

Automated Image and Video Data Analysis with Python

University of Lucerne, Digital Skills

Dr. Natalia Umansky (University of Zurich)

Outline

1. Why do Images Matter?
2. Types of Existing Research with Images as Data
3. Available Automated Image Analysis Methods
 - what we'll cover
 - what we'll not cover
4. Good practices and limitations

Why do Images Matter?

People are more likely to pay attention to visuals (Dahmen, 2012)

FEATURE

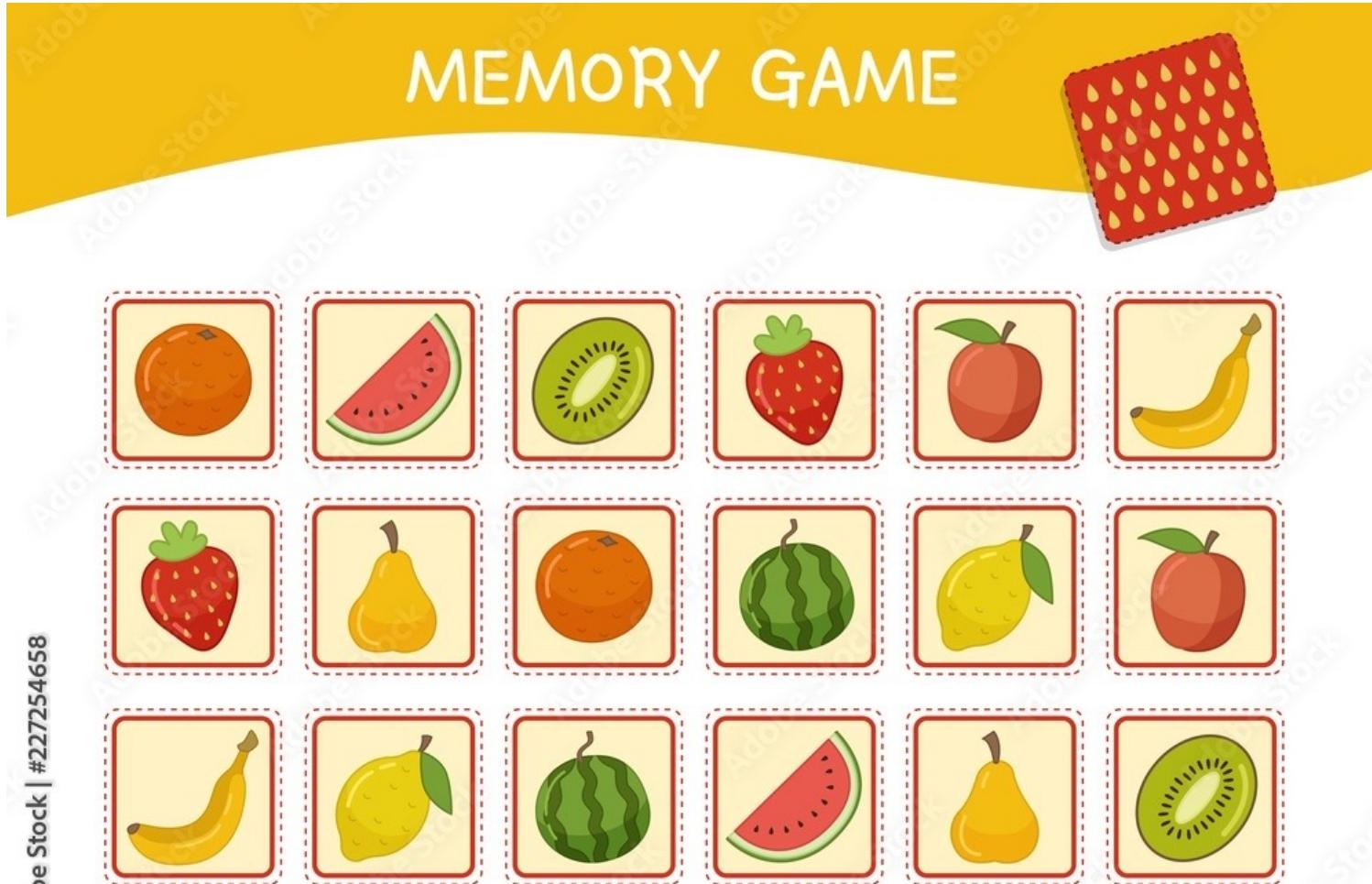
How America Got to 'Zero Tolerance' on Immigration: The Inside Story

Battles have raged within the Trump administration over family separations, ICE raids and the president's obsession with a wall. Together, they have remade homeland security.



