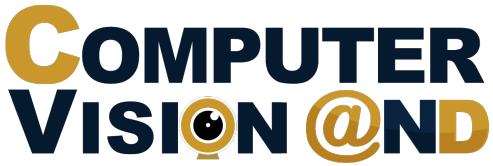
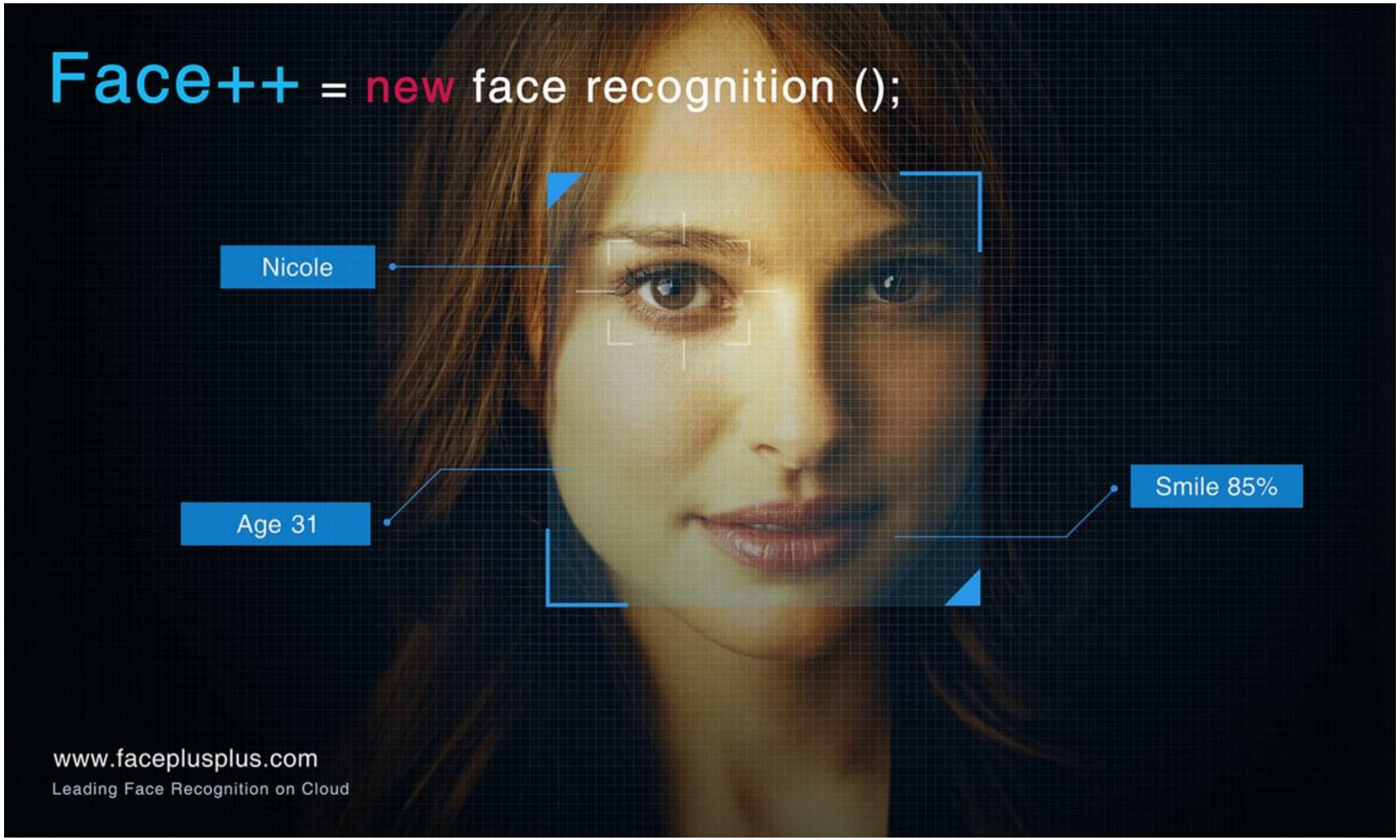
The “Criminality from Face” Illusion

Walter J. Scheirer

Computer Vision Research Laboratory
Department of Computer Science and Engineering



Facial Analysis



https://www.faceplusplus.com/

Kevin Bowyer

Mike King

The “Criminality From Face” Illusion

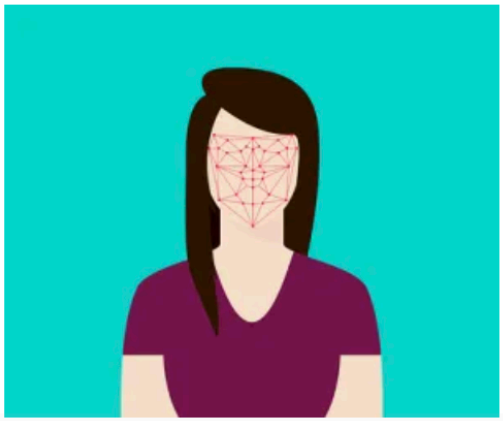
IEEE Transactions on Technology and Society
To Appear

Be on guard for technology that can’t possibly work

HU facial recognition software predicts criminality

HU facial recognition software predicts criminality

A group of Harrisburg University professors and a Ph.D. student have developed automated computer facial recognition software capable of predicting whether someone is likely going to be a criminal.



With 80 percent accuracy and with no racial bias, the software can predict if someone is a criminal based solely on a picture of their face. The software is intended to help law enforcement prevent crime.

Ph.D. student and NYPD veteran Jonathan W. Korn, Prof. Nathaniel J.S. Ashby, and Prof. Roozbeh Sadeghian titled their research “A Deep Neural Network Model to Predict Criminality Using Image Processing.”

http://archive.is/N1HVe

Facial recognition software paper not being published

Facial recognition software paper not being published

If you are looking for a news release outlining the paper "A Deep Neural Network Model to Predict Criminality Using Image Processing," it was removed from the website at the request of the faculty involved in the research. The editor of the publication where it was scheduled to appear has decided to not publish the work.

At Harrisburg University of Science and Technology, we harness the fruits of our research with the intention to empower people and communities. As a world-class university, we are committed to the continuous cycle of scientific inquiries and discoveries subject to commonly accepted research standards within higher education. Each new finding potentially leads to the next scientific breakthrough in an ongoing search for innovations that contribute to uplifting and improving all lives.

