 / 8 / 10 bit dynamic range
3 / 4 additive primaries (RGB, rarely +yellow)
Defines a gamut



50-150:1 contrast ratio
??? dynamic range
4 subtractive primaries (CYMK)
Defines a gamut

We want:

Colors we see with our eyes in the world

=

Colors we see with our eyes in the reproduction

How do we calibrate these?

Time Warp: Film processing

http://www.picture-newsletter.com/kodak/

Kodak’s test input + output

- ‘Shirley cards’ – 1950s/60s
- Shirley was photographed hundreds of times by Kodak.
- One negative was processed as per Kodak specifications.
- A new unexposed negative + processed output was sent to each printer lab.
- Printer colors were calibrated on site as per the target Shirley card.

Circa 1960

Kodak's test input + output

- 'Shirley cards' – 1950s/60s

- Any issues with this approach?

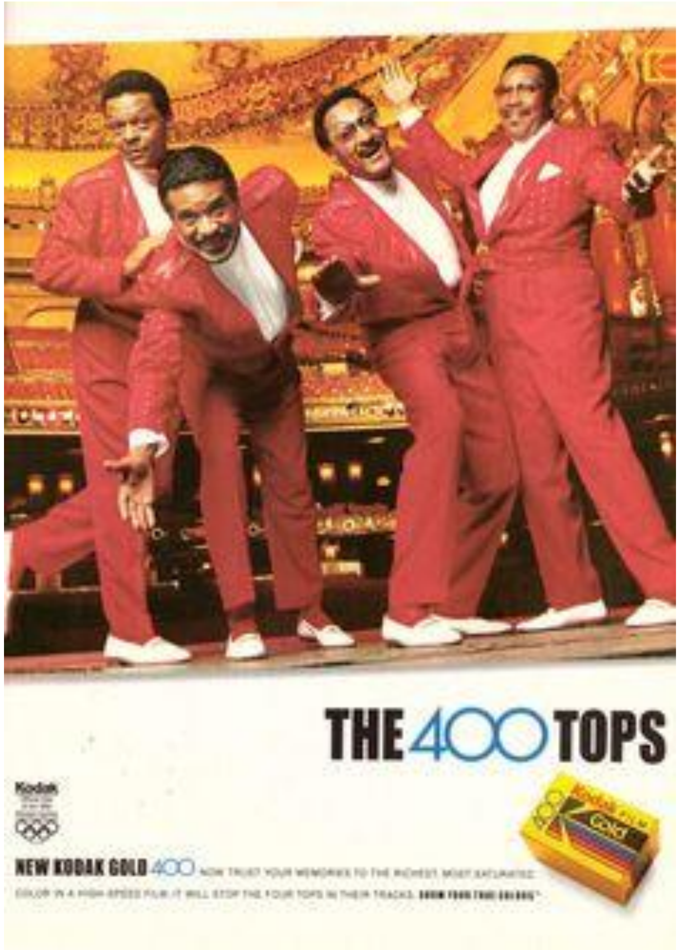
Circa 1960

Over time

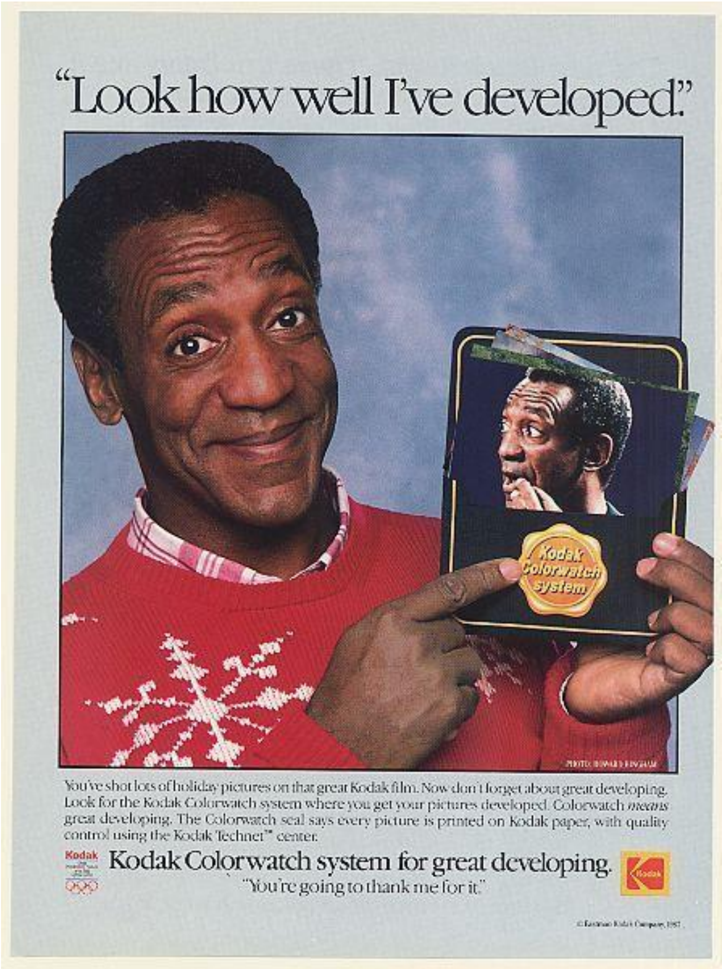
- 1978: Filmmaker Jean-Luc Godard refuses to use Kodachrome film in Mozambique.
- 1980s: Chocolate and furniture manufacturers complain.
- 1986: Kodacolor VR-G (or Gold) – film for dark browns.
    - “Photograph the details of a dark horse in low light.”