Why do Images Matter?

People are more likely to recall information learned through visuals (Paivio et al., 1968)



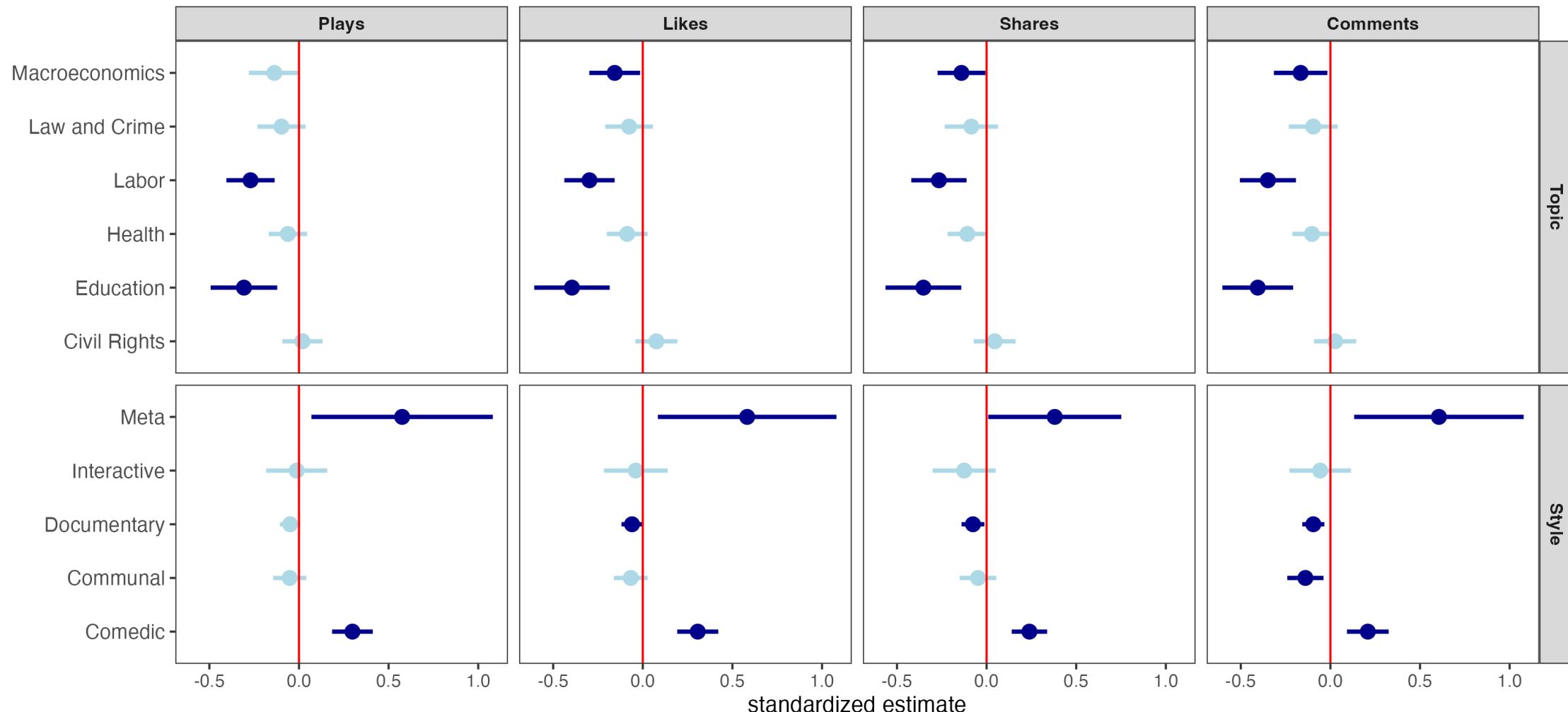
Why do Images Matter?

Visuals evoke stronger emotional reactions (Grabe & Bucy, 2009)



Why do Images Matter?

Image effects in politics (Umansky & Pipal, 2023)



Why do Images Matter?

