



# Images-as-data:

## *An Intro to Computer Vision for Political Research*

Emil D. Bello-Pardo, PhD  
American University Institute for Data Science  
January 9, 2026

# Overview

- **Images and political research**
- **Studying images at scale: images-as-data**
  - Neural networks overview
  - Convolutional Neural Networks
  - Autotaggers
- **Applications of computer vision**
  - Transfer learning using social media data
  - Autotagger example
  - A word about LLMs
- **Conclusions**

# Images-as-data

# Why study images?



<https://www.smithsonianmag.com/smart-news/woman-iconic-v-j-photo-died-age-92-heres-her-story-180960435/>



[https://www.npr.org/2010/01/  
18/122701268/i-have-a-drea  
m-speech-in-its-entirety](https://www.npr.org/2010/01/18/122701268/i-have-a-dream-speech-in-its-entirety)



Saul Loeb | AFP | Getty Images



<https://www.cnn.com/2019/04/12/us/crying-girl-john-moore-immigration-photo-of-the-year/index.html>



<https://unsplash.com/photos/jSKjkV4Oc5Q>



<https://www.vox.com/culture/360711/trump-fist-pump-photo-explained-expert-media-savvy-politics>

**TL;DR:**

Images are powerful  
tools that convey  
significant amounts of  
information

# On the power of images...

- People pay more attention to images (Dahmen 2012) and are processed more quickly than text (Graber 1990, 2012, Whitehouse et al 2006).
- Images evoke emotions that influence evaluations (e.g. Wright and Citrin 2011, Renshon et al 2015), so they can be constructed to persuade and manipulate (Nahon-Nerfatty, 2018).
- Images can increase attention and online diffusion of political movements (e.g. Casas and Webb Williams 2018) and offline mobilization (Geise et al 2021).
- Images can also capture text in them that can be analyzed after being processed – ex: election results in Mexico (Torres-Pacheco and Cantu, 2022)
- In multimodal environments, images and text have a symbiotic, mutually-amplifying relationship (Geise and Baden 2015; Lee and Ho 2018)

# Previous limitations on studying images...

- **Image-based analyses are largely absent from political academic research, particularly on disinformation (Bucy and Joo 2021)**
- **Images are not easily studied quantitatively, but it's now possible**
  - Previous techniques for content analysis relied on hand-coding
  - With the growth in computing power, however, it is possible to use computational methods to analyze them at scale.

# How can we study images at scale?

- **Images are easily accessible and collectable through APIs**
  - Using APIs allows users to collect data
  - One of the pieces of information that the old Twitter API provides is whether a Tweet contains an image and, if so, it provides a link through which the image can be downloaded.
  - Disclaimer: check with IRB before collecting any data!

# Example on images: Casas and Webb Williams (2018)

- Study diffusion of Black Lives Matter online and the role of images
- They argue that images evoking anger, fear, and enthusiasm are mobilizing for online engagement
- But how do the authors get there?
  - Combining Twitter data with images-as-data strategy!
  - Authors coded ~9,500 images using undergrads and AMT workers to get labels based on emotions of interest, then analyzed the data

# Example on images: Torres-Pacheco and Cantu (2022)

- Methods paper showcasing CNNs to study political phenomena.
- Election results:
  - “In the case of Mexico, capturing vote results with a CNN approach may significantly decrease accidental errors when adding up the votes, which actually occurs in almost 40% of the tallies in the country (Challú, Seira and Simpser, 2018)”
- Police violence:
  - Front pages of 450 U.S. newspapers after the El Paso 2019 shooting, with the aim of studying the visual frames that communicate news about gun violence.
  - “This exercise shows that, when reporting the same event, newspapers are more likely to show pictures with heavily armed police in states where gun popularity is high, and vice versa in states where such popularity is low.”

# Example on images: Casas and Webb Williams (2018)

Figure 6: Predicting attention to BLM over range of evoked emotions

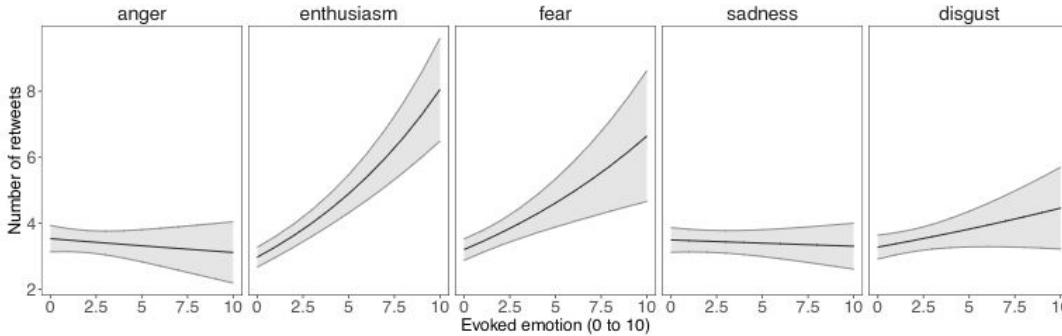
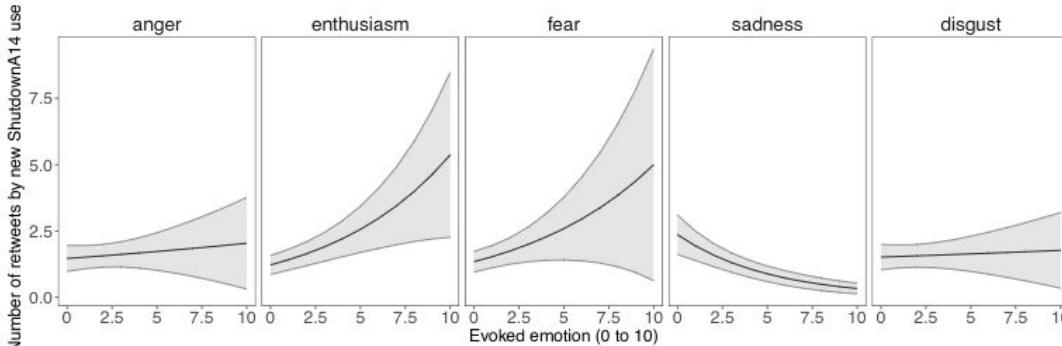