Shirley card, 1996

1980s – adverts



The Four Tops!



Bill Cosby!
Some other issues here too : (

What are the underlying problems?

- …and how might we overcome them?
- Think-pair-share.

Issues

- Dynamic range: not enough!
- Color balance:

So digital fixes this, right?

- Well…

"[The hardest part of being in a biracial relationship is taking a picture together."]

[whatthecaptcha]

So digital fixes this, right?

…it’s a lot better.

- 14-bit sensors (≈ eye’s static range)
- High-dynamic range by combining low-dynamic range
- Digital post-processing for color balance

References

Canadian Journal of Communication:

Roth et al., Looking at Shirley, the Ultimate Norm: Colour Balance, Image Technologies, and Cognitive Equity

http://www.cjc-online.ca/index.php/journal/article/view/2196

http://www.npr.org/2014/11/13/363517842/for-decades-kodak-s-shirley-cards-set-photography-s-skin-tone-standard/

https://priceonomics.com/how-photography-was-optimized-for-white-skin/

https://www.buzzfeed.com/syreetamcfadden/teaching-the-camera-to-see-my-skin/

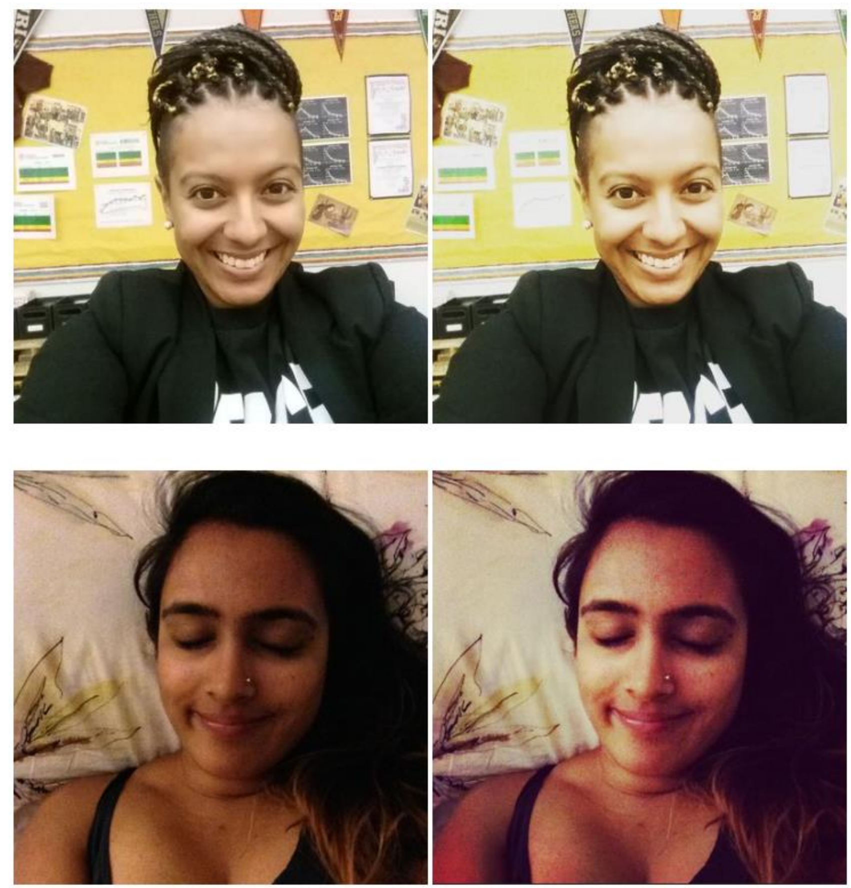
Word of warning

- Around 2013/2014 there were a lot of articles about this issue.
- Many articles rewrite the same few sources.
- Most do not have a technical background, and sometimes technical issues are confused.
- ‘Take care.’

Instagram filters

- Filters that brighten
- Filters that darken

- Filters can do anything!



Snapchat

4/20

"Anime inspired"

Dataset Bias

Computer vision domain

In the beginning, there was light.