Image effects in politics: images → framing → attitudes

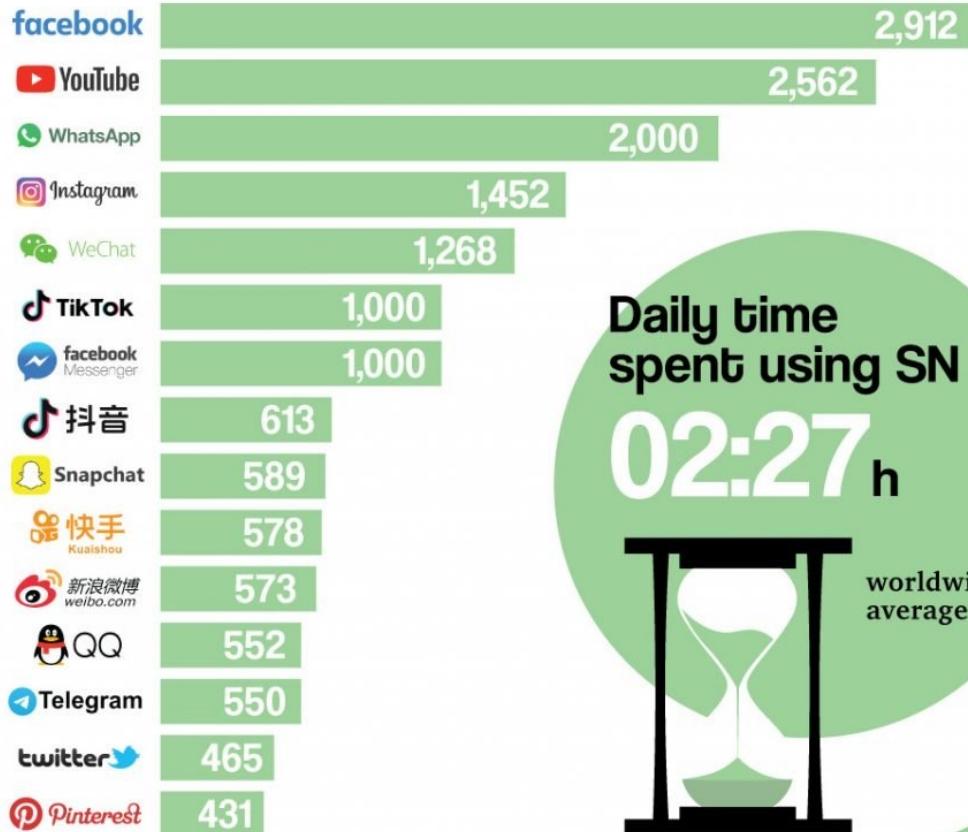


Why do Images Matter?

Images are more central than even in our life

Most popular social networks

(in millions)



Daily time spent using SN

02:27 h

worldwide average



2022

Why do Images Matter?

Images are more central than even in our life



A L E X Den Stoerste @Alex... · 6d ...

En respuesta a @MaxBoot

This guy agrees "we need more **content moderation**"



A L E X Den Stoerste @Alex... · 6d ...

En respuesta a @MaxBoot

This guy agrees "we need more **content moderation**"



Types of Existing Research with Images as Data

Causal Framework

Images as explanatory/independent variable

- Boussalis et al. (2021): How candidate emotional expressions in televised debate affect voting preferences?
- Casas and Webb Williams (2018): Which Black Lives Matter images mobilized more supporters?

Images as dependent variable

- Dietrich and Ko (2022): TV coverage of Dr. Fauci during the covid pandemic
- Michelle Torres (2023): How do different news organizations choose different pictures to accompany articles about Black Lives Matter?

Types of Existing Research with Images as Data

As a Measurement Strategy

- Images can contain information about electoral incidents and fraud (Callen and Long (2015); Cantú (2019))
- Images can help us identify and classify protest events (Zhang and Pan (2018), Won, Steinert-Threlkeld and Joo (2017))
- Night-time lights imagery as a proxy for economic development (many authors)
- Digitized historical maps as evidence of road quality variation (Hunziker et al (working paper))
- Videos/Images can help us measure cooperation in legislative politics (Dietrich 2020)