Figure 7: Predicting diffusion of ShutdownA14 over range of evoked emotions



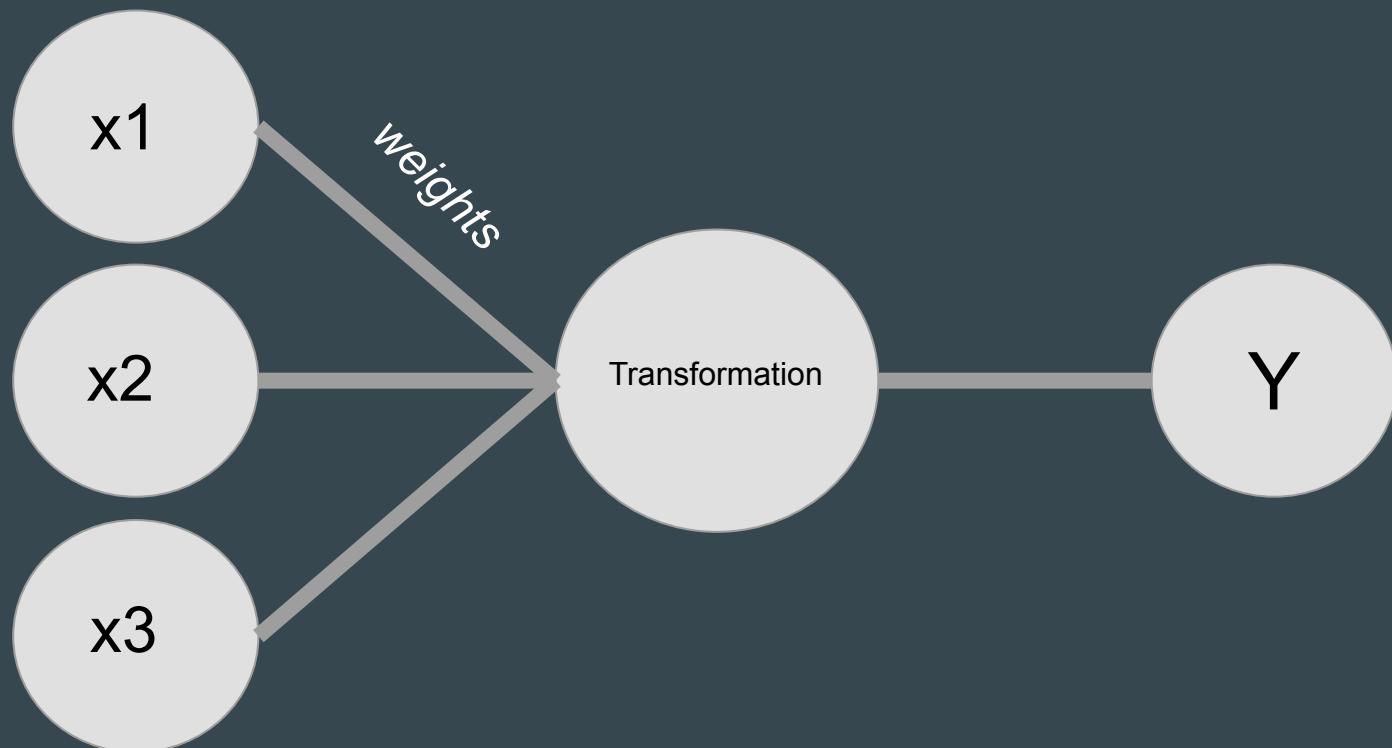
<https://journals.sagepub.com/doi/10.1177/1065912918786805>