→

Eye
- Optics
- Dynamic range
- Color

Camera
- Optics
- Dynamic range
- Color

→

Digital processing

Low-level features

Calibration
- Geometric
- Photometric
- Radiometric

→

Data collection

Labeling

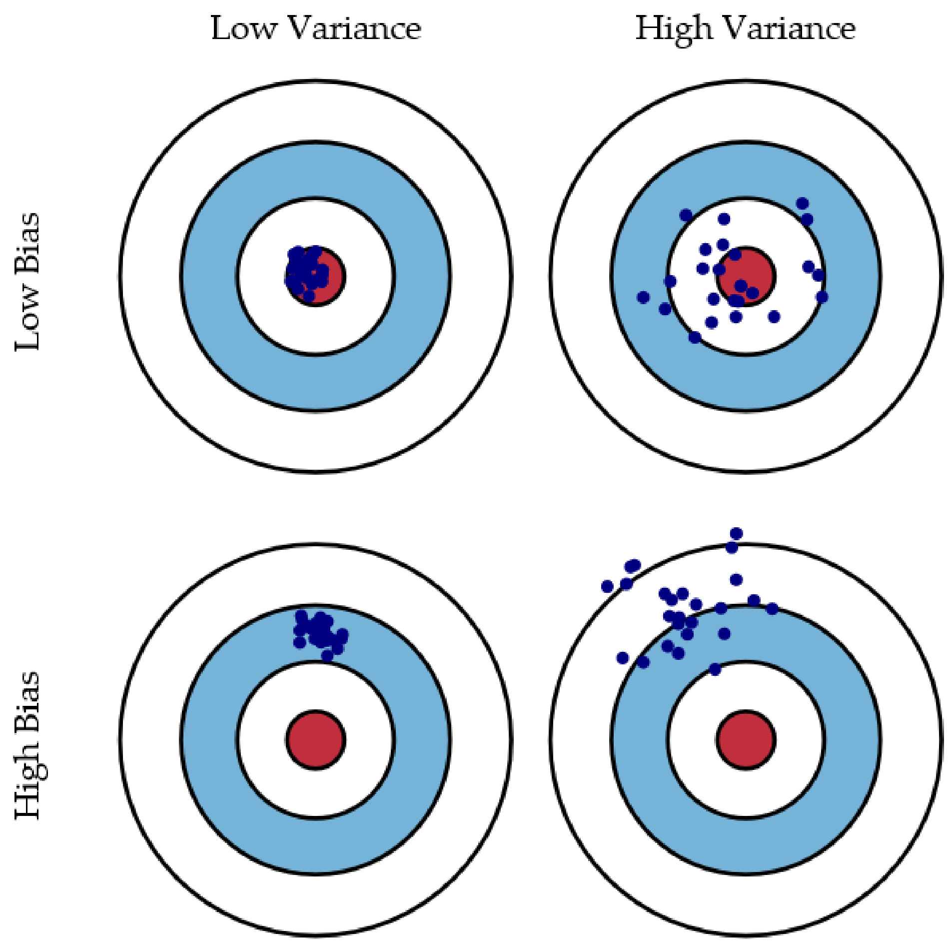
→

Model fitting

Recognition

Classification

Bias/variance trade-off



Bias = accuracy
Variance = precision

Scott Fortmann-Roe

Unbiased Look at Dataset Bias

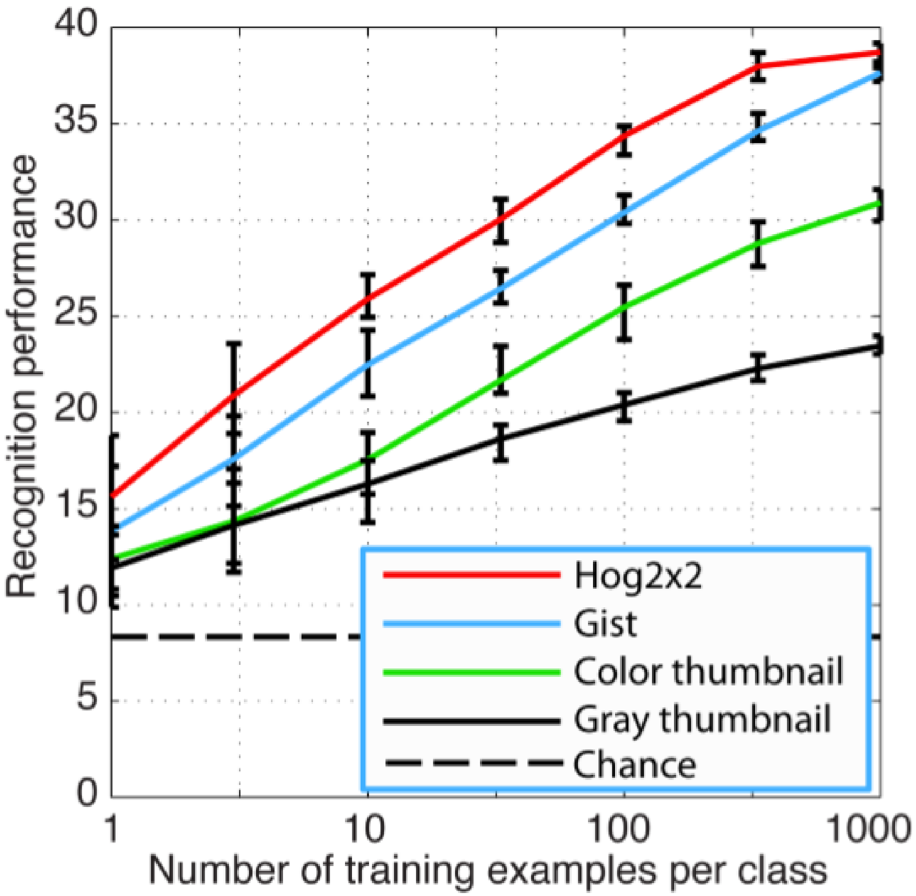
Torralba and Efros, CVPR 2011

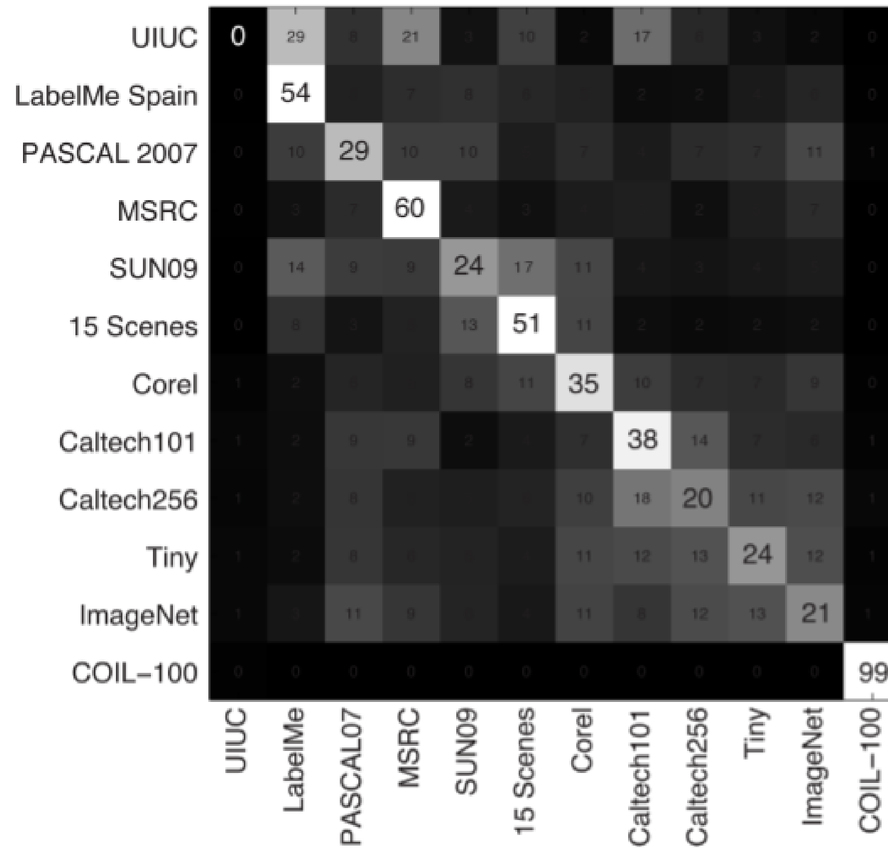
“The authors would like to thank the Eyjafjallajokull volcano as well as the wonderful kirs at the Buvette in Jardin du Luxembourg for the motivation (former) and the inspiration (latter) to write this paper.“

Next few slide contents are from the paper

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!





Progression of dataset complexity

- COIL-100:

- 15 scenes: Out of the lab, backgrounds
- Caltech-101: Google-mined, single object in middle.
- LabelMe: Multiple objects, anywhere
- PASCAL VOC: More rigorous testing standards
- ImageNet: Internet-scale, real-world

