**What is the Effect of Facial Recognition on Racial Bias?**

**Noah Buchanan**

**Rockdale Magnet School for Science and Technology**

**Table of Contents**

Abstract: Pg. 3

Introduction: Pg. 4 – 5

Results: Pg. 5

Graphs: Pg 6 – 8

Conclusion: Pg. 8 - 9

Appendix: Pg. 9 – 10

Reference: Pg. 10 - 11

Acknowledgements: Pg. 11

What Is the Effect of *Facial Recognition* on *Racial* *Bias*?

Systems Software

**Abstract**

The research hypothesis is that facial recognition is inherently biased to different demographic groups, due to machine learning not having sufficient data from other demographics being tested in facial recognition's initial creation. This created a cascade of biases as other models were being built on previous ones. These biases can make criminal investigation inaccurate if facial assessment is incorrect. Using the 2 databases I was granted access to; I could run trials on over 10,000 images. For each image in each database, I used Google's facial recognition solution, MediaPipe, to perform facial recognition and return a numerical value of how confident the program was in its facial assessment in a numerical percentage value. I repeated this when splitting the databases into different demographic groups (white/black, male/female), and used python's file managing systems, to store the data. Using Matplotlib, NumPy, and pandas (python packages specializing in data visualization/analysis), I was able to find the median, mean, and mode of the different demographics, and create different graphs to visualize the data. Over a total of 10,168 images analyzed, different conclusions were made for all images as well for each specific demographic. I found that the accuracy was slightly skewed in favor of females, while male face detection typically performed worse. In conclusion, the research conducted supported the hypothesis that not only is facial recognition biased to different demographics groups, but also that technology may not be ready to be solely responsible for facial assessments, especially in criminal cases.

**Introduction**

Facial recognition is a technology used to identify human faces or match a given face to another face in a database (Najibi, 2020). This is commonly used in police investigations, to identify a suspect from a list of similar people. However, this is a danger, for the many innocent people who might be caught in harm’s way. An inaccurate assessment of a person could result in an innocent person’s incarceration. As of 2016, It is estimated that over 117 million people have photos actively used in a facial database by law enforcement (Alex Najibi, 2020). The purpose is to investigate a commonly used facial recognition solution to see how accurate it is on different demographics, and how this accuracy may be skewed based on race or gender. There’s already a belief that there’s less accuracy with African Americans (Coutin,2020). Using Opencv (OpenCV, 2020) I will be able to find data on how accurate computers are at finding faces. The research hypothesis is that the accuracy of the commonly used facial database will have a significant change in accuracy based on different demographics. The sub hypothesis being that Caucasian females will have the most significant increase in accuracy. The null hypothesis being that different demographics will have no significant effect on the accuracy.

**Method**

Over 10,000 trials were run on images from 2 databases.  Google’s facial recognition solution, MediaPipe, performed facial recognition on each image and returned a numerical value of how confident the program was in its facial assessment. The databases were separated into different demographic groups (white/black, male/female), and the facial recognition process was repeated. Using python’s file managing systems, data was stored for each image in each demographic group. Matplotlib, Pandas, and NumPy (python packages specializing in data visualization/analysis), were used to find the median, mean and mode for each demographic, and create graphs to visualize data. Conclusions were made for all 10,000 images, and separate conclusions were drawn for the separated images in each demographic group.

**Results**

The resulting data was extremely diverse for all 10,000 images (Figure 1.). There were low outliers which signify that the accuracy of the facial recognition is likely too low to be used in a real-world setting. On all 10,000 images, the highest value of % confidence was 96.98%. This value was the outlier for the upper data. The mode of all 10k image assessments (mode is used because as seen in Figure. 1, the mean isn’t a viable measurement due to the amount of inaccurate assessment outliers) was 86.93% confidence score. The mode of the Caucasian images (Figure 2.) was 75.2% confidence score. The mode of the African American images (Figure 2.) was 76.94% confidence score. This difference isn’t enough to conclusively support the hypothesis alone, as 1% variation could be blamed on experiment errors. More conclusive data can be seen when comparing male (Figure 3.) and female (Figure 3.) demographics. The mode of the male images’ confidence scores was significantly less than the female values. The male’s mode was 75.2% in confidence assessment. While the female mode was 83.13% in confidence assessment. This did support not only the hypothesis but the sub hypothesis as well.