# Images at scale: Computer Vision Methods

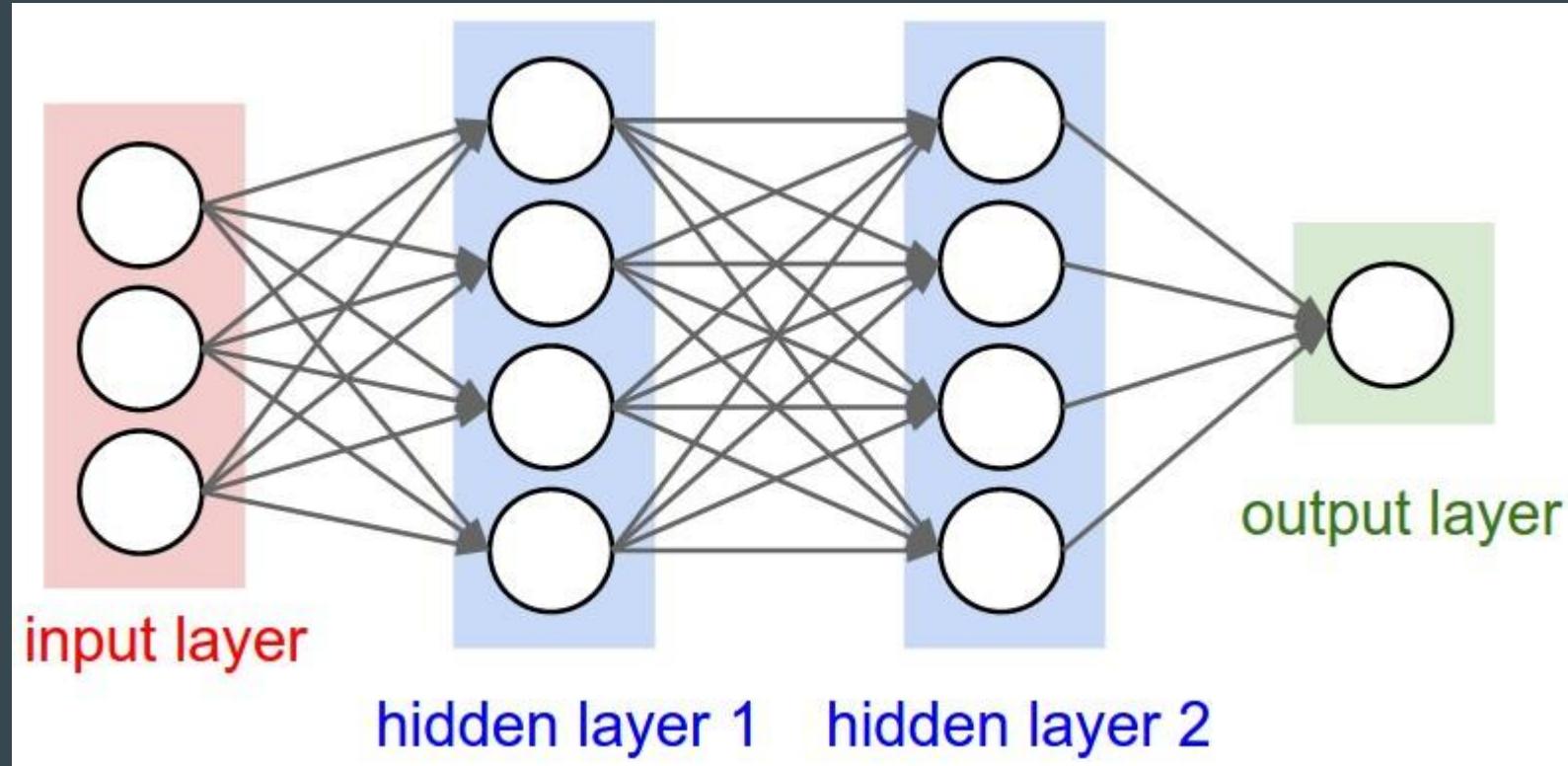
# Neural Nets Explained

- Neural networks use matrix algebra to transform inputs through multiple layers producing a final output.
  - Today we'll focus on intuition and examples, not the math/programming. If you want to go in depth, [see here](#)
- Black Box Model : The coefficients and results of internal layers are *high-dimensional and difficult to interpret directly* , only the final prediction is observable and most interpretable.
- These networks works like a human brain, learning by optimizing parameters through a process called training.

# A Simple One-Layer Network



# A Multi-Layer Network



# Training a neural net

- **Phases:**
  - Training
  - Validating
  - Testing
- **You will need:**
  - Training data that comes with labels
  - Validation data with labels, held off during training.
  - Testing data that comes with labels, held off from the network while it's being trained and validated

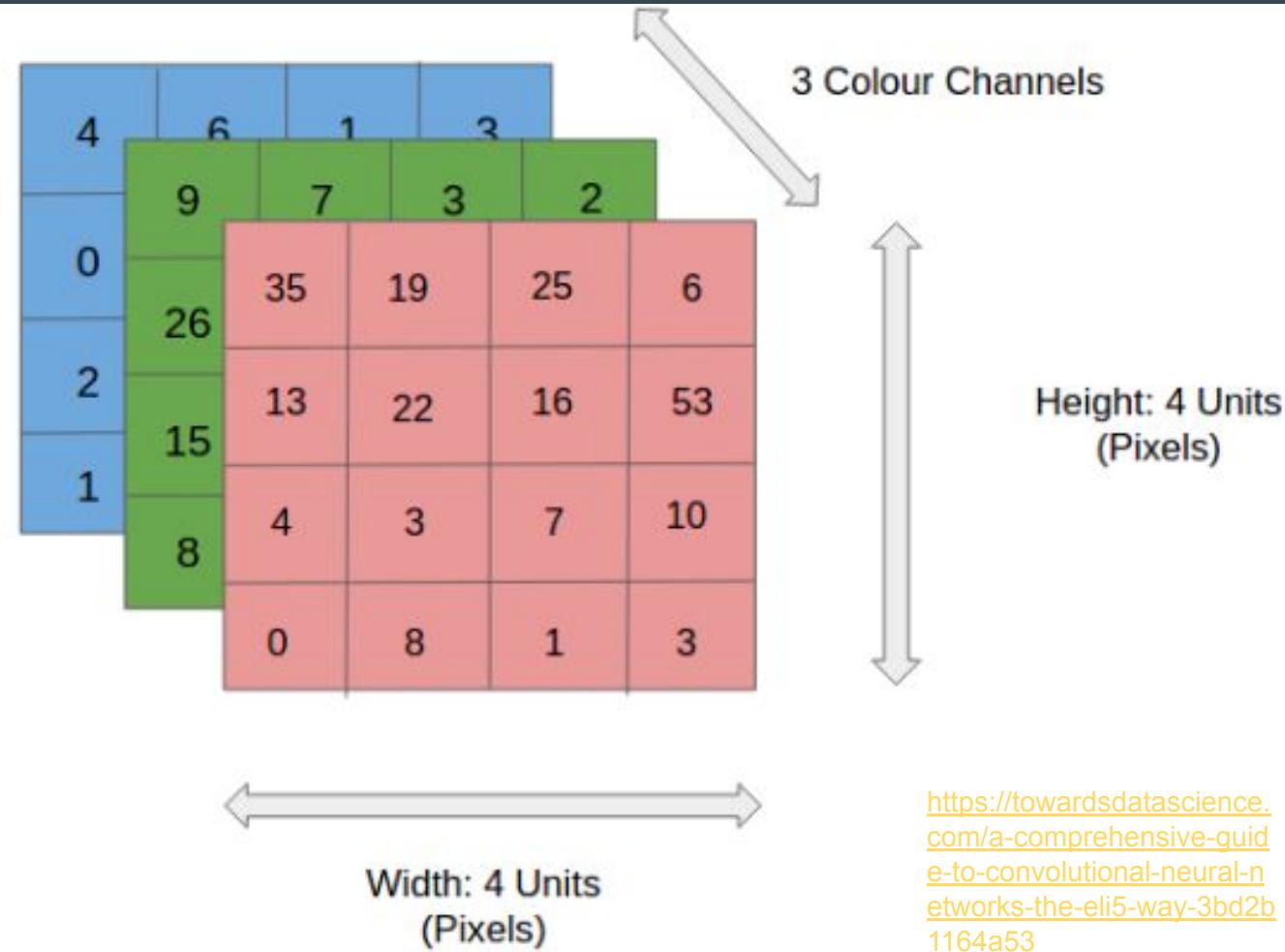
# Example: Classifier about dog/cat text