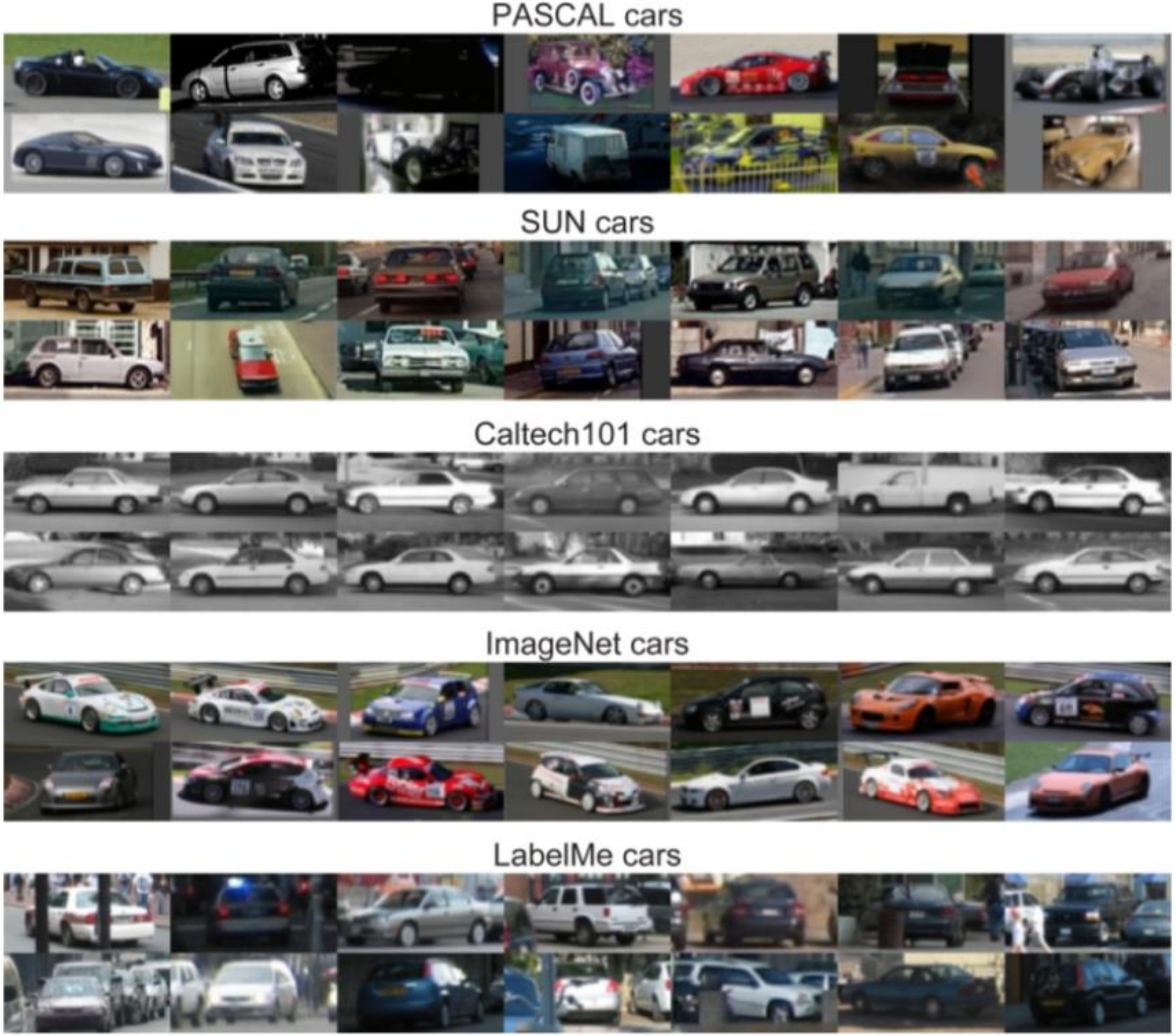


Figure 4. Most discriminative cars from 5 datasets

Measuring Dataset Bias

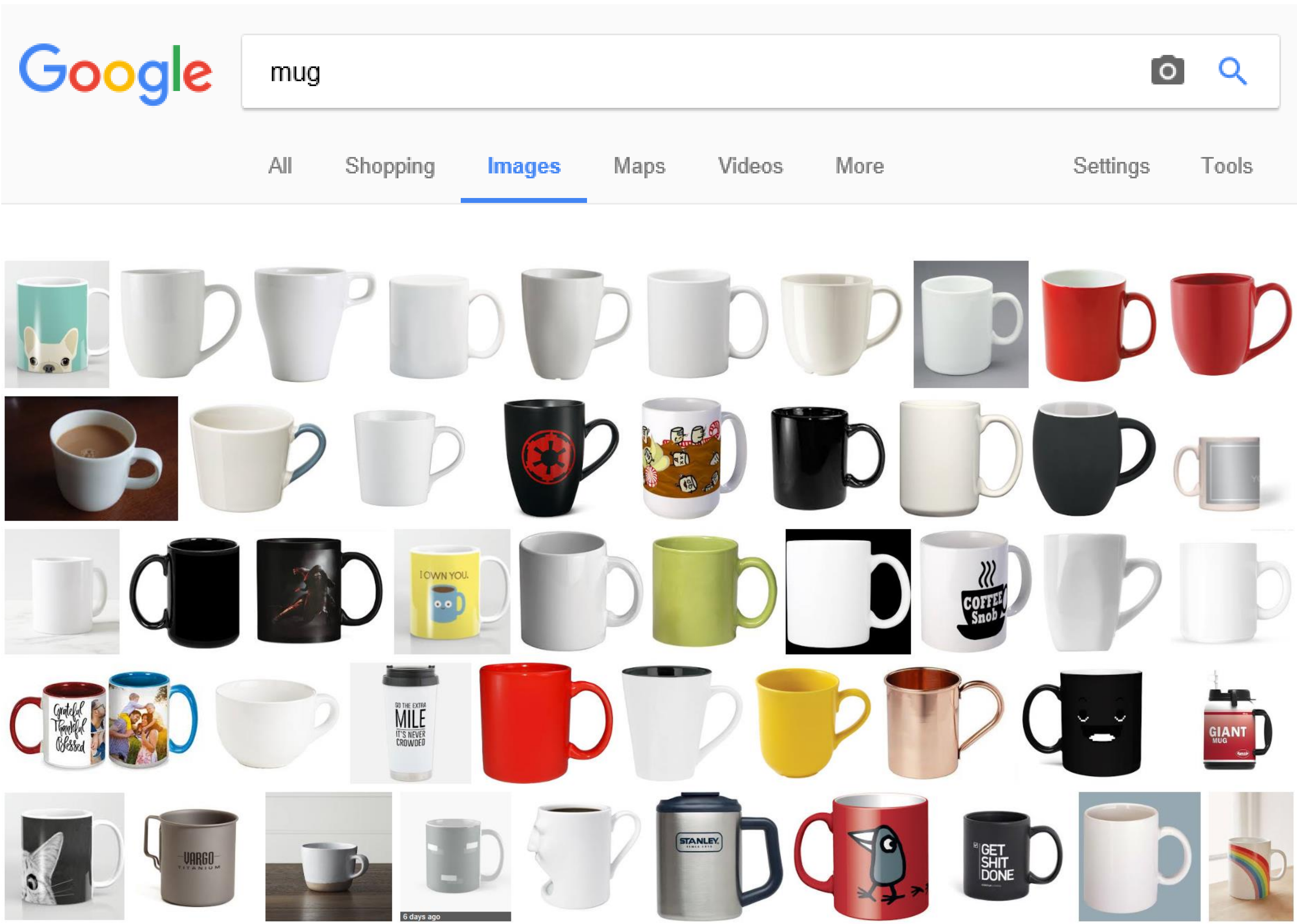
- Idea: cross-dataset generalization

- Train an object classifier on one dataset
- Test on the same object class on another dataset
- Observe performance as measure of bias

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%**

Different kinds of bias

- *Selection bias*
  - Retrieve different kinds of images; keywords/search engines can bias.
- *Capture bias*
  - Objects photographed in similar ways that do not generalize, e.g., object always in center, race track car vs. street car, mugs.


Google
mug
All
Shopping
Images
Maps
Videos
More
Settings
Tools
I OWN YOU.
COFFEE Snob
GO THE EXTRA MILE IT'S NEVER CROWDED
GIANT MUG
VARGO TITANIUM
6 days ago
STANLEY
GET SHIT DONE

Different kinds of bias

- *Selection bias*
  - Retrieve different kinds of images; keywords/search engines can bias.
- *Capture bias*
  - Objects photographed in similar ways that do not generalize, e.g., object always in center, race track car vs. street car, mugs.
- *Category/label bias*
  - Poorly-defined classes, e.g., painting vs. picture
- *Negative set bias*
  - In one vs. all classification, 'all' or "the rest of the world" is not well represented.
  - "Are features which helps classify 'boat' object really the boat, or are they the water it sits on?"
    - Low bias negative set would include many boat-free images of rivers and lakes.

