Facial feature detection technology (and its inherent bias)

EE550 Mini-project Final Presentation

Haeeun Kim

15/01/2021

Outline

1. Motivation
2. Viola-Jones Algorithm
3. Convolutional Neural Network
4. Twitter: a case study
5. Discussion.
6. Questions to explore

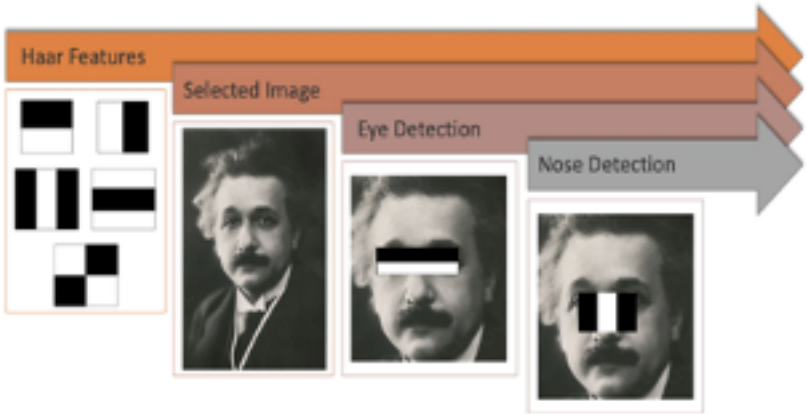
Motivation of Face Detection Technology

https://www.sachamber.org/blog/2019/10/majesty-egyptian-art-artifacts/
https://encykorea.aks.ac.kr/Contents/Item/E0019993
https://www.moma.org/collection/works/61240

https://www.koreatimes.co.kr/www/opinion/2019/11/715_255208.html
https://twitter.com/theellenshow/status/440322224407314432

Viola-Jones Algorithm

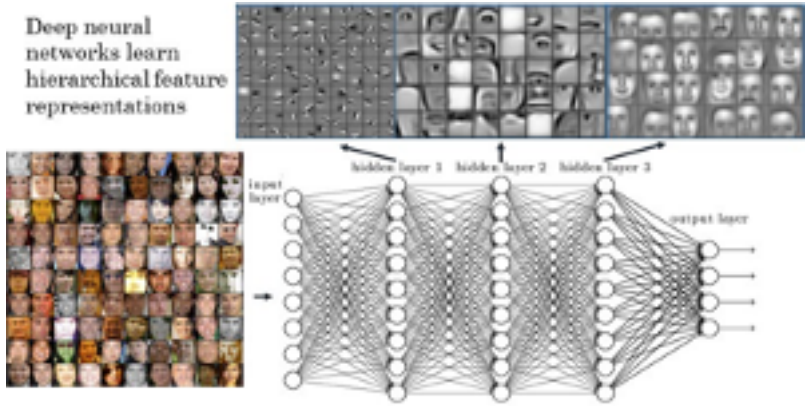
- ***“Faces have certain skeletal structures”***
  - Protrusions, valleys, and plateaus
  - Contrast between shadows and highlights
- Haar feature
  - Adjacent rectangular regions with sufficiently different average brightness
- Model-based method
  - Cascade of feature detection
  - “Reject the image if the feature is not found”
  - Face vs. Non-face classification



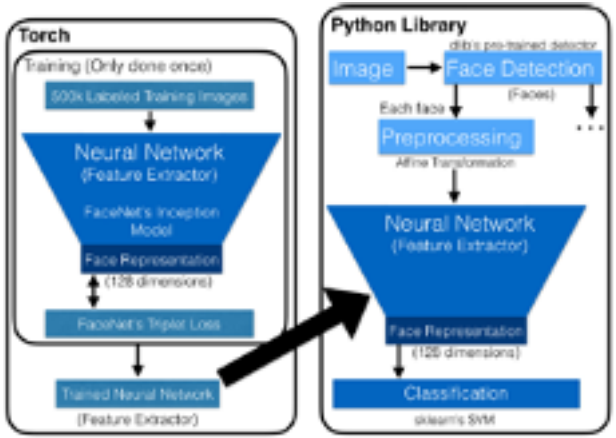
Leslie, D. (2020). Understanding bias in facial recognition technologies: an explainer. The Alan Turing Institute.

Convolutional Neural Network

- ***"Face images contain snippets of facial elements"***
  - Eyes, nose, entire face, etc
- Convolutional layers
  - Convolution operation with different face-detecting filters.
  - Multiple hidden layers with multiple filters.
- Training-based model
  - Face-detecting filters are trained from large training data
  - Face detection and recognition