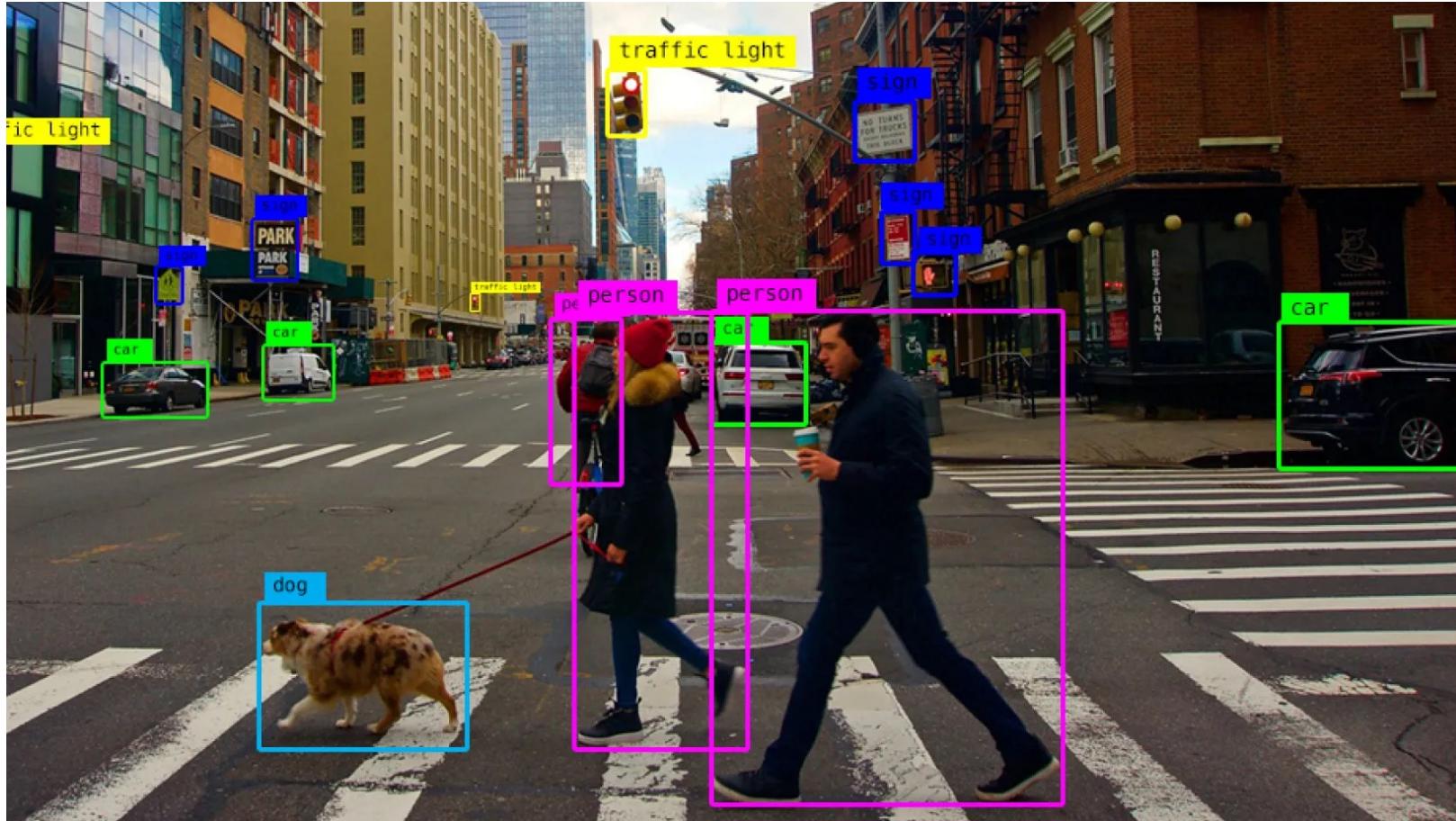
Types of Existing Research with Images as Data

Methodological contributions

- Methodological reviews (Webb Williams et al. 2020; Torres & Cantu 2021)
- Unsupervised clustering (Zhang Peng 2022; Casas et al.(working paper); Umansky et al. (working paper))
- Limitations & biases (Schwemmer et al. 2020)
- Extracting/leveraging aesthetic features (Peng 2021)

