

Isabel O'Connor
12/09/24

Building a CNN Model to Correctly Identify Celebrity Facial Images

Facial recognition technology has become increasingly prevalent in industries such as law enforcement, retail, and banking. However, this tool is hotly contested in the world of data ethics, posing challenges for individual privacy.

Facial recognition tech has immense potential to improve efficiency in retail spaces. Amazon launched their first “Amazon Go” store in 2018, utilizing facial recognition technology to create a checkout-less shopping experience (Wingfield, 2018). However, the company encountered several challenges when implementing the technology at a large scale. As of 2018, Amazon Go relied on human data technicians to verify thousands of images taken in stores (Wingfield, 2018). This suggests that current facial recognition models are insufficient at tracking the retail environment with complete accuracy.

Facial recognition is also utilized in the banking industry. As of 2021, 15 to 20 percent of financial institutions in the United States utilized selfie photo recognition as a tool to authenticate users (Heun, 2021). This measure may improve security in online banking. However, it has also sparked controversy in the realm of data privacy, as users question how their personal data is stored and managed.

On the law enforcement side, the NYPD utilizes facial detection to match mugshot photos with crime scene images. The department denies that any false arrests have been made as a result of this technology (NYPD, n.d.). However, with growing evidence of racial bias in training data, it is crucial to monitor and improve possible defects in facial recognition models, particularly when implementing these measures in legal procedures.

Deliverable:

In this case study, you will train a CNN model to correctly identify celebrity images. Throughout the course of this endeavor, you will also explore some of the model's weaknesses and biases. Your initial database will consist of google images from 25 different celebrities, with 100 images per individual. Using the ImageDataGenerator class, you will conduct image augmentation (flipping, cropping, and resizing photos) in order to generate a larger database. Subsequently, you will build and train a facial identification model using Keras. As a final measure, you will investigate the overall accuracy rate of identification, as well as conduct EDA on misidentified images in order to evaluate particular weaknesses in the model.

Github Link: <https://github.com/oak50/CS3-DS4002>

Works Cited

- Heun, D. (2021, October 4). *Facial recognition tech is catching on with Banks*. American Banker.
<https://www.americanbanker.com/news/facial-recognition-tech-is-catching-on-with-banks>
- NYPD questions and Answers facial Recognition*. Facial Recognition - NYPD. (n.d.).
<https://www.nyc.gov/site/nypd/about/about-nypd/equipment-tech/facial-recognition.page>
- Wingfield, N. (2018, January 21). *Inside amazon go, a store of the future*. The New York Times.
<https://www.nytimes.com/2018/01/21/technology/inside-amazon-go-a-store-of-the-future.html>