**Graphs**

Figure 1.

Chart, histogram

Description automatically generated

Figure 2.

Chart, histogram

Description automatically generated

Figure 3.

Chart, histogram

Description automatically generated

**Conclusion**

The research hypothesis is that facial recognition is inherently biased to different demographic groups. The results support the hypothesis that facial recognition is biased towards different demographic groups. From the modes of the data, we see that the most common result of the data is, generously, about 80%. To be used in real-world in police investigations, it is believed that this number should be higher to hold any value as a piece of evidence. Especially in criminal cases where stakes are even higher. However, another interpretation of the results is that facial recognition may sometimes be a decent place to start when searching for a suspect to narrow down the field. On the other hand, we see that 5% of the time, the detection score was less than 30%, which is unacceptable. A way to make the results more consistent and acceptable would be to have: higher quality images (the ones used in the 10k adult face database were quite small), constant, stagnant lighting for all individuals, and images where all facial expressions are constant (similar to the pictures might be inside a facial database). In the future, a database satisfying all these criteria, as well as having many more images (the more images the more accurate the results), would improve the quality of results. Regardless of which interpretation is accepted, it is true that there is lots more progress able to b**e** made and manymore discussions about the ethicalness and privacy to be had.

**Appendix A (Caucasian Demographic % Confidence Scores)**

Due to the constraints of a research paper’s appropriate length, here’s a short sample of my raw data which comes from an 190+ page document with over 50,000 characters:

The data indicates the path to the 10k US adult database (Bainbridge, W.A; 2013), and the specific image inside the dataset. For each image path the detection score or percent confidence was added to the end for data analysis later on. This data is specifically from the Caucasian demographic. One’s own custom data analysis can be performed on the numerical values below, however, access to the 10k US adult database (Bainbridge, W.A; 2013) must be granted by the owner.

C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_Danielle Goble\_5\_oval.jpg,75.20707249641418,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_Phillip Owensby\_9\_oval.jpg,83.12661647796631,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_Douglas Ziegler\_3\_oval.jpg,78.97548079490662,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_Peter Grigsby\_13\_oval.jpg,86.87063455581665,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_Nathan Beauregard\_3\_oval.jpg,66.53748750686646,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Brian\_Furlong\_5\_oval.jpg,67.88977980613708,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_Paul Wills\_11\_oval.jpg,82.30311274528503,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_Jesse Lindgren\_7\_oval.jpg,91.55169129371643,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_David Ferris\_3\_oval.jpg,79.23969626426697,C:\Users\user\Documents\research\10k adult database\10kfaces\10k US Adult Faces Database\Face Images\Google\_1\_Bonnie Fullmer\_5\_oval.jpg,69.89702582359314

**References**

Alex, Najibi. (2020, October 26). BLOG, SCIENCY POLICY, SPECIAL EDITION: SCIENCE POLICY AND SOCIAL JUSTICE: Racial discrimination in face recognition technology. <https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/>

Brad Coutin, Nist study evaluates effects of race, age, sex on face recognition software. NIST. (2020, May 18). <https://www.nist.gov/news-events/news/2019/12/nist-study-evaluates-effects-race-age-sex-face-recognition-software>. This article from the National institute for standards and technology outlined different sources of error inside of an algorithm. It explains the difference of a false-positve and a false-negeative.

Bainbridge, W.A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face images. Journal of Experimental Psychology: General. Journal of Experimental Psychology: General, 142(4), 1323-1334.

OpenCV. (2010, May 11 – 2020, April 13) OpenCV/data/haarcascades. github.com/opencv/opencv/commits/master/data/haarcascades

Stephen Fontenot. (2020, December 4). Study outlines what creates Racial Bias in Facial Recognition Technology. <https://news.utdallas.edu/science-technology/racial-bias-facial-recognition-2020/> Examination into facial recognition technologies; progressions in present technologies and the downfalls of past technology.

**Acknowledgements**

Portions of the research in this paper use the SCface database of facial images. Credit is hereby given to the University of Zagreb, Faculty of Electrical Engineering and Computing for providing the database of facial images.

Bainbridge, W.A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face images. Journal of Experimental Psychology: General. Journal of Experimental Psychology: General, 142(4), 1323-1334.s