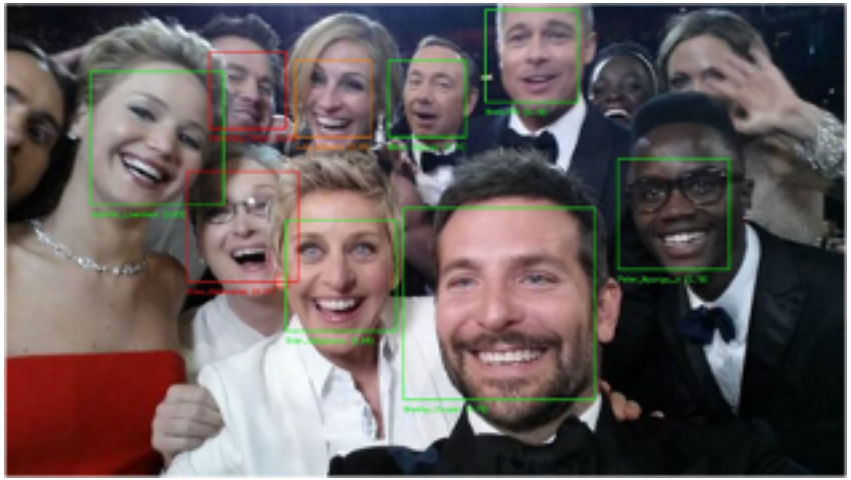


https://medium.com/@fenjiro/face-id-deep-learning-for-face-recognition-324b50d916d1



https://github.com/eashanadhikarla/Facial-Recognition-with-DNN

A Huge, Biased Success



https://laptrinhx.com/an-intro-to-deep-learning-for-face-recognition-1809043462

https://www.codedbias.com/

/

Twitter: A Tiny Experiment

- Twitter image cropping system
  - ‘Important’ part of images are cropped as previews
- Mitch McConnell vs. Barack Obama
  - The former president cropped out.

https://twitter.com/bascule/status/1307440596668182528

Twitter: Public Reaction

- Public Outrage
  - 79.9k retweets (including 16.6k quotes)
  - 197k likes
- Thousands of voluntary follow-up experiments

https://twitter.com/TheArtGun/status/1307508831363104768
https://twitter.com/TheArtGun/status/1307508831363104768
https://twitter.com/JackCouvela/status/1307602747718594563

Twitter: How it Settled (so far)

- Official response from Twitter
  - No bias detected while testing
  - Future improvements to be made
- No improvements made yet

Product

Transparency around image cropping and changes to come

By Parag Agrawal and Dantley Davis

Thursday, 1 October 2020

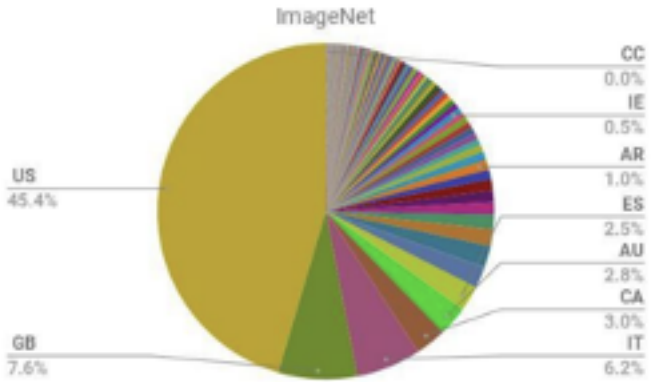
We're always striving to work in a way that's transparent and easy to understand, but we don't always get this right. Recent conversation around our photo cropping methods brought this to the forefront, and over the past week, we've been reviewing the way we test for bias in our systems and discussing ways we can improve how we display images on Twitter. So, while there's a lot still to do, today we want to share how we're developing a solution for each of these areas.

https://blog.twitter.com/official/en_us/topics/product/2020/transparency-image-cropping.html

Discussion(1): Bias is a Feature, not a Bug

- Viola-Jones: “Faces have certain skeletal structures and those regions can be distinguished with brightness”
  - Does it apply for all ethnic groups?
- CNN: “Faces have facial elements that match the filters, constructed from the training dataset”
  - Is the training set well-balanced?

https://www.shutterstock.com/image-vector/elegant-silhouettes-african-american-asian-caucasian-1023992569



Breakdown of the geographic origin of the images in ImageNet 23.[1]

Discussion(2): Default Privileging of Paleness in Photography

- *Why are faces of people of color usually brighter in photographs? (even nowadays?)*
- *Has humankind ever made enough effort to capture darkness of human skin in photography?*

https://www.boisestatepublicradio.org/post/martin-luther-king-jrs-legacy-idaho#stream/0

Slave family picking cotton in the fields near Savannah, circa 1860s, Bettmann Archives/Getty Images

Discussion(3): “It’s just an error. Why should we care?”

- Face is not ‘just an object’.
- Face detection failure is not ‘just an error’.
- Unequal system performance between different ethnic groups is not ‘just an error’ either.
- Subjective evaluation counts.

https://theater-masks.com/

Questions to Explore

- How should we define facial feature?
- How humans perceive faces?
- What is an appropriate metric?
- What is the mission of this technology?

References

[1] Freney B el at., 2020, Profiling and Convention 108+ : Report on developments after the adoption of Recommendation (2010)13 on profiling. 15 p.

[2] Leslie, D., 2020, Understanding bias in facial recognition technologies: an explainer. The Alan Turing Institute.

[3] F Schroff et al., 2015, FaceNet: A Unified Embedding for Face Recognition and Clustering, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), arXiv:1503.03832 [cs.CV]

[4] Honglak L et al., 2011. Unsupervised Learning of Hierarchical Representations with Convolutional Deep Belief Networks.Commun. ACM 54, 10 (October 2011), 95–103. DOI:https://doi.org/10.1145/2001269.2001295

[5] Socret L, (2020, May 22), Understanding Face Detection with the Viola-Jones Object Detection Framework [Medium post] Retreived from https://towardsdatascience.com/understanding-face-detection-with-the-viola-jones-object-detection-framework-c55cc2a9da14