See the petition against this work: https://medium.com/@CoalitionForCriticalTechnology/abolish-the-techtoprisonpipeline-9b5b14366b16

Hashemi and Hall *J Big Data* (2020) 7:2
https://doi.org/10.1186/s40537-019-0282-4

Journal of Big Data

RESEARCH Open Access

Criminal tendency detection from facial images and the gender bias effect

Mahdi Hashemi[1*] and Margeret Hall[2]

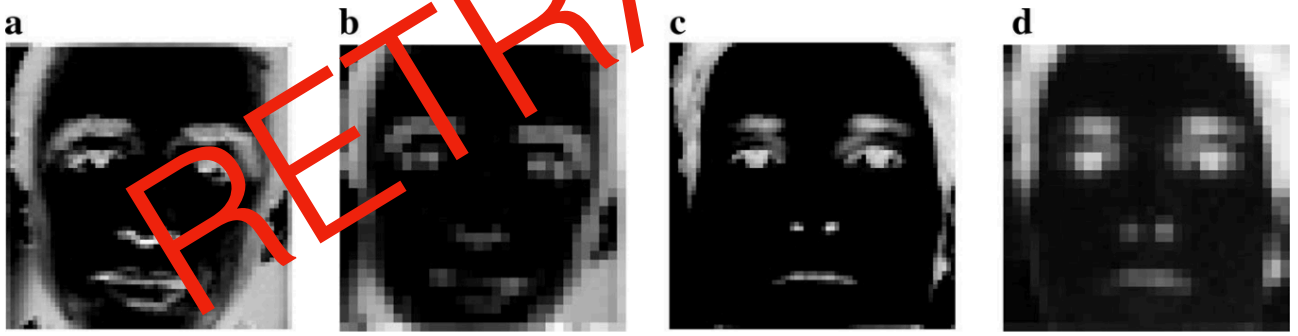


Fig. 6 Facial features detected by the first (**a**, **c**) and second (**b**, **d**) convolutional layers in CNN, for a criminal (**a**, **b**) vs. non-criminal (**c**, **d**) face shot

M. Hashemi and M. Hall. Criminal tendency detection from facial images and the gender bias effect. Journal of Big Data, 7(2), 2020.

Automated Inference on Criminality using Face Images

Xiaolin Wu
McMaster University
Shanghai Jiao Tong University
xwu510@gmail.com

Xi Zhang
Shanghai Jiao Tong University
zhangxi_19930818@sjtu.edu.cn

Abstract

We study, for the first time, automated inference on criminality based solely on still face images, which is free of any biases of subjective judgments of human observers. Via supervised machine learning, we build four classifiers (logistic regression, KNN, SVM, CNN) using facial images of 1856 real persons controlled for race, gender, age and facial expressions, nearly half of whom were convicted criminals, for discriminating between criminals and non-criminals. All four classifiers perform consistently well and empirically establish the validity of automated face-induced inference on criminality, despite the historical controversy surrounding this line of enquiry. Also, some discriminating structural features for predicting criminality have been found by machine learning. Above all, the most important discovery of this research is that criminal and non-criminal face images populate two quite distinctive manifolds. The variation among criminal faces is significantly greater than that of the non-criminal faces. The two manifolds consisting of criminal and non-criminal faces appear to be concentric, with the non-criminal manifold lying in the kernel with a smaller span, exhibiting a law of "normality" for faces of non-criminals. In other words, the faces of general law-biding public have a greater degree of resemblance compared with the faces of criminals, or criminals have a higher degree of dissimilarity in facial appearance than non-criminals.

1. Introduction

Motivated by many commercial applications of artificial intelligence and man-machine interfaces, the research communities of pattern recognition and computer vision have devoted a great deal of efforts to the recognition and manipulation of human faces [11, 31, 40, 35], and achieved measured successes. But very little research has been done on analyzing and quantifying social perception and attributes of faces [33], although this subject is of great importance to many academic disciplines, such as social psychology, management science, criminology, etc.

In all cultures and all periods of recorded human history, people share the belief that the face alone suffices to reveal innate traits of a person. Aristotle in his famous work Prior Analytics asserted, "It is possible to infer character from features, if it is granted that the body and the soul are changed together by the natural affections". Psychologists have known, for as long as a millennium, the human tendency of inferring innate traits and social attributes (e.g., the trustworthiness, dominance) of a person from his/her facial appearance, and a robust consensus of individuals' inferences . These are the facts found through numerous studies [3, 39, 5, 6, 10, 26, 27, 34, 32].

Independent of the validity of pedestrian belief in the (pseudo)science of physiognomy, a tantalizing question naturally arises: what facial features influence average Joes' impulsive and yet consensual judgments on social attributes of a non-acquaintance member of their own specie? Attempting to answer the question, Todorov and Oosterhof proposed a data-driven statistical modeling method to find visual determinants of social attributes by asking human subjects to score four percepts: dominance, attractiveness, trustworthiness, and extroversion, based on first impression of static face images [33]. This method can synthesize a representative (average) face image for a set of input face images scored closely on any of the four aforementioned social percepts. The ranking of these synthesized face images by subjective scores (e.g., from least to most trustworthy looking) apparently agrees with the intuition of most people.

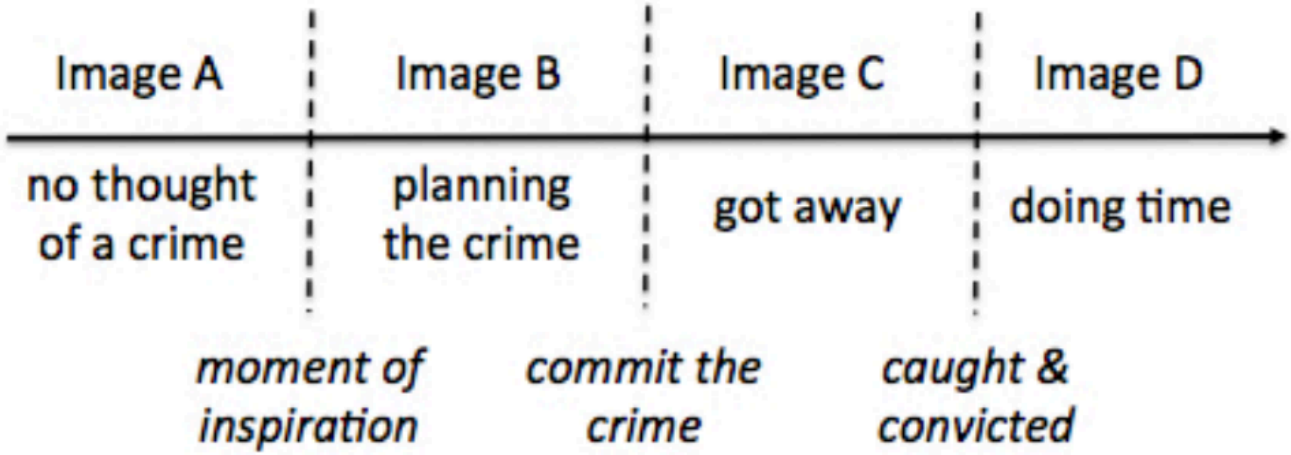
Following the consensus in social perception from facial appearance, arrives the next even bigger speculation: is there any diagnostic merit of the face-induced inferences on an individual's social attributes? In this paper we intend not to nor are we qualified to discuss or debate on societal stereotypes, rather we want to satisfy our curiosity in the accuracy of fully automated inference on criminality. At the onset of this study our gut feeling is that modern tools of machine learning and computer vision will refute the validity of physiognomy, although the outcomes

