INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD ISSN: 2455-0620 Volume - 7, Issue - 8, Aug – 2021
Monthly, Peer-Reviewed, Refereed, Indexed Journal with IC Value: 86.87 Impact Factor: 6.719
Received Date: 28/07/2021 Acceptance Date: 13/08/2021 Publication Date: 15/08/2021

Bias in Industry Leading Facial Recognition Services: A Regional Analysis Across South Asian Regions

Armeet Singh Jatyani
Independent Researcher – San Jose, California, United States.
Email – armeetjatyani@gmail.com

Abstract: *The objective of this study is to assess the extent to which bias is present, if any, in facial recognition services offered by FacePlusPlus, Google Cloud, and Microsoft Azure. This study assesses the selected services across eight different South Asian regions: Kashmir (North), Ladakh (North), Punjab (North), Rajasthan (Northwest), Jharkhand (East), Telangana (South), Tamil Nadu (Deep South), and Gujarat (West). Our results reveal interesting and concerning patterns between characteristics of regions, and final accuracy scores. FacePlusPlus had unevenly distributed beauty scores, with a range of approximately 9.49, was more likely to correctly identify the gender of males, and severely struggled to accurately detect the gender and faces of groups with heavy facial hair, such as males in the Punjab (North) region. Microsoft Azure was more likely to accurately predict the gender and face of females, and struggled the most (out of all three services) with groups that had heavy facial hair, with a gender detection accuracy of just 63% for the Punjab (North) region. Finally, Google Cloud performed phenomenally, with facial detection accuracy percentages higher than 90% across all eight regions and genders. The results reveal disturbing biases present in the FacePlusPlus and Microsoft Azure facial recognition/detection services, that should be addressed to maintain ethical integrity.*

Keywords: *bias, ethnic bias, regional bias, gender bias, computer vision, faceplusplus, google cloud, microsoft azure, gender classification, detection accuracy.*

1. INTRODUCTION:

Facial Recognition: A Quick Overview

Modern facial recognition is a revolutionary technology that is able to detect human faces in images, videos, and other widely consumable media. Industry leaders in the facial recognition services field, build computer vision applications and services, powered by machine learning and artificial intelligence. Today, facial recognition systems have the ability to predict ethnicity, gender, age, among other factors. Some systems even label faces with highly subjective values, such as beauty scores.

Facial Recognition Services: Industry Practices

Companies rely on the use of machine learning algorithms to train artificial intelligence models. As such, they require large datasets of faces to train, optimize, and evaluate their artificial intelligence models. Thus, naturally, if these systems are trained with inadequate data, they will produce inaccurate and biased results. We consider "inadequate" data as data that favors a certain group with similar features or weights one group with higher priority/importance than another in training. Most simply, bias can arise when a certain group is overrepresented or underrepresented in a dataset (Cowgill et al., 2020). A model that has been trained on a group that is overrepresented in a dataset, will perform vastly better and make more accurate predictions on that group, than on an underrepresented group in the training data.

Facial recognition services are used to confirm identity, aid missing persons investigations, and identify and track criminals. Outside the focus of this study, facial recognition services are also used in the biocryptography, education, and entertainment sectors. Facial recognition services have been used to apprehend and track criminals (Choraś, 2007). Real-time criminal surveillance systems, running on consumer mobile devices have been proposed in the past (Elrefaei, 2017). The vast implications of this technology compel us to consider the risks that come with such a novel technology, that can easily misidentify faces and fails to prove, without reasonable doubt, a match in identity between a criminal and suspect in question. Questions about the ethical considerations and implications of a technology are important to ask (King et al, 2019). When Facebook, for instance, launched a new feature that used facial recognition to tag people in images, it generated much controversy for failing to adequately announce and inform users (Buckley and Hunter, 2011). Should these technologies be considered in a court of law? What are the guidelines for the use of AI facial recognition models on the general public (Hayashi and Wakabayashi, 2017)?

Facial Recognition Bias: A Quick Look

Take for example a hypothetical situation, in which a machine learning model is trained on a dataset with 10,000 images. If 8,000 of these images are of males, and the final 2,000 of these images are of females, we can expect our model to show bias by gender. As the model has seen more examples of male subjects, when our model sees an unknown image, it may be more likely to classify the subject in that image to be a male. Additionally, if our model has been trained to predict more features, such as age and emotion, its predictions for these parameters may be less accurate on testing data with female subjects, as it has seen less examples of this group.

2. SERVICES :

Selection Criteria

Services to be analyzed and scrutinized throughout this research were chosen for a multitude of reasons, which can be simplified to just two criteria. Firstly, we want to include companies that offer facial recognition services in domestic, as well as foreign markets. Secondly, companies must have a substantial reputation in their respective fields and offer a developer API. This criteria is in place so their services can be thoroughly assessed.