- Let's say you have a classifier (neural net) that has been trained to recognize whether the subject of a text is a dog or a cat.
- A neural network accurately trained on enough textual data will be able to predict whether the following sentences fall into each of those categories:
  - *The dog ran through the backyard.*
  - *Free from his masters, the feline sits on his throne in the room.*
- This same classification idea can also apply to pictures!

But images are a bit  
more complicated...

# Why images are more complicated...

- Images are made out of thousands to millions of pixels, each with a value.
- Individual pixels have no meaning of their own.
  - In text, a word carries meaning. In images, pixels do not.
- Meaning comes from relationships between pixels, nearer ones are more likely to be related to each other than distant ones.
- This is much more complicated, so needs specialized neural net.



<https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

# Computer Vision

It's been applied before in academic research, and it is extensively used in industry

- Convolutional Nets are most often used!
-

# Convolutional Neural Network: ReLu example

0	0	0	0	0	0	0	...
0	156	155	156	158	158	158	...
0	153	154	157	159	159	159	...
0	149	151	155	158	159	159	...
0	146	146	149	153	158	158	...
0	145	143	143	148	158	158	...
...	...	...	...	...	...	...	...

Input Channel #1 (Red)

0	0	0	0	0	0	0	...
0	167	166	167	169	169	169	...
0	164	165	168	170	170	170	...
0	160	162	166	169	170	170	...
0	156	156	159	163	168	168	...
0	155	153	153	158	168	168	...
...	...	...	...	...	...	...	...

Input Channel #2 (Green)

0	0	0	0	0	0	0	...
0	163	162	163	165	165	165	...
0	160	161	164	166	166	166	...
0	156	158	162	165	166	166	...
0	155	155	158	162	167	167	...
0	154	152	152	157	167	167	...
...	...	...	...	...	...	...	...

Input Channel #3 (Blue)