An Illusory Problem Definition

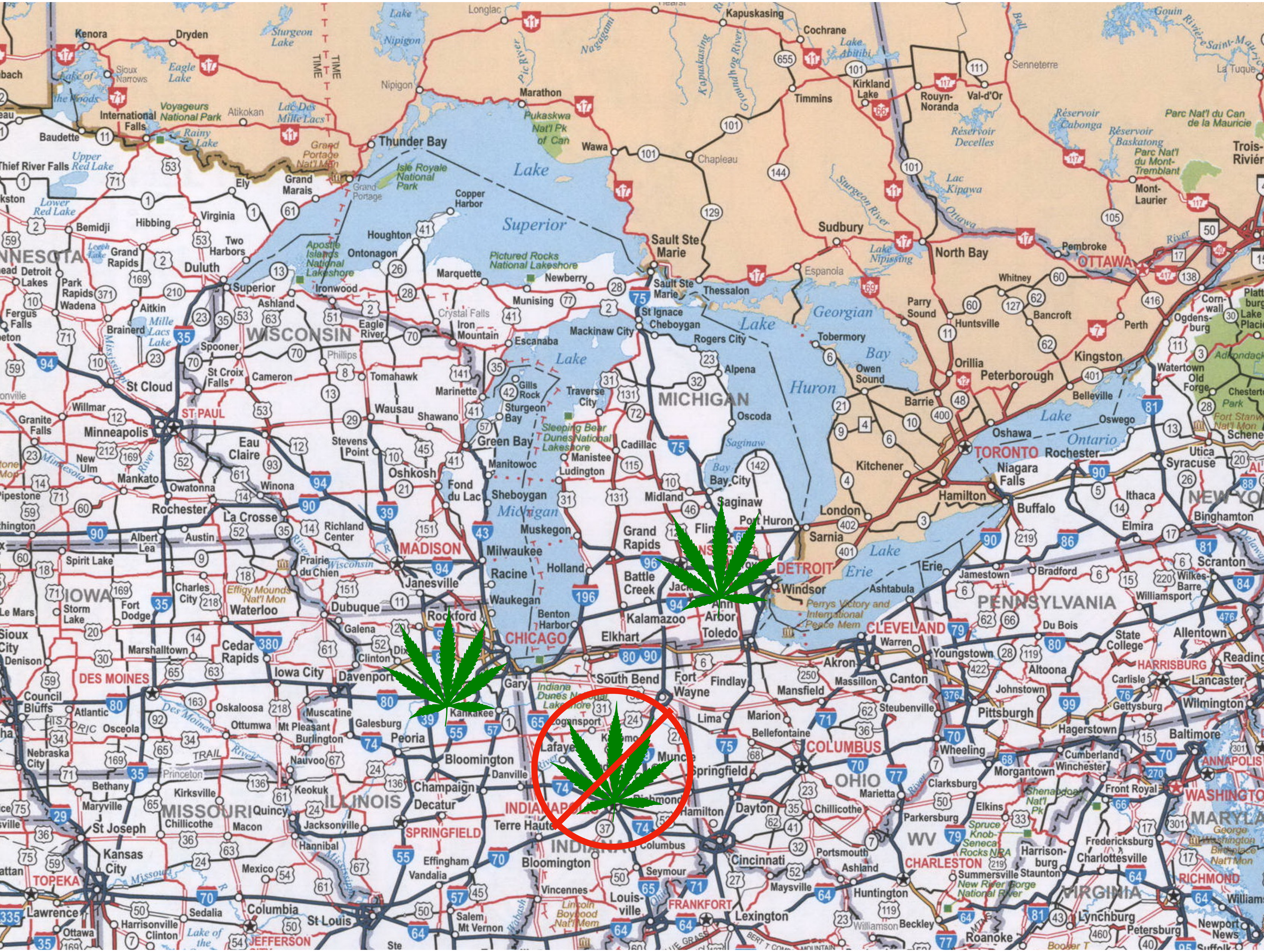
Consider a person who to a given point in their life has never even thought of committing a crime

Let’s call this person Raskolnikov

Images from four different points in Raskolnikov's life



What are the ground-truth labels for these images?



Lake Superior
Lake Michigan
Lake Huron
Georgian Bay
Lake Erie
Lake Ontario
WISCONSIN
MICHIGAN
IOWA
ILLINOIS
INDIANA
OHIO
PENNSYLVANIA
MISSOURI
WV
VIRGINIA
CHICAGO
DETROIT
CLEVELAND
MADISON
DES MOINES
COLUMBUS
SPRINGFIELD
INDIANAPOLIS
CHARLESTON
FRANKFORT
HARRISBURG
TORONTO
OTTAWA
ST PAUL
TOPEKA
JEFFERSON CITY
WASHINGTON
RICHMOND
ANNAPOLIS
Thunder Bay
Sault Ste Marie
Sudbury
North Bay
Minneapolis
Milwaukee
Grand Rapids
Lansing
Toledo
Windsor
Buffalo
Pittsburgh
Cincinnati
St Louis
Kansas City

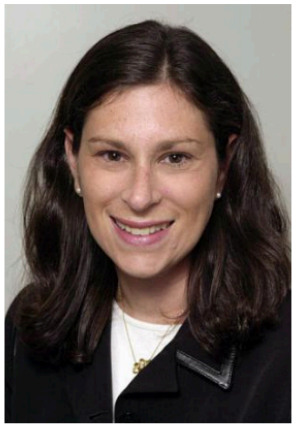
Illusory Experimental Results

The dataset used by Hashemi and Hall

NIST Special
DB 18 →

Criminal

(a)



Non-Criminal

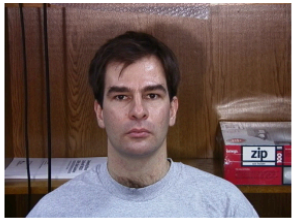
(b)

Non-Criminal

(c)

(d)

Non-Criminal



(e)

Non-Criminal

(f)

Non-Criminal

What's wrong with this picture?

Some trouble with the labeling of Criminal / Non-Criminal photos

- In the US, a mugshot **does not** indicate that a person has been convicted of a crime
    - NIST Special DB 18 does not indicate how each subject's case resolved
- 15% of all exonerees in the US originally pled guilty
- There is no way to verify that persons labeled non-criminal have never committed a crime

Confusion with Models of "First Impressions"

Wu and Zhang's motivation

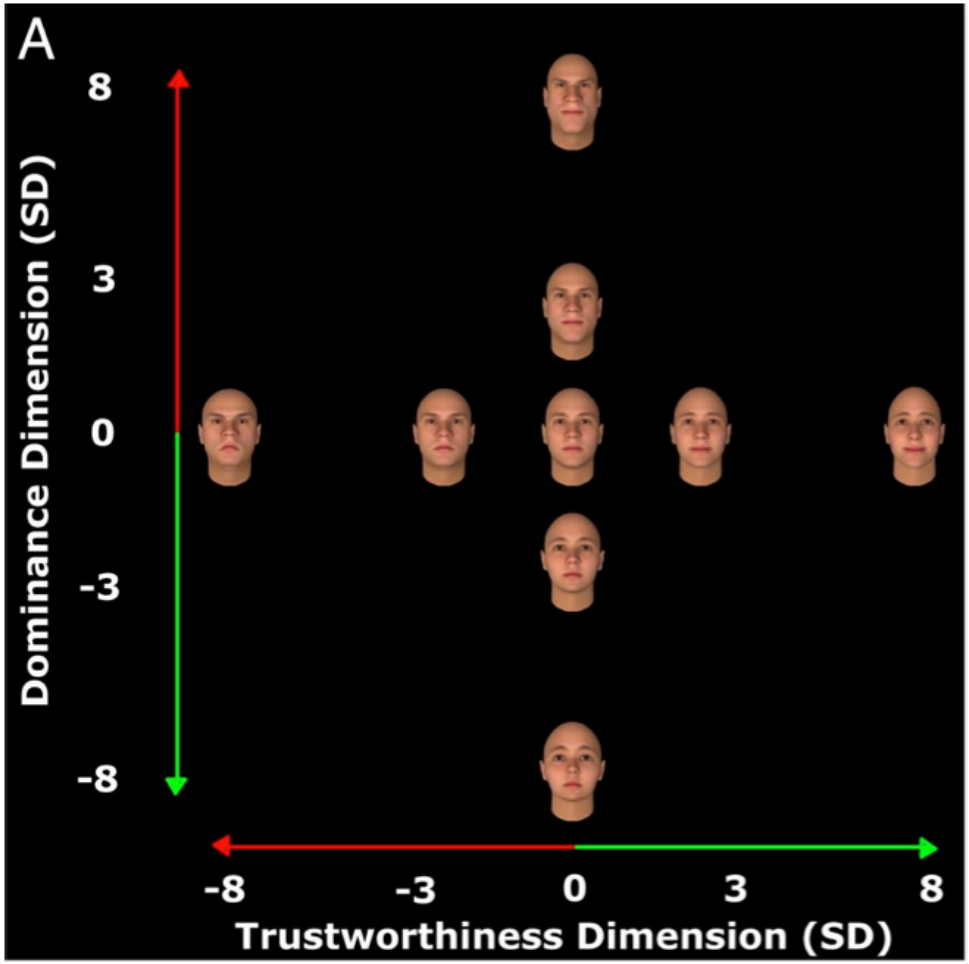
"Todorov and Oosterhof proposed a data-driven statistical modeling method to find visual determinants of social attributes by asking human subjects to score four percepts: dominance, attractiveness, trustworthiness, and extroversion, based on first impression of static face images [33].