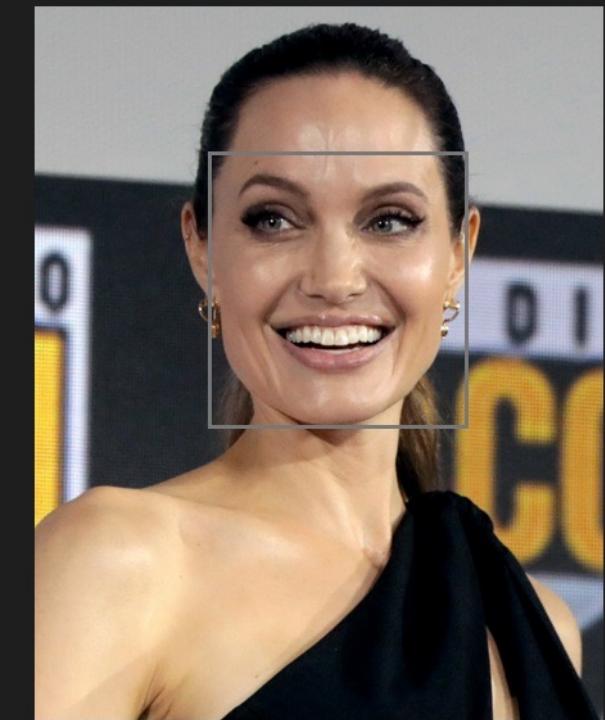
Available Automated Image Analysis Methods

Object detection & recognition



Available Automated Image Analysis Methods

Face detection & recognition



target: img1.jpg



#1

id: img4.jpg
distance: 0.205



#2

id: img2.jpg
distance: 0.234



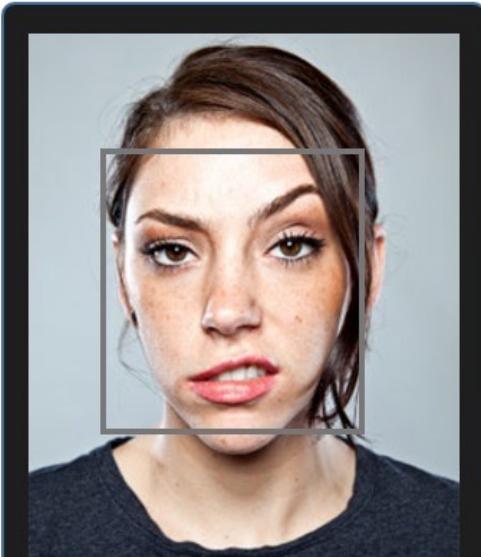
#3

id: img6.jpg
distance: 0.254

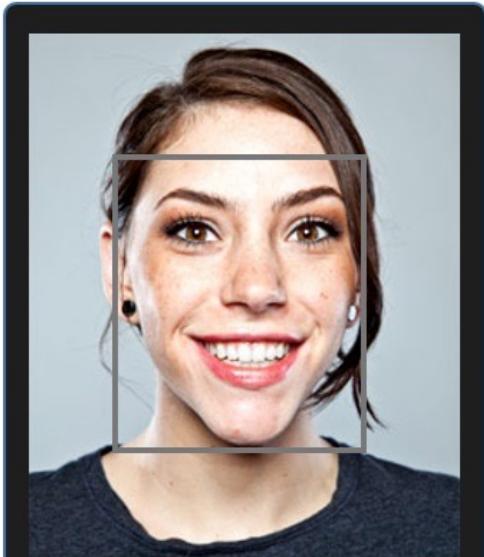
found

Available Automated Image Analysis Methods

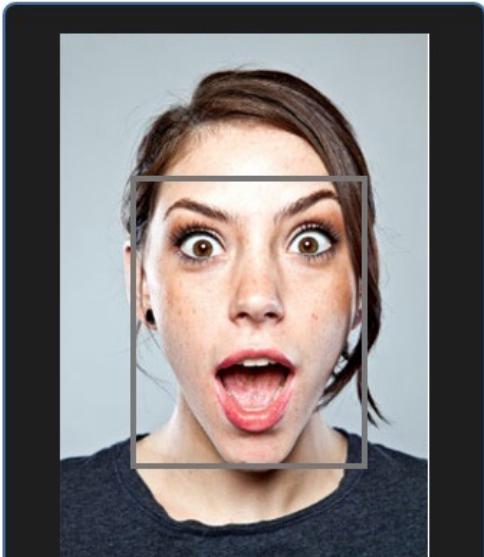
Face analysis



```
{  
  "age": 28.66,  
  "emotion": "neutral",  
  "gender": "Woman",  
  "race": "latino hispanic"  
}
```



```
{  
  "age": 29.27,  
  "emotion": "happy",  
  "gender": "Woman",  
  "race": "white"  
}
```