-1	-1	1
0	1	-1
0	1	1

Kernel Channel #1

1	0	0
1	-1	-1
1	0	-1

Kernel Channel #2

0	1	1
0	1	0
1	-1	1

Kernel Channel #3



308

+

-498

+

164

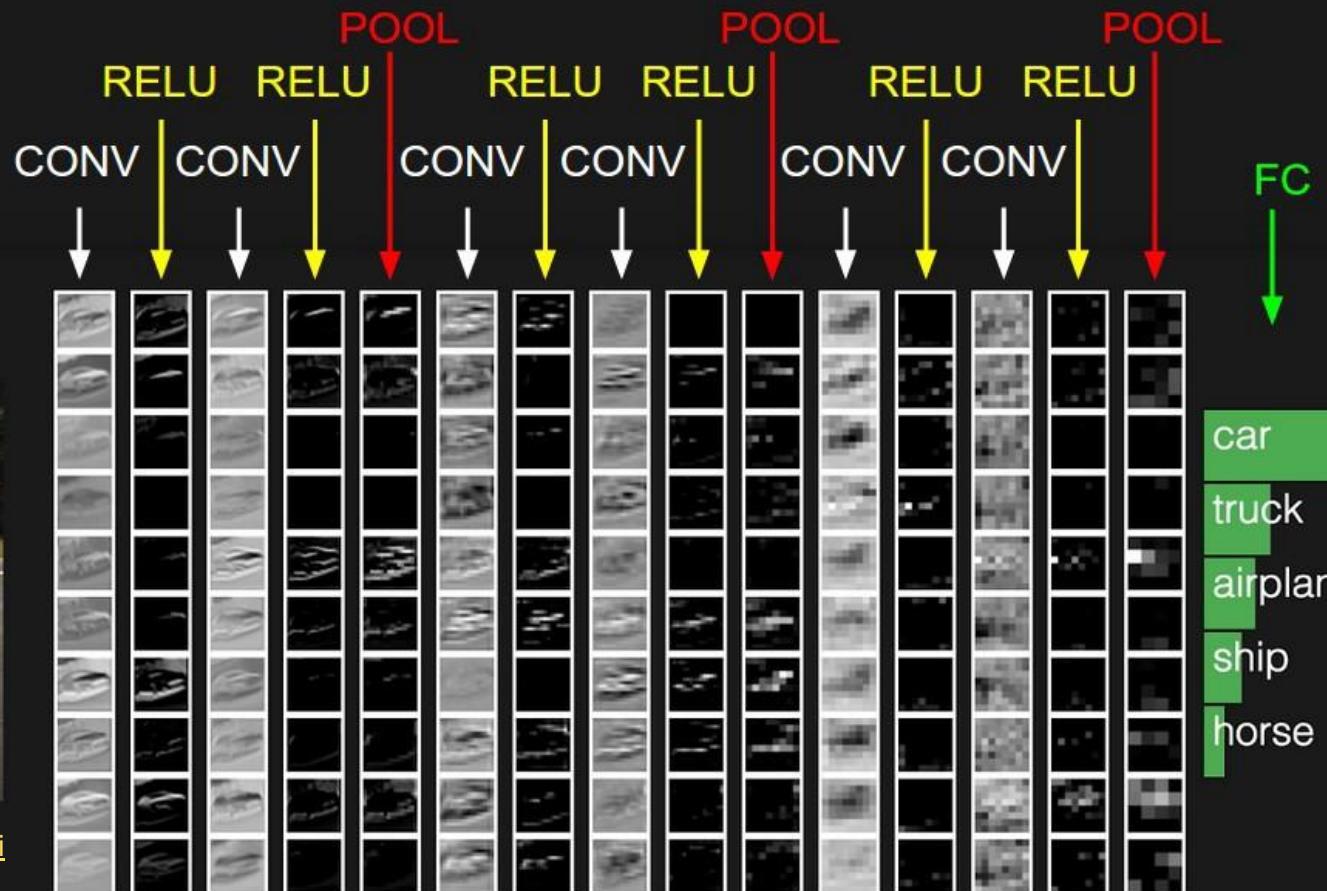
+ 1 = -25



Bias = 1

-25				...
				...
				...
				...
...	...	...	...	...

# Convolutional Neural Network



# Tools that allow automatic analysis of images

- **Machine learning: convolutional neural networks (CNN)**
  - They are very hard to train and calibrate from scratch, need a lot of labeled images to create a new CNN, and are time-consuming and resource-intensive to build.
  - However, we can do transfer learning and re-train the last layer of a CNN that has already been trained so you don't have to do the entire process from scratch(e.g. ResNet18 or VGG16).
  - There are various tools -- pytorch, keras, tensorflow.

# Application 1: Transfer learning on pictures from Venezuelan public officials' Tweets\*

\* Proof of concept

# Transfer Learning: Retraining ResNet 18 on political figures

- Tweets from 145 Venezuelan MCs in May 2019
- Randomly selected 1,000 pictures from these tweets.
- Retrained ResNet18 to recognize Nicolas Maduro, Juan Guaido, and crowds in these 1,000 pictures.

# Transfer Learning: Retraining ResNet 18 on political figures

- Retrained ResNet18 to recognize Nicolas Maduro, Juan Guaido, and crowds.
- Training data (yes, hand-coded)
  - Maduro = 40 images
  - Guaido = 15 images
  - Crowds = 80 images

# Transfer Learning: Retraining ResNet 18 on political figures

- Retrained ResNet18 to recognize Nicolas Maduro, Juan Guaido, and crowds.
- Testing data (yes, hand-coded too...)
  - Maduro = 10 images
  - Guaido = 10 images
  - Crowds = 10 images

# Transfer Learning: Retraining ResNet 18 on political figures

- Retrained ResNet18 to recognize Nicolas Maduro, Juan Guaido, and crowds.
- Applying this retrained CNN on 1,000 pictures off Twitter.

**Findings: The CNN became highly accurate in test data!  
(though I used very little test data, should have had more)**

- Correctly recognized Guaido 10/10 times
- Correctly recognized Maduro 10/10 times
- Correctly recognized crowds 10/10 times

# Examples: Guaidó test data



Prediction = 0.9997



Prediction = 0.7103

# Examples: Maduro test data



Maduro = 0.9978



Prediction = 0.9905

# Findings: Retrained CNN processes 1,000 pictures

- Recognized Guaido 32% of the time
- Recognized Maduro ~ 1% of the time
- Recognized crowds as top category 67% of the time

# Examples: OK predictions



Crowd = 1.00



Guaido = 0.42  
Crowd = 0.54

# Examples: Need more training!

INFORMACIÓN

**PAÍS**

DOMINGO 02 de Abril de 2017 -- /3

## Gracias Venezuela

**HENRY RAMOS ALLUP**

**Sin Censura**

Gracias, conciudadanos sin militancia partidista, por haber confiado en nosotros y participado en torrente validando a Acción Democrática que no es sólo de quienes militamos en este partido indestructible, sino perteneciente de la mejor Venezuela. Gracias a quienes sin dudar ni temblar creyeron en nosotros su fe por creer que somos un instrumento eficiente para salir de esta tragedia y lograr el futuro grande que nos merecemos. Gracias por el valiente y solidario trabajo fraternal de los profesionales y quando amigos de la diáspora que desde todo el mundo sienten y padecen la tragedia de la Patria adolorida cuya sanación depende de todos nosotros. Gracias a las mujeres y hombres de toda condición que con su trabajo y su dedicación contribuyeron a la gran validación para expresar sus deseos de cambio democrático y progreso. Iguales a las que hacen cotidianamente para adquirir los menudos alimentos y medicinas que se pueden encontrar. Gracias a los voluntarios que con su trabajo de buena fe que hicieron un alto en sus reparos y vinieron a validarlos entregándonos una incomparable lección del respaldo y solidaridad. Gracias a quienes ni remotamente creíamos que podían validar por la fuerza Democrática lo que sucedía en su país con apoyo de interesarado. Gracias a los hijos, nietos y parentes de nuestros queridos fundadores desaparecidos que supieron el vacío y multiplicaron su presencia. Gracias a nuestra maquinaria silenciosa y eficiente, que sin recursos materiales pero con energía infinita, esparcida a lo largo y ancho de toda Venezuela, fue capaz de entusiasmarse y entusiasmar millones por encima de todas las desdades y males y convencer a miles de personas a subirse a una fiesta de equipo imbatible con misticismo, activismo, disciplina y organización hoy es capaz de repetir los prodigios que la hicieron mítica. Gracias a los activistas y sobre todo a los de la Juventud del Partido que han hecho cargo de la reivindicación y apoderado de la fuerza que se contagió al evento la alegría y energía que los distinguen. Gracias a Venezuela que aceptó nuestro llamado a validar nuestro partido que es patrimonio nacional. Gracias a la inmensa mayoría de nuestros conciudadanos creyentes en la solución constitucional, democrática, pacífica y electoral a este intenso drama que agobia a Venezuela y que por la mano de la dictadura ha llevado a la violencia a las que el gobierno nos quiere desviar.