…

Following the consensus in social perception from facial appearance, arrives the next even bigger speculation: is there any **diagnostic merit** of the face-induced inferences on an individual's social attributes?"

(emphasis mine)

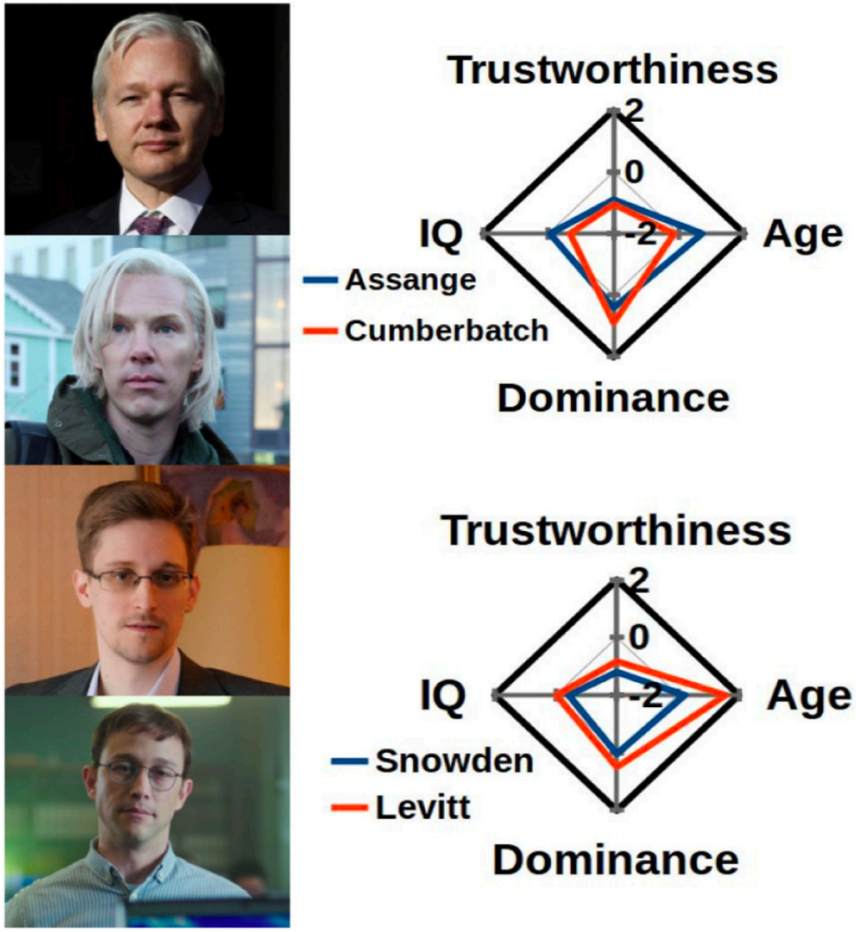
Trustworthiness and Dominance



Oosterhof & Todorov, PNAS (2008)