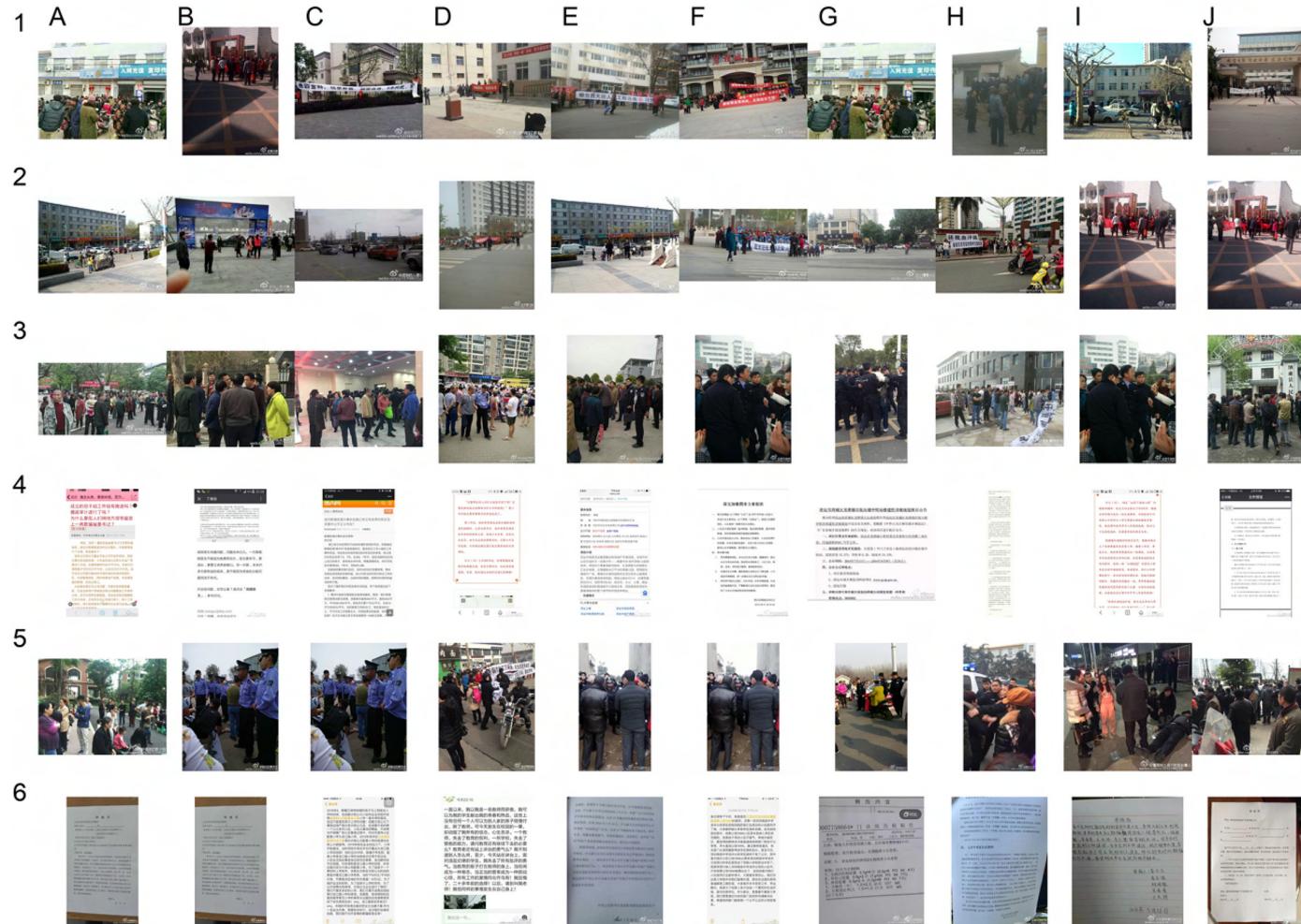
```
{  
  "age": 29.27,  
  "emotion": "surprise",  
  "gender": "Woman",  
  "race": "white"  
}
```



```
{  
  "age": 29.74,  
  "emotion": "neutral",  
  "gender": "Woman",  
  "race": "white"  
}
```

Available Automated Image Analysis Methods

Unsupervised Clustering (Zhang & Peng, 2022)



Available Automated Image Analysis Methods

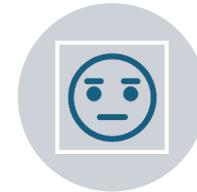
And many others...