De lamentar que se quedara sin validar una cantidad de personas quizás igual o mayor a la que participó en la votación, y que por el momento no más capturadas, por su forma de distinción, por las pocas existentes en los sitios en que eran presumibles mayores afluencias, por interrupciones en su funcionamiento por diversas causas y por los intentos de la dictadura de impedirlo. De lamentar que AD tuvimos disponibles menos máquinas y horarios más restringidos. No obstante, ahí están los resultados. Y ese es uno de los aspectos que demuestran la voluntad antidemocrática del régimen cuando trataba de impedir que le desgraciado restringe, limita y complica todo.

Con este invento de la validación, el gobierno quiso por partida doble crear un obstáculo para difundir las elecciones de gobernadores ya vencidas y que no tiene sentido y consecuencia oficial a los partidos de la oposición porque, aunque para él un contrasentido, legalizarlos legalmente. Pero el tiro va a salirse por la culata. Gracias Venezuela.



Guaido = 0.45

Guaido = 0.7029

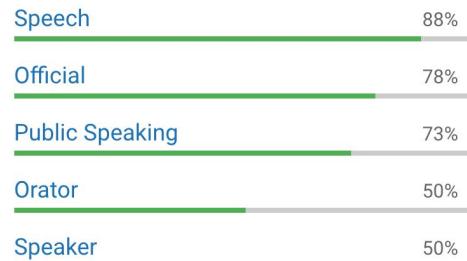
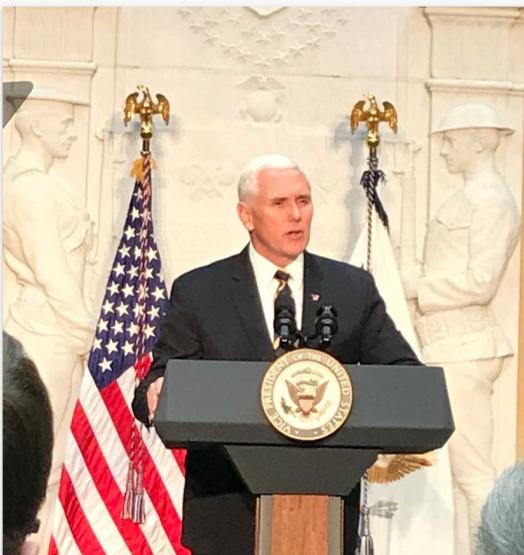
# Commercial Auto-Taggers: Potentials & Limitations

# Tools that allow automatic analysis of images

- Machine learning: convolutional neural networks (CNN)
- Commercial platforms like Google Cloud Vision API allows to
  - Upload a picture,
  - Have the platform analyze it, and
  - Obtain labels (and more information) for the image.

[Why Google](#)[Products](#)[Solutions](#)[Launcher](#)[Pricing](#)[Security](#)[Customers](#)[Document](#)

&gt;

[Faces](#)[Labels](#)[Web](#)[Document](#)[Properties](#)[Safe Search](#)[JSON](#)

An older example of the Google Cloud Vision API (GUI)

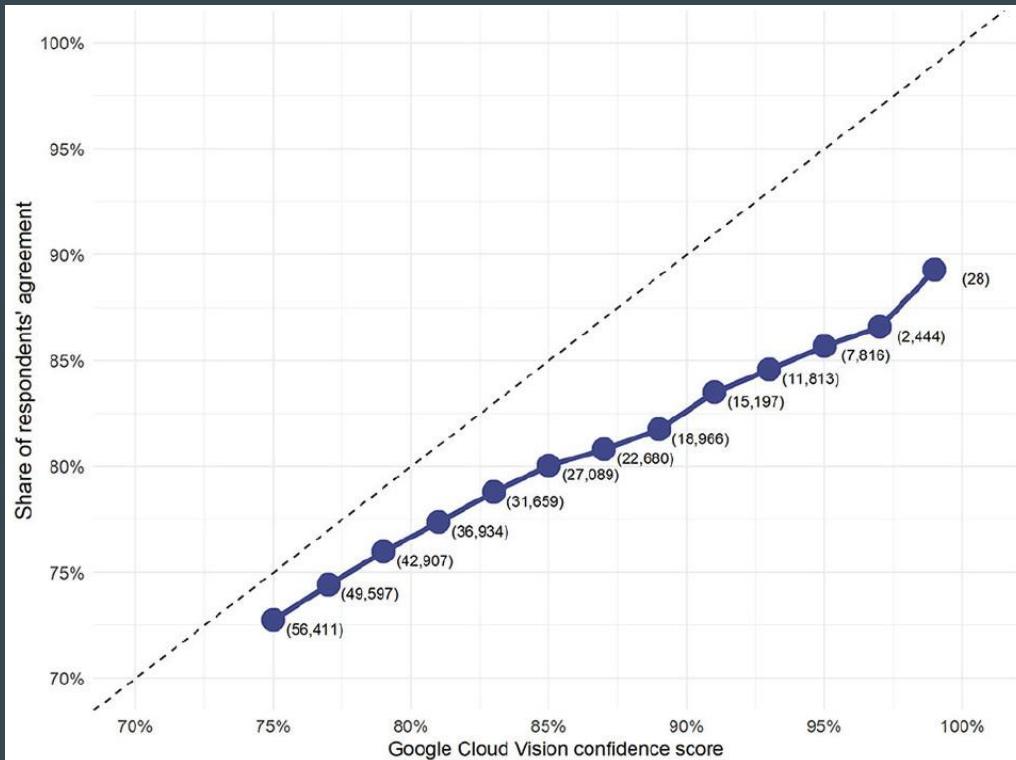
# When you do this programmatically via API, you get usable data

	img_labels	img_values
1	senior_citizen, official, event, interaction, fun	0.9065596461296082, 0.7297561764717102, 0.71...
2	infrastructure, tree, walking, road, recreation, street...	0.8797321319580078, 0.8417906165122986, 0.70...
3	room, standing, furniture, electronic_device, technol...	0.8771613240242004, 0.7967374920845032, 0.79...