Predicting *first impressions* is possible

Ponce-López et al. ECCV Workshops 2016



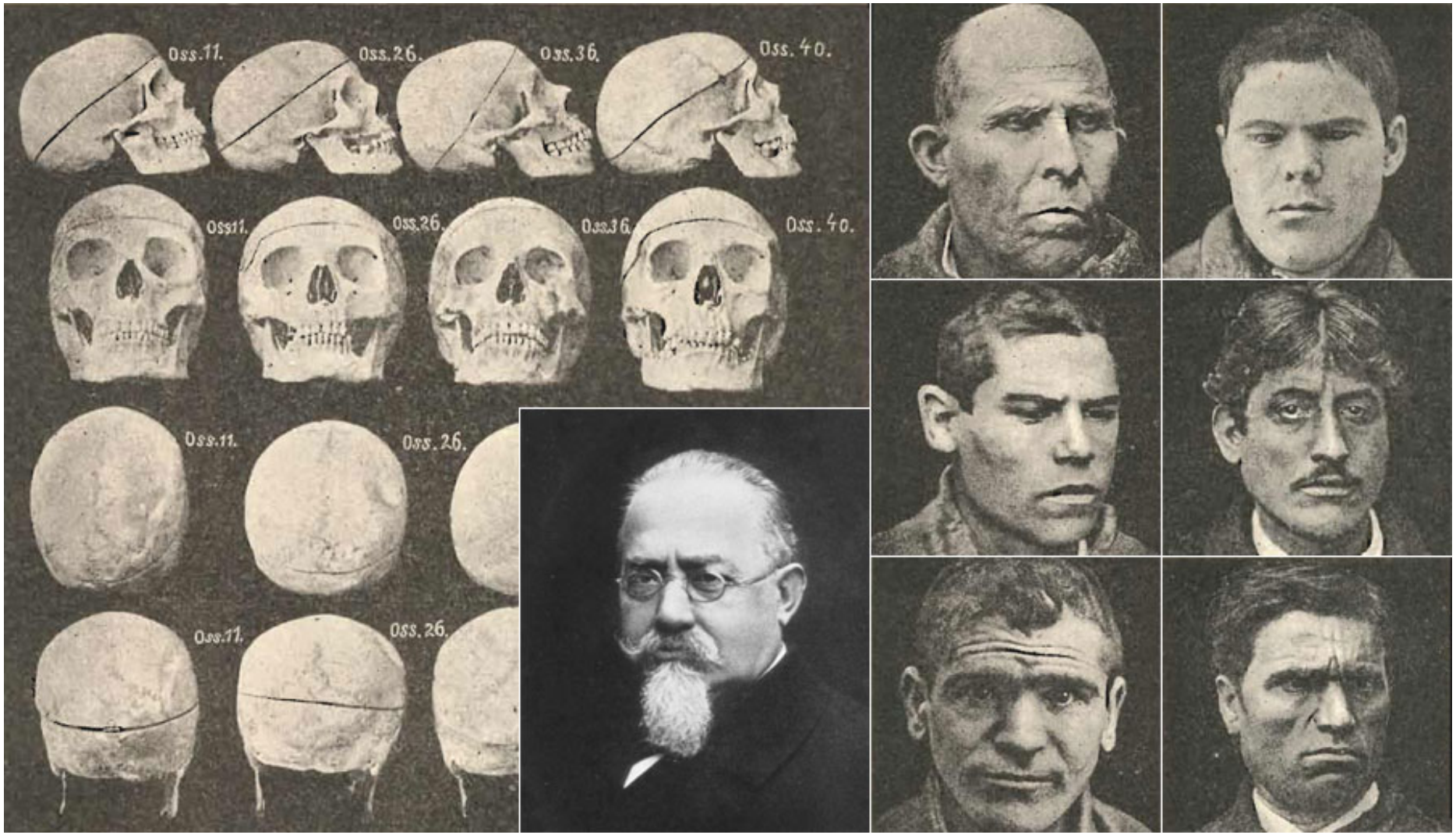
McCurrie et al. IVC 2018

Todorov in his book: *Face Value: the Irresistible Influence of First Impressions*

"Psychologists in the early twentieth century found little evidence for the accuracy of first impressions, but the past decade has seen a resurgence of physiognomic claims in scientific journals. We are told that it is possible to discern a person's political leanings, religious affiliation, sexual orientation, and even criminal inclinations from images of their face… A closer look at the modern studies shows that the claims of the new physiognomy are almost as exaggerated as those in the eighteenth and nineteenth centuries."

Social Implications of Criminality from Face Technologies

Lombroso and the Positivist School

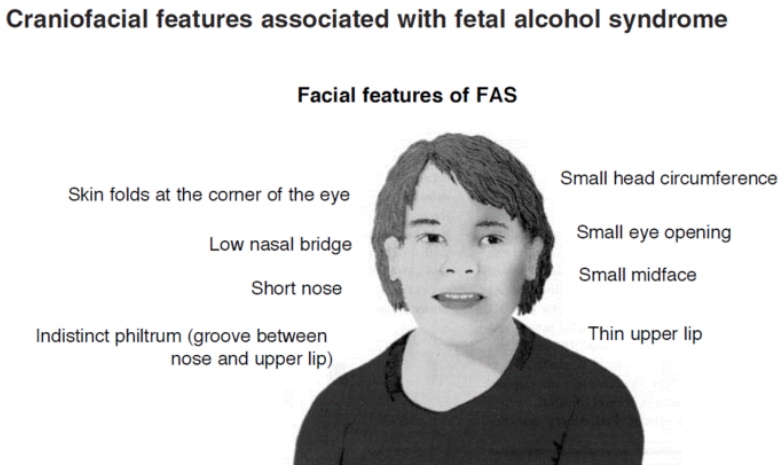


https://psychlite.wordpress.com/2015/04/09/origins-of-criminology-cesare-lombroso-and-the-face-of-crime/

General confusion about genetics, mental illness and anatomy

Mental illness is prevalent in prison populations

But common mood disorders and schizophrenia do not present with physical markers



Maybe fetal alcohol spectrum disorders?

Vast majority of criminals do not suffer from these disorders

Market Incentives for Bad Tech

- Large market for new law enforcement technology: $59.9B by 2025*
- Not all of it is legitimate

*https://www.asdreports.com/market-research-report-495298/law-enforcement-police-modernization-market

Image Credit: https://www.muckrock.com/news/archives/2017/nov/02/iacp-2017/

Unethical Applications

- Surveillance
- Predictive Policing
- Hiring

Mel McCurrie

Sam Anthony

CNNs for Subjective Face Attributes

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

Todorov Revisited…

Automatic perception of social face cues

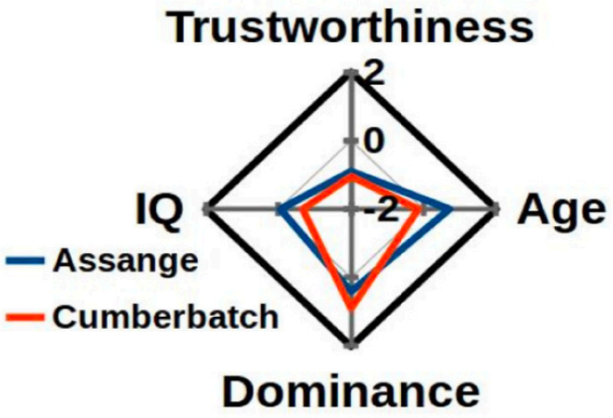
Attended characteristics of stimulus

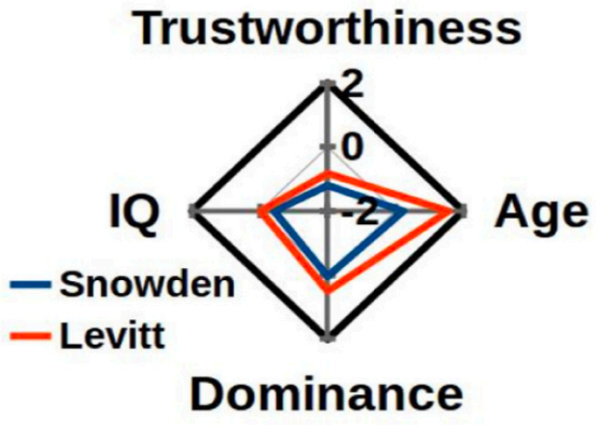
Type of social cue	**Extrinsic, deceptive, non-veridical**	**Intrinsic, truthful, veridical**
Emotional states (labile, instantaneous)	Duchenne, Ekman, etc.	Darwin, Duchenne, Ekman, Baron-Cohen, etc.
Personality traits (stable, consistent)	?	Todorov, Ambady

Automatic perception of social face cues

	Extrinsic, deceptive, non-veridical	**Intrinsic, truthful, veridical**
Trustworthiness, Dominance, IQ, Age	Pose, expression, glasses, hairstyle, lens length, image contrast, makeup, facial hair, facial jewelry, tattoos, dental work, lighting direction, gaze direction, pupil dilation	 Physiognomic cues