Measuring Negative Set Bias

- Take negative examples from other datasets and add to superset; train against this.

task	Positive Set: / Negative Set:	SUN09	LabelMe	PASCAL	ImageNet	Caltech101	MSRC	Mean
"car" detection	self	67.6	62.4	56.3	60.5	97.7	74.5	70.0
	all	53.8	51.3	47.1	65.2	97.7	70.0	64.1
	percent drop	20%	18%	16%	-8%	0%	6%	8%
"person" detection	self	67.4	68.6	53.8	60.4	100	76.7	71.1
	all	52.2	58.0	42.6	63.4	100	71.5	64.6
	percent drop	22%	15%	21%	-5%	0%	7%	9%

- Drop in performance of 'all' suggests negative examples are being misclassified

Overcoming bias at collection time

- *Selection bias*
  - Multiple keywords, search engines, countries.
  - Collect unknown images and label them by crowd-sourcing.
- *Capture bias*
  - Better sampling
  - Different transforms: noise, flips, rotations, affine, crops.

Overcoming bias at collection time

- *Category/label bias*
  - Clear instruction to turkers; unambigous classes (possible?)
  - Pre-label clustering, or multiple acceptable answers.

- *Negative set bias*
  - Cross-dataset mining
  - Mine for hard negatives from unlabeled set using a reliable algorithm and high threshold.

Undoing the Damage of Dataset Bias

Khosla et al., ECCV 2012

"While it remains in question whether creating an unbiased dataset is possible given limited resources, we propose a discriminative framework that directly exploits dataset bias during training."

More examples

https://www.quora.com/What-are-examples-of-computer-vision-bugs-related-to-race

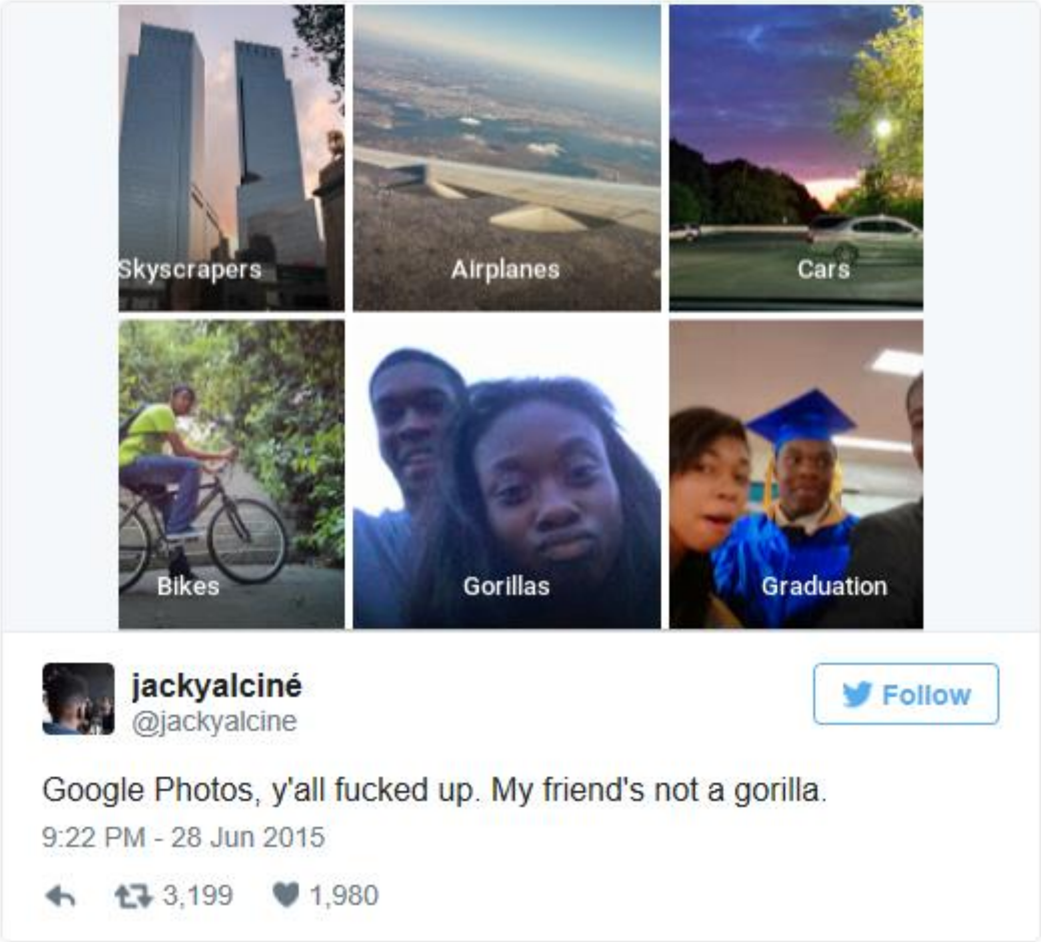
http://www.telegraph.co.uk/technology/2016/12/07/robot-passport-checker-rejects-asian-mans-photo-having-eyes/

Thank you Tiffany Chen

Viola-Jones with a bad training database

wzamen01

Google Photos (2015)



[Jacky Alciné]

Google Photos (2015)

- What do you think the problem was?
- How could you fix it?

- Has it been fixed? Anyone use Google Photos?

Google Photos (2015)

(((Yonatan Zunger)))
@yonatanzunger

Follow

@jackyalcine Quick update: we shouldn't be making piles with that label anymore, and searches are mostly fixed, but they can still turn up..

(((Yonatan Zunger)))
@yonatanzunger

Follow

@jackyalcine ..photos where we failed to recognize that there was a face there at all. We're working on that issue now.