# Google Cloud Vision in research

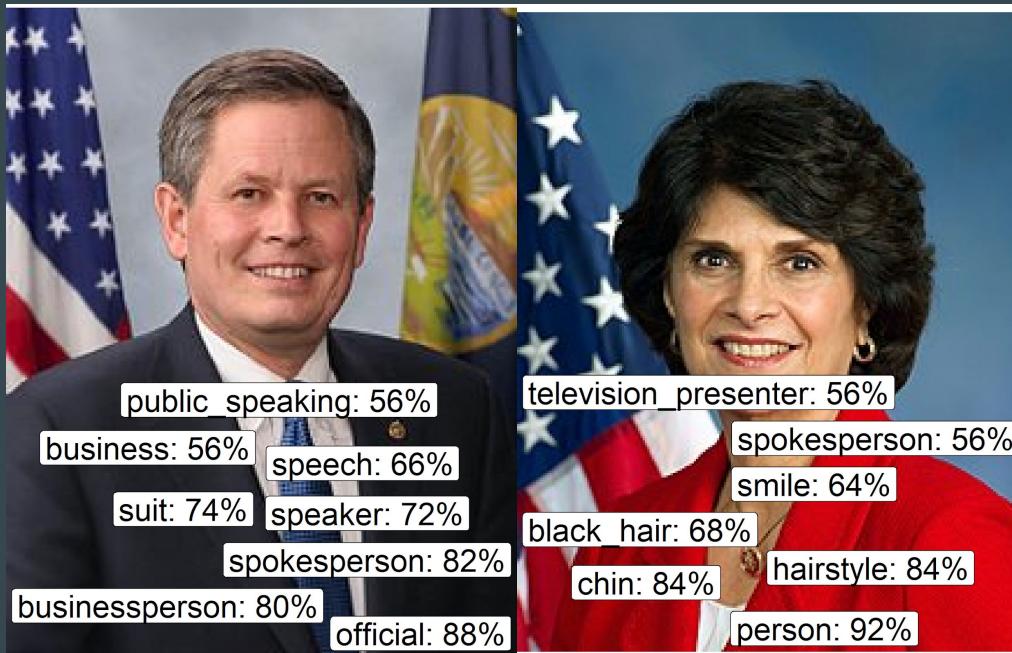
- Few studies have looked into using Google Cloud Vision in research and found its accuracy being relatively close to that of human coders
- Bosch, Revilla, Paura (2018):
  - Between 52.4% and 65.0% of the images were similarly codified by the Google Vision API and the human coder.
  - The API codified 1,818 images in less than 5 min, whereas the human coder spent nearly 35 hours to complete the same task.
- Could political images also be accurately coded by the API?

# The good news: can agree with human coders quite a lot



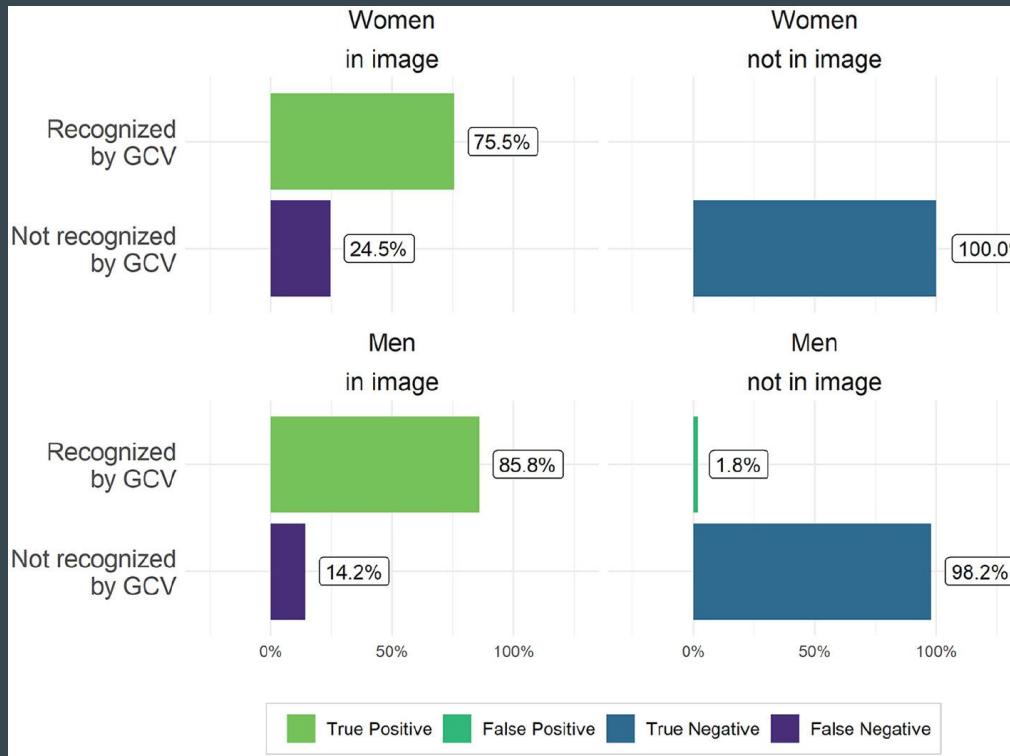
Analysis from Schwemmer, C., Knight, C., Bello-Pardo, E. D., Oklobdzija, S., Schoonvelde, M., & Lockhart, J. W. (2020). Diagnosing gender bias in image recognition systems. *Socius*, 6, 2378023120967171.