Subjective Attribute Predictions

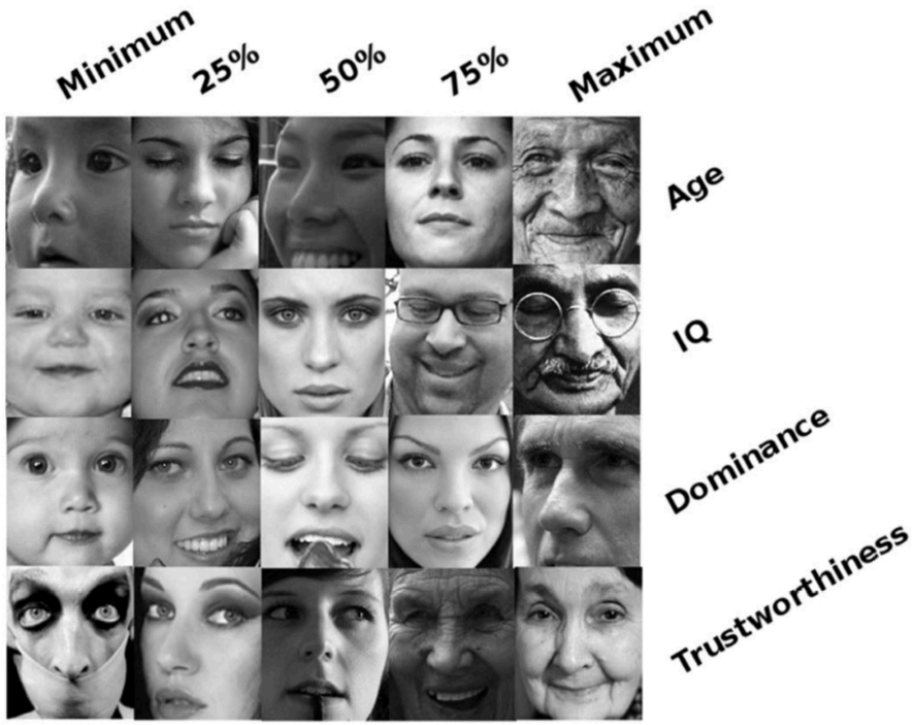


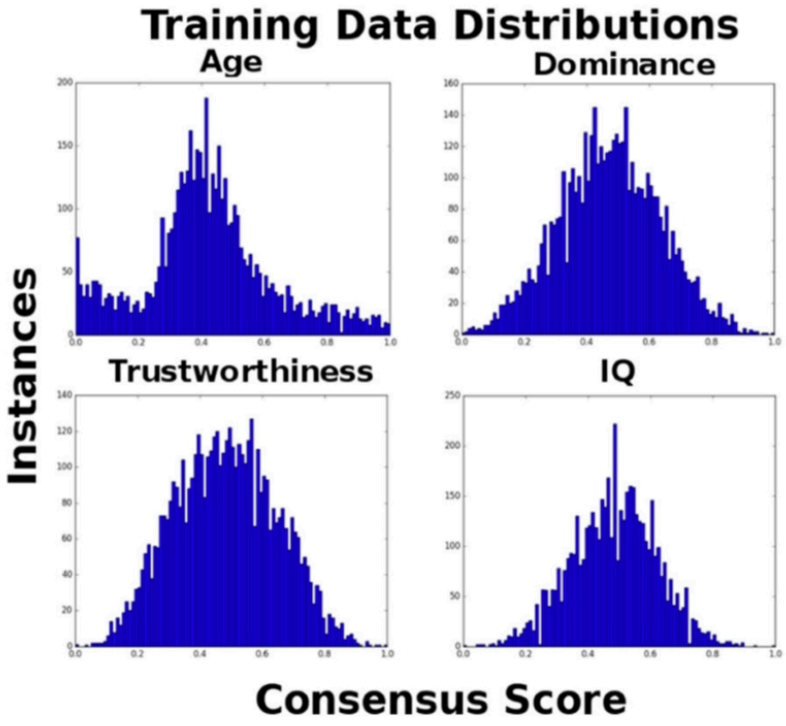


N.N. Oosterhof, A. Todorov, “The functional basis of face evaluation,” PNAS vol. 105, no. 32, 2008