Text extraction
(OCR)



Caption generation



Sentiment analysis
(evoked emotions)



Visual aesthetics
analysis



etc...

A large orange circle is positioned on the left side of the slide, covering approximately one-third of the vertical space.

In this
workshop
we'll focus
on...

Some basic image (& video)
manipulation/processing

Supervised image
classification

Unsupervised image
classification

Face
detection/recognition/analysis

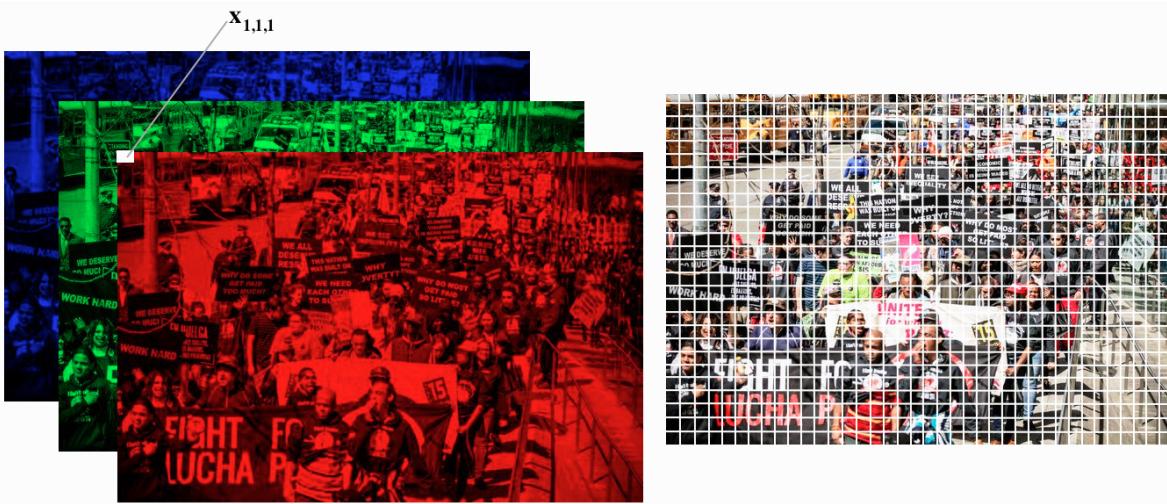
Convolutional Neural Networks

The Basics



(1) Images as inputs:

3-dimensional matrices (width x height x depth)