# The bad news: they can amplify biases that are hard-coded into the data to an extent that can be difficult to quantify



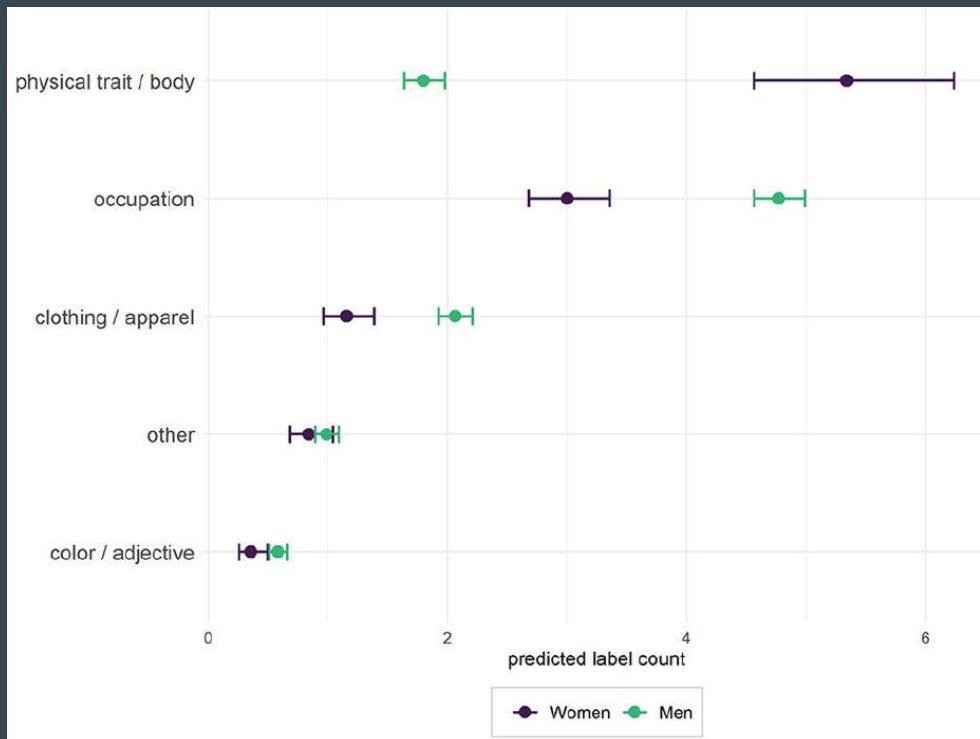
Example: headshots for Steve Daines (R, MT) and Lucylle Royball-Allard (D, CA-40), from Schwemmer, C., Knight, C., Bello-Pardo, E. D., Oklobdzija, S., Schoonvelde, M., & Lockhart, J. W. (2020). Diagnosing gender bias in image recognition systems. *Socius*, 6, 2378023120967171.

# The bad: GCV “sees” women in images at different rates...



Analysis from Schwemmer, C., Knight, C., Bello-Pardo, E. D., Oklobdzija, S., Schoonvelde, M., & Lockhart, J. W. (2020). Diagnosing gender bias in image recognition systems. *Socius*, 6, 2378023120967171.

# The bad: GCV is biased in the type of labels it gives images...



Analysis from Schwemmer, C., Knight, C., Bello-Pardo, E. D., Oklobdzija, S., Schoonvelde, M., & Lockhart, J. W. (2020). Diagnosing gender bias in image recognition systems. *Socius*, 6, 2378023120967171.

## TL;DR:

Use commercial auto-taggers at your own risk depending on your own goals and ideally understanding the kind of biases that could be baked into it

# A word about LLMs



pretend you're a commercial autotagger like google cloud vision and provide me labels that apply to this image i'm attaching here

## Analysis ▾

If I were processing this image as a commercial computer vision API, here are the labels and confidence scores I would generate based on the visual features, composition, and context of the photo.

**Image Analysis: Official Portrait**

Label	Confidence	Category
<b>Person</b>	99%	Entity
<b>Suit</b>	98%	Clothing
<b>Official</b>	94%	Professional
<b>White-collar worker</b>	92%	Occupation
<b>Formal wear</b>	91%	Style
<b>Government</b>	89%	Context
<b>United States Flag</b>	88%	Object
<b>Tie</b>	87%	Clothing
<b>Portrait</b>	85%	Photography
<b>Politician</b>	82%	Persona
<b>Spokesperson</b>	78%	Role

 Export to Sheets

## Additional Metadata

- **Dominant Colors:** Dark Blue (Navy Suit), White (Shirt), Red/White/Blue (Flag).
- **Composition:** Chest-up medium shot, centered subject, flags in the background (typical of "head of state" or "public servant" iconography).
- **Facial Expression:** Smiling, Neutral, Professional.

## Entity Detection (Knowledge Graph)

Beyond general labels, a commercial tagger would likely identify the specific individual:

- **Entity:** Steve Daines
- **Description:** United States Senator from Montana.

# CNNs vs LLMs

- LLM doesn't "see" a picture like a CNN does – i.e. doesn't use pixels.
- Hallucinations do happen with LLMs – it can confidently generate a label that is not in the actual picture.
- Outputs can be unpredictable – even plausible labels may be wrong or biased, and it's hard to measure errors like the previous paper did, since LLMs are always learning and adapting.

## TL;DR:

Use LLMs depending on your own goals  
and understanding the kind of biases that  
could still be baked into it (and keep in  
mind possible hallucinations!)

# Conclusions

# Recap

- Images-as-data is a very exciting new frontier in research with many interesting applications.
- You can use convolutional neural networks (especially transfer learning) to code images at scale.
- You can also use commercial autotaggers or even LLMs but please be careful with those, as research has shown these could prone to significant biases.
- Think of the ethics of big data analysis before doing any of this (see *Bit by Bit* by Matt Salganik for further discussion)

# More info...



# Thank you!

Questions? Feedback?  
Emil D. Bello-Pardo, PhD  
[ebellopardo18@gmail.com](mailto:ebellopardo18@gmail.com)