Data collection via TestMyBrain.org

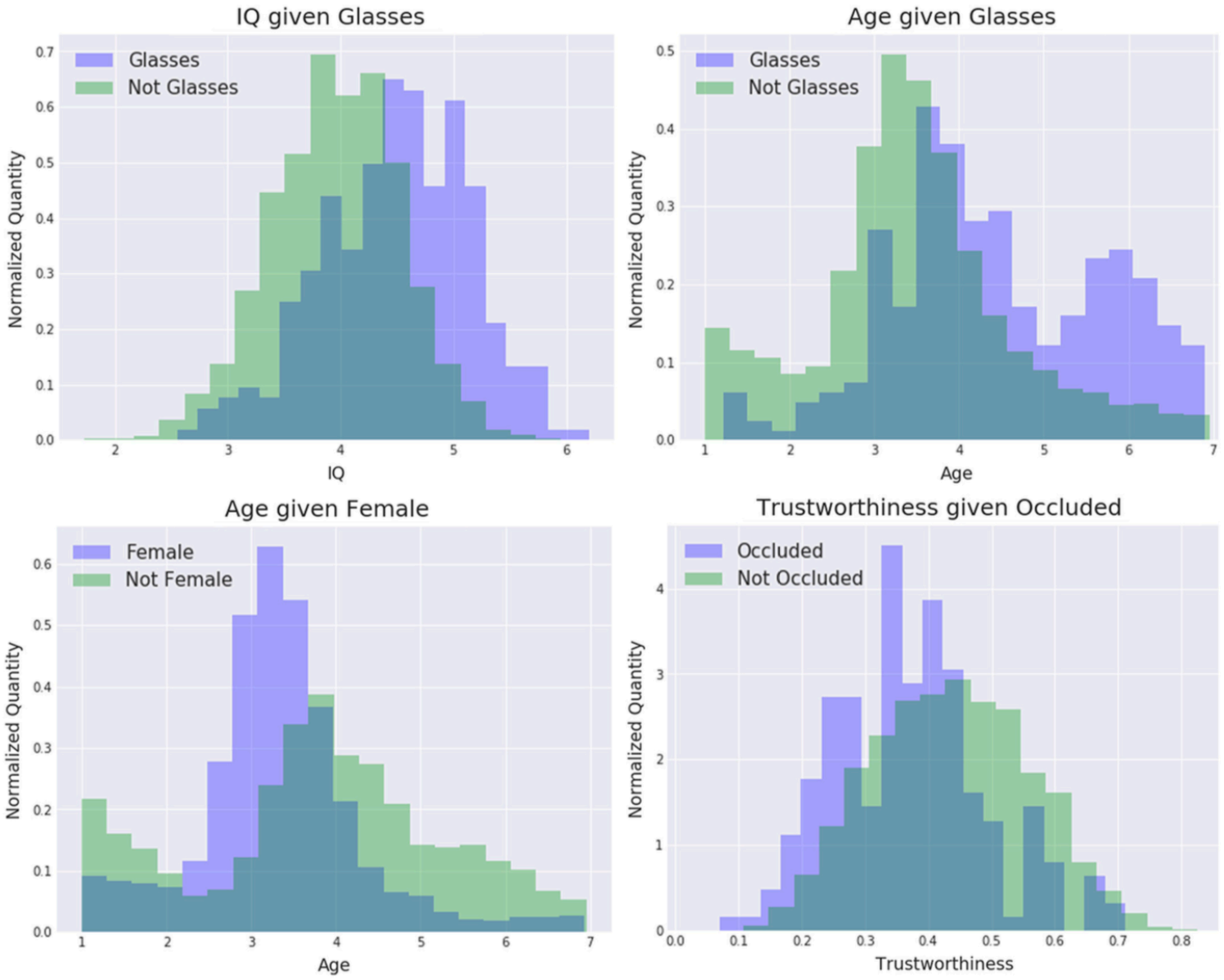
Data Distributions



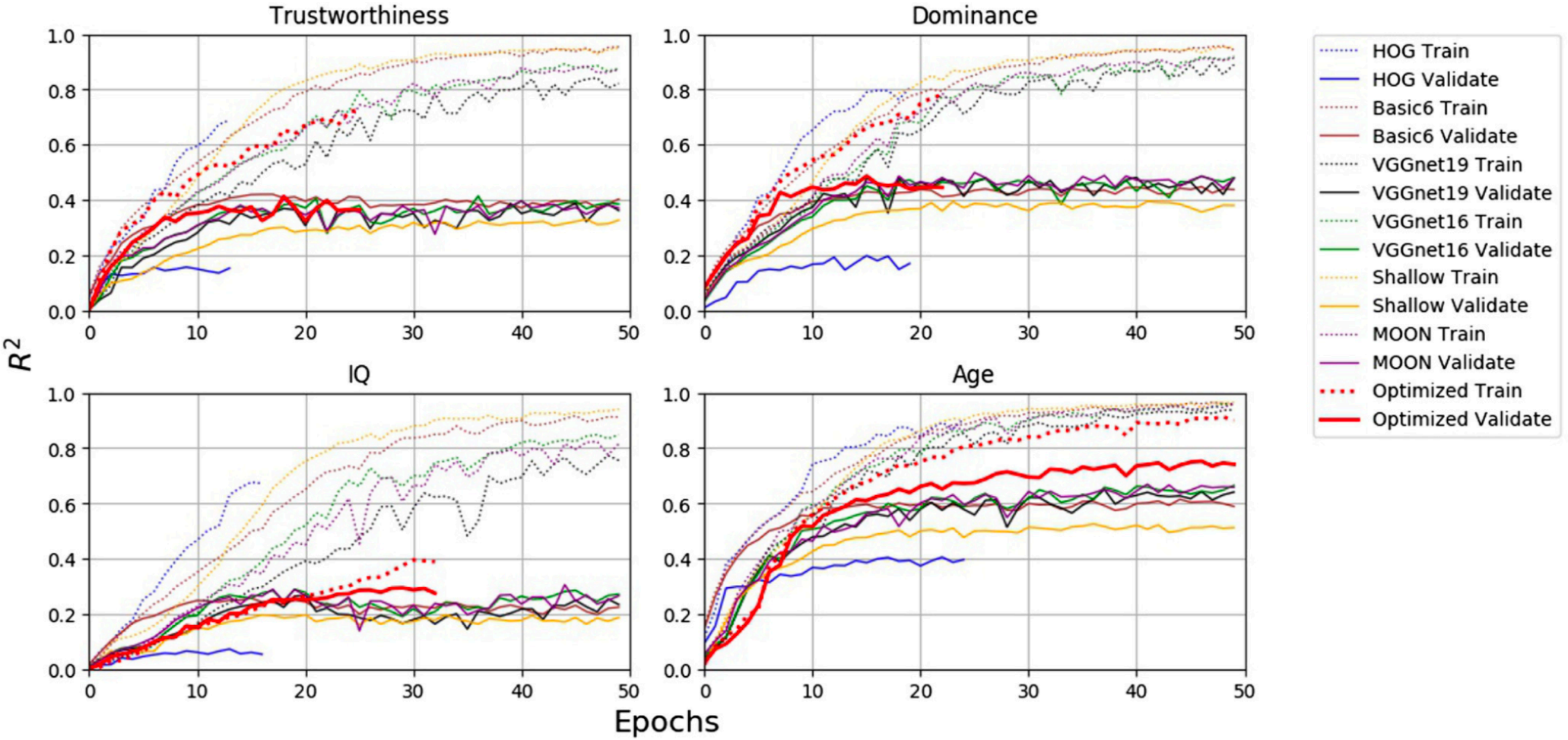


5040 Total Images

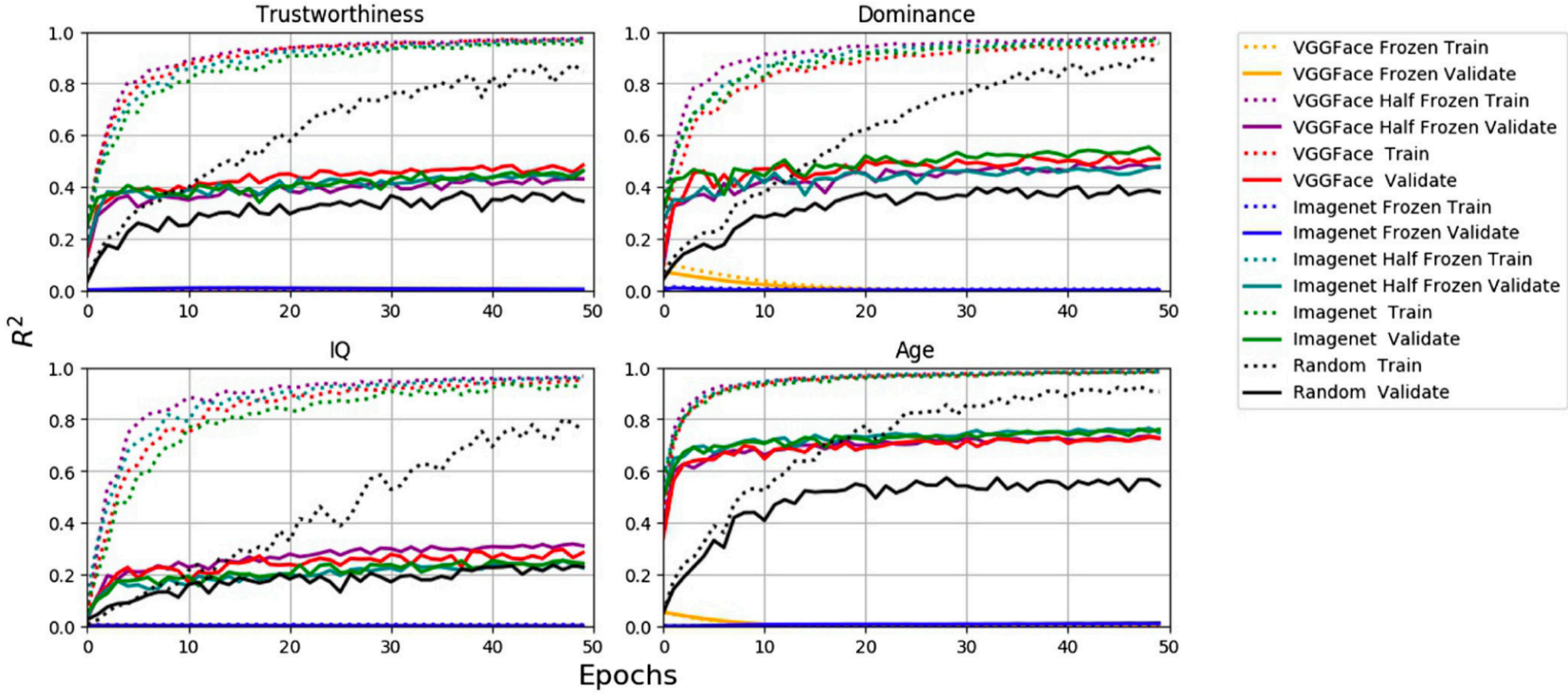
Dataset Patterns



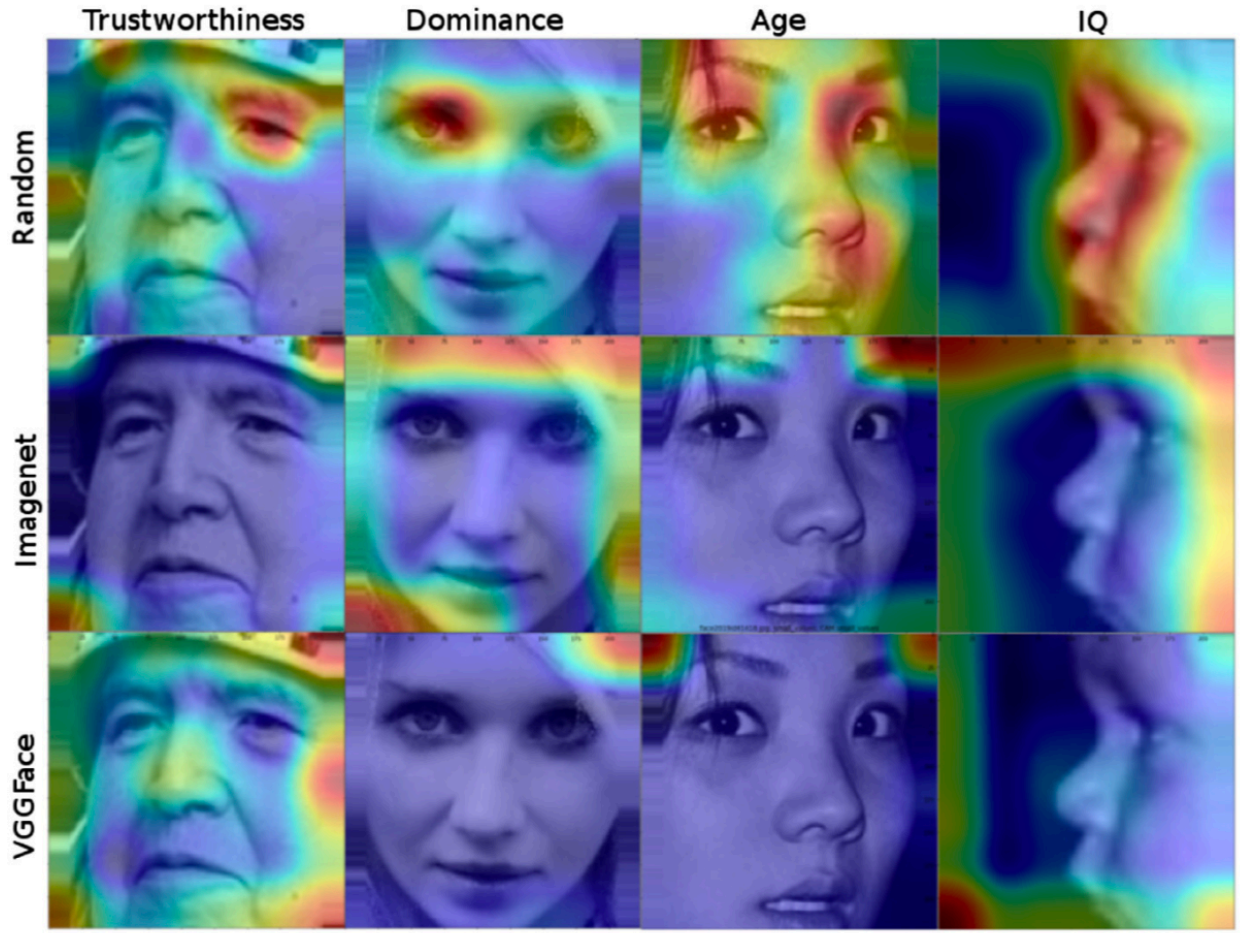
Which architecture is the best?



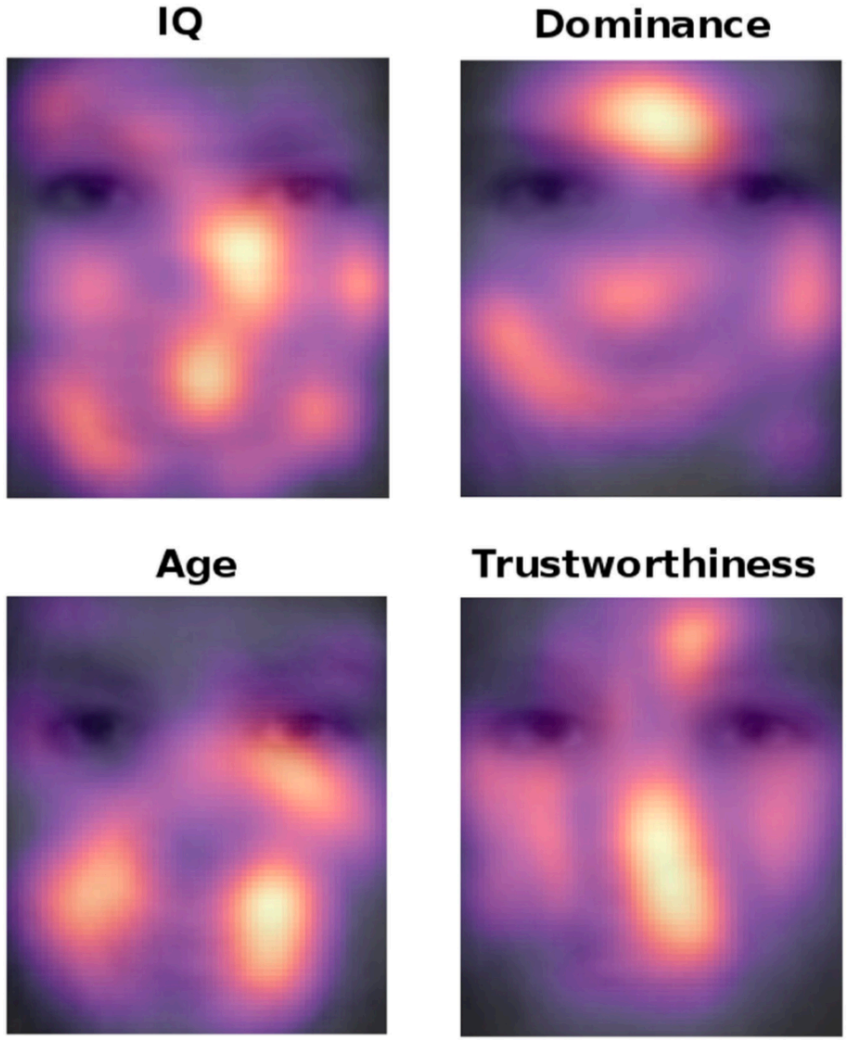
The dataset is small. Does transfer learning help?



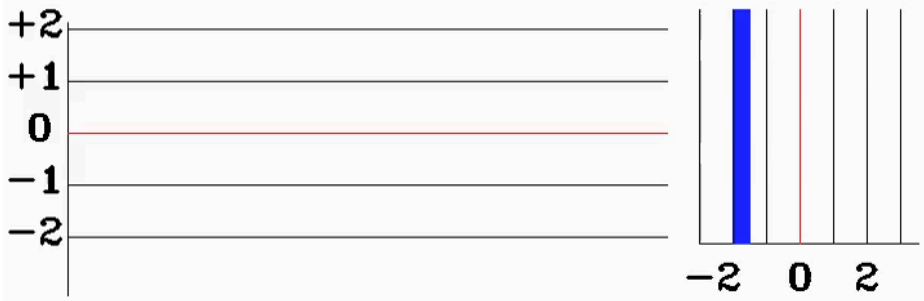
Regions of the face most important to the models (transfer learning saliency)



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



Trustworthiness



Ethical Applications

- Scientific Inquiry
- Robotics
- Human-Computer Interaction

What other attributes are interesting?