LIKES
8

11:17 AM - 29 Jun 2015

2 8

AI 'Safety'

Concrete Problems in AI Safety

- https://arxiv.org/abs/1606.06565

In context of robots, but promising ideas

  - Regularizer based on expert 'risk' of class confusion

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

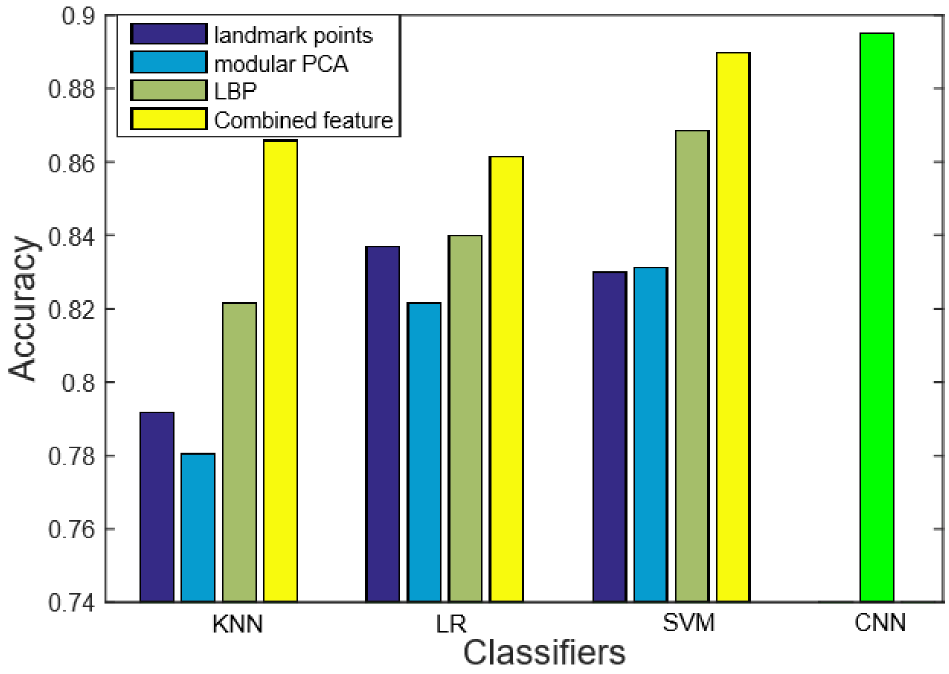
Figure 1. Sample ID photos in our data set.

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."

Criminality

- 1100 non-criminal, 730 criminal Chinese face photos
- Tested various features + classifiers



Criminality

K-means, averaging clusters

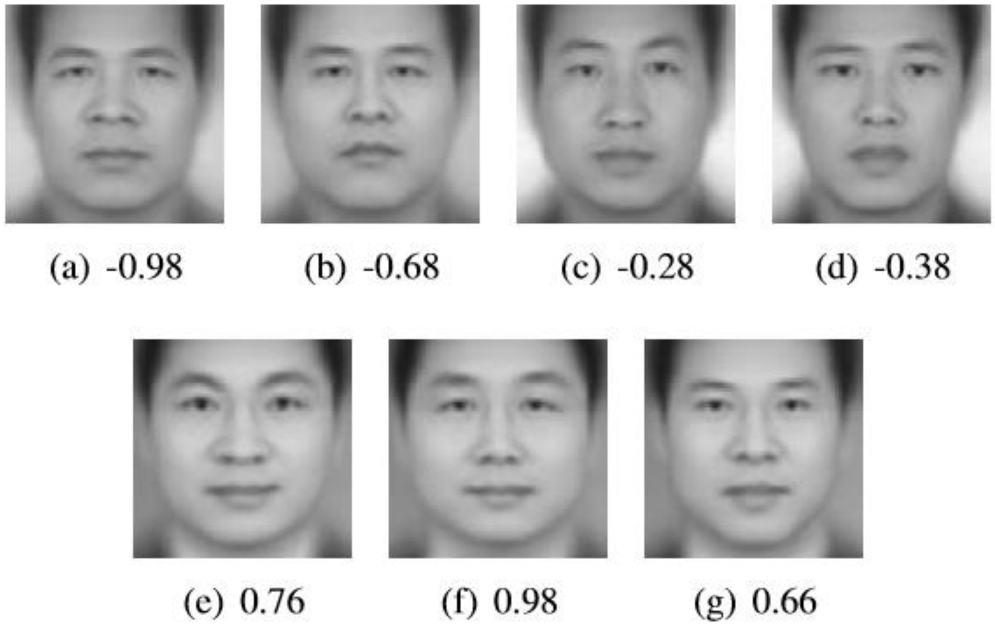


Figure 13. (a), (b), (c) and (d) are the four subtypes of criminal faces corresponding to four cluster centroids on the manifold of S_c; (e), (f) and (g) are the three subtypes of non-criminal faces corresponding to three cluster centroids on the manifold of S_n. The number associated with each face is the average score of human judges (-1 for criminals; 1 for non-criminals).

What biases might exist? Discuss!

- Selection bias
- Capture bias
- Category/label bias
- Negative set bias

Is this real?

- Whatever the case, it needs care! Significant ramifications.

- Humans might be able to do this:
  - *Small but statistically significant ability to tell criminal from non-criminal in photo.*

Valla, J., Williams, W., & Ceci, S. J. (2011).

The accuracy of inferences about criminality based on facial appearance.

Journal of Social, Evolutionary, and Cultural Psychology, 5(1), 66-91.