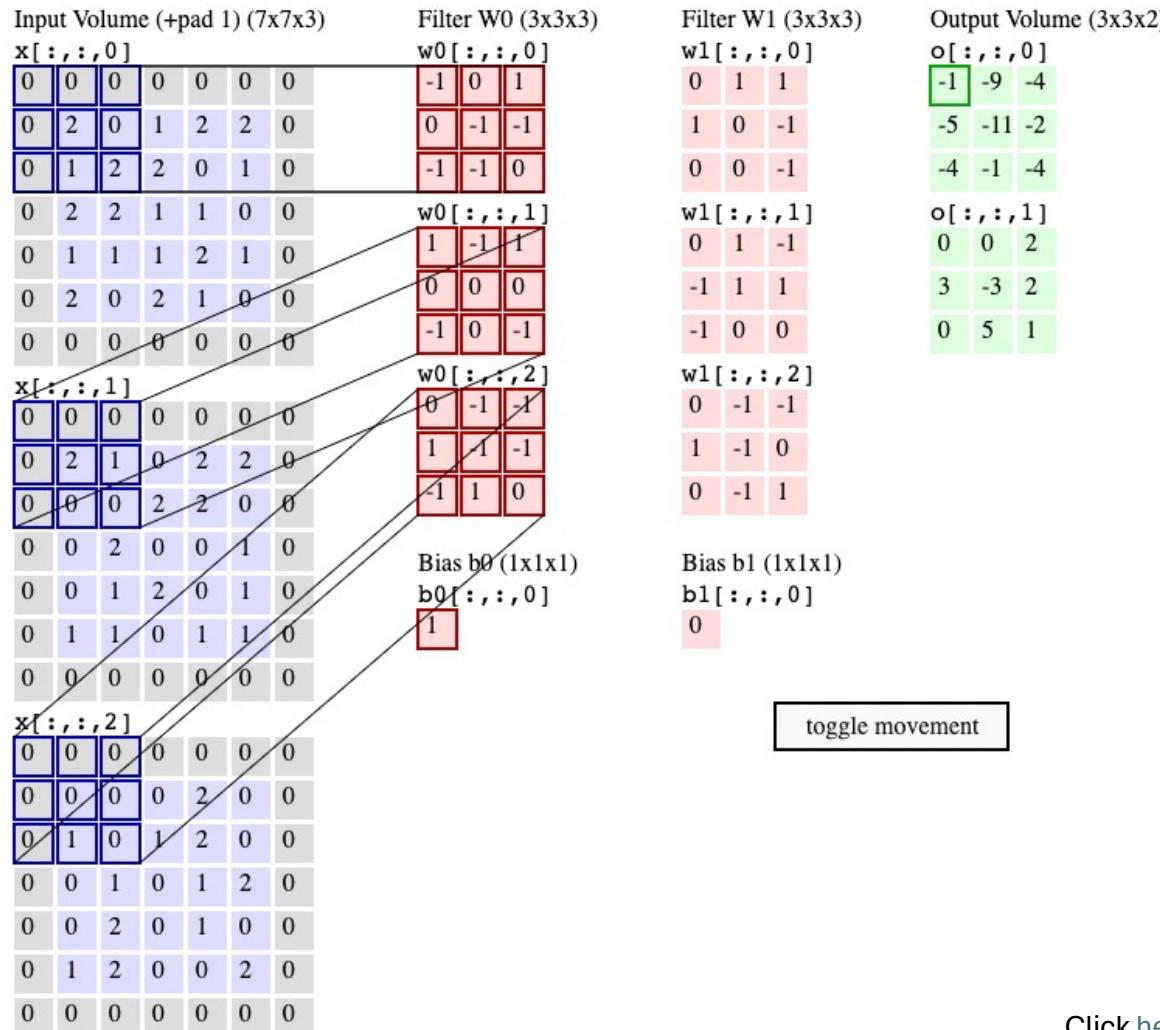


$\mathbf{X} =$

$$\begin{bmatrix} x_{111} & x_{112} & \dots & x_{11n} \\ x_{121} & x_{122} & \dots & x_{12n} \\ x_{131} & x_{132} & \dots & x_{13n} \\ x_{141} & x_{142} & \dots & x_{14n} \\ \vdots & \vdots & & \vdots \\ x_{1n1} & x_{1n2} & \dots & x_{1nn} \end{bmatrix}, \begin{bmatrix} x_{211} & x_{212} & \dots & x_{21n} \\ x_{221} & x_{222} & \dots & x_{22n} \\ x_{231} & x_{232} & \dots & x_{23n} \\ x_{241} & x_{242} & \dots & x_{24n} \\ \vdots & \vdots & & \vdots \\ x_{2n1} & x_{2n2} & \dots & x_{2nn} \end{bmatrix}, \begin{bmatrix} x_{311} & x_{312} & \dots & x_{31n} \\ x_{321} & x_{322} & \dots & x_{32n} \\ x_{331} & x_{332} & \dots & x_{33n} \\ x_{341} & x_{342} & \dots & x_{34n} \\ \vdots & \vdots & & \vdots \\ x_{3n1} & x_{3n2} & \dots & x_{3nn} \end{bmatrix}$$

(2) Convolutional layers:

weights (filters) are not connected to the whole input volume:
convolution.



Click [here](#) for a full visualization by the Stanford cs231 folks.

Convolutional Neural Nets for Computer Vision

Some new terminology...
and more
hyperparameters

- **Input volume:** a 3-dimensional input
- **Convolutional layer:** a 4-dimensional parameter layer where convolutional filters are applied to the input volume; of size $F \times F \times N \times K$ where F is the width and height of the filter, N is the number of filter dimensions, and K is the number of filters → $3 \times 3 \times 3 \times 2$ in the previous example
- **Stride:** the number of pixels we move the filter at a time. This is 2 in the previous example
- **Zero-padding:** adding zeros around the input border (often done to avoid deforming input images)
- **Pooling layer:** a layer where we reduce the size the output of a convolutional layer. From $224 \times 224 \times 3 \times 64$ to $112 \times 112 \times 3 \times 64$ for example.
- **Fully connected layer:** a layer of weights that is connected to the whole input volume. These are usually at the end of a network.
- **Softmax:** a multi-class classifier. This is basically a multinomial logit model that uses the output of the last fully-connected layer to predict the final classes of interest

Convolutional Neural Nets for Computer Vision

This is how a ConvNet looks like