MIT Technology Review has a good overview:

https://www.technologyreview.com/s/602955/neural-network-learns-to-identify-criminals-by-their-faces/

1

2

3

4

5

6

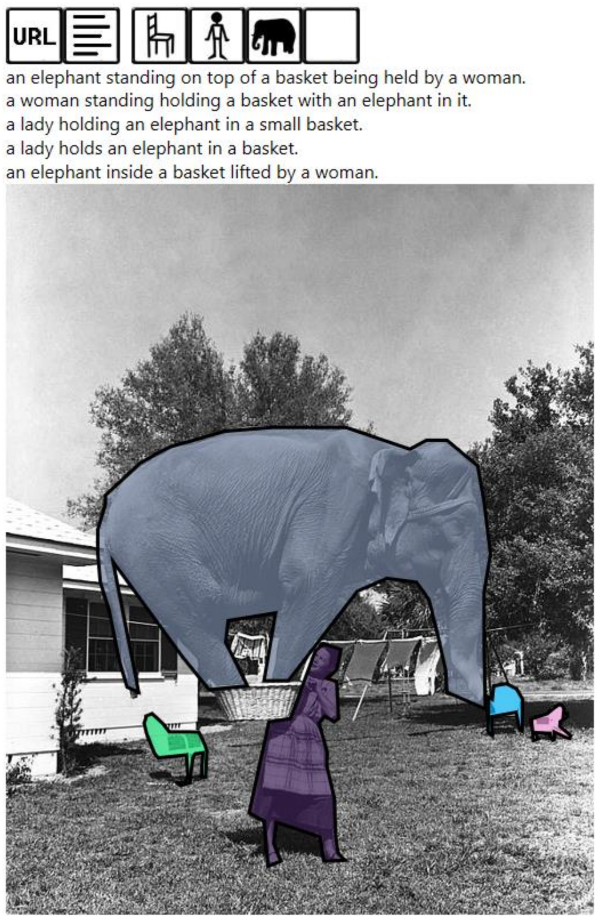
Valla, J., Williams, W., & Ceci, S. J. (2011)

“Guns don't kill people, *people* kill people!”

“Machine learning doesn’t kill people,
training data kills people!”

- ML community, all the time.

Dataset improvement: MS COCO



What is COCO?

COCO is a new image recognition, segmentation, and captioning dataset. COCO has several features:

- **Object segmentation**
- **Recognition in Context**
- **Multiple objects per image**
- **More than 300,000 images**
- **More than 2 Million instances**
- **80 object categories**
- **5 captions per image**
- **Keypoints on 100,000 people**

Help Do Something About It: Upcoming Hackathon

May 11th @ Microsoft New England in Boston

Organized by New England Machine Learning Day

Hacking Bias in ML

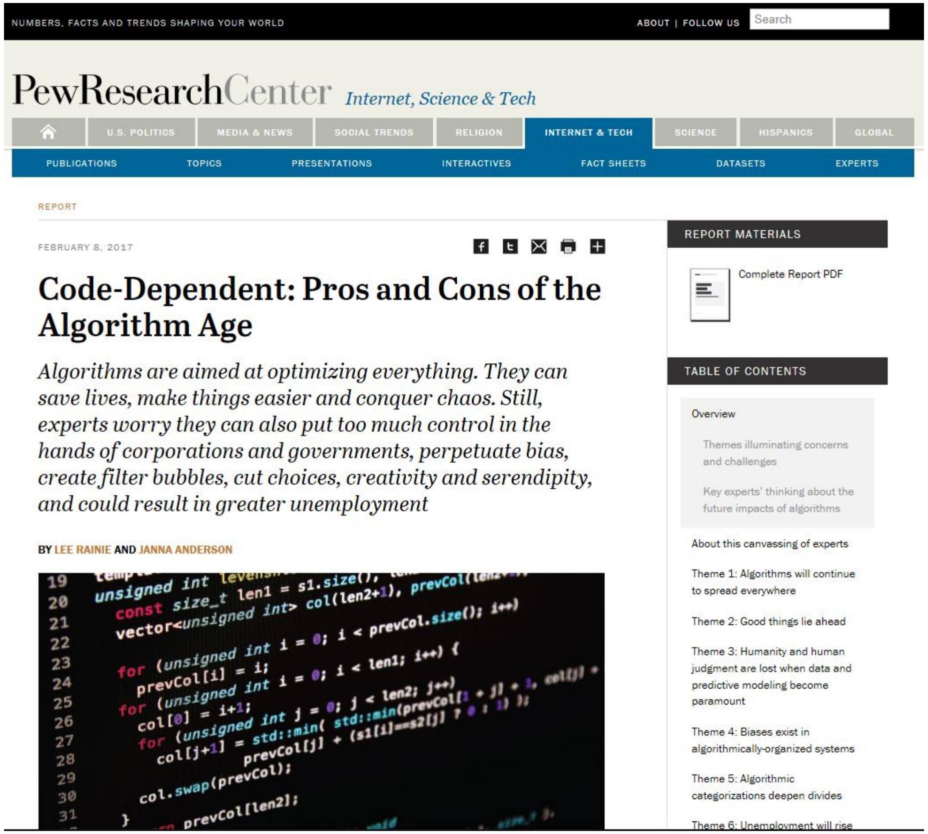
Ex-Brown PhD will be there - *Genevieve Patterson*

https://www.eventbrite.com/e/new-england-machine-learning-hackathon-hacking-bias-in-ml-tickets-32951771636?aff=NEML

Decent Pew Overview on Big Picture

Rainie and Anderson

Code-Dependent: Pros and Cons of the Algorithm Age



http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/