FacePlusPlus

FacePlusPlus is the first service selected for this research. FacePlusPlus is a China based company, valued in the billions (“Face Detection - Face++ Cognitive Services,” 2021). It is a dominant leader in the Asian/Chinese market, with one of the largest (Jacobs and Ralph, 2018). FacePlusPlus offers many computer vision services, some of which include human body recognition, text recognition, and image enhancing. Moreover, the company’s facial recognition services offer novel features such as predicted beauty score, among other conventional features such as age, gender, ethnicity, and emotion detection.

Microsoft Azure

Microsoft Azure is a popular cloud computing service offered by Microsoft, serving millions of customers. Microsoft Azure offers its “Azure Cognitive Services” which consist of high quality AI models as web APIs. Azure’s Face API contains conventional features like the ability to label face pose, facial hair, and gender. Microsoft invests over $1B yearly on development and cybersecurity research of its services and employs thousands of security experts to ensure the integrity of its platform (“Facial Recognition | Microsoft Azure,” 2021). It is an industry leader in the cloud computing field.

Google Cloud

Google Cloud is an industry leading cloud computing service offered by Google, serving millions of customers. Google Cloud offers its “Cloud Vision API” which provides a facial detection service with traditional features such as face pose, bounding boxes, and critical points on the face. Additionally, the service offers the ability to predict emotion (“Detect Faces | Cloud Vision API | Google Cloud,” 2021). More interestingly, Google Cloud is the only service that doesn’t predict gender. Google has decided to leave gender labels out of their service to reduce the chances of “unfair biases,” recognizing that “male” and “female” labels don’t apply to everyone (Lyons, 2020). For this, among other reasons, Google Cloud’s Cloud Vision API is seen as one of the most progressive services in the market.

3. PROCEDURE :

Data Collection

To evaluate the three services in question, we compiled a dataset of faces. The dataset was organized and numbered by regions: 1_north _kashmir, 2_north_ladakh, 3_north_punjab, 4_northwest_rajasthan, 5_east_jharkhand, 6_south_telangana, 7_deepsouth_tamil_nadu, 8_west_gujrat. These regions were selected to maximize variance in the dataset. Faces in each different regional group have unique skin tones and facial features (Zhuang et al, 2010). For each region, 100 male and 100 female faces were selected. In total 1600 image files with 1600 faces were compiled into a dataset used in the study.

Script Overview: Nameformat

To keep our dataset organized and ensure that every face/image is trackable, we numbered every image in the dataset. This Python script scrambled the order of images in the dataset, and labeled these images using the following template: “{group_name}/in/{gender}/{gender}_in_{number}.png.”
For example the path “/3_north_punjab/in/male/male_in_8.png” describes the 8th face in the 3_north_punjab group of male faces, pictured below.

Images were assigned an unchanging ID. Since the analyze scripts, described later, pass dataset images to the cloud services in order of ID, the images are sent to all services in the same order every time. This was done to decrease the probability of confounding factors affecting results.

Sample Image in Dataset (North Punjab Region) "/3_north_punjab/in/male/male_in_8.png"

Script Overview: Analyze

After the dataset was compiled and organized, I wrote "analyze" scripts that sent API requests to the cloud services from FacePlusPlus, Azure, and Google Cloud, saving the JSON responses. In simpler, non technical terms, these scripts submitted images in our dataset to the services offered by these companies, and saved the responses (which contain predictions and facial recognition data) into JSON files, described below.

JSON, JavaScript Object Notation, is a file format standard that uses "key-value" pairs and arrays to display information. JSON files are the industry standard for API requests and responses (Nurseitov and Nurzhan, 2009). Below is a sample JSON response from FacePlusPlus, when passed in the image file "/3_north_punjab/in/male/male_in_8.png" pictured to the left, truncated for convenience.


```json
{
    "request_id": "***",
    "time_used": 143,
    "faces": [
        {
            "attributes": {
                "gender": {
                    "value": "Male"
                },
                "beauty": {
                    "male_score": 50.126,
                    "female_score": 55.933
                }
            }
        }
    ],
    "image_id": "***",
    "face_num": 1
}
```


After all "analyze" scripts have been run, every image in the 1600 image dataset has been passed to all 3 cloud services. The 4800 JSON responses from these services have now been saved.

The below JSON response, received from FacePlusPlus when given the subject in "/3_north_punjab/in/male/male_in_8.png," predicted that the subject was male, with a male beauty score of 50.126 and female beauty score of 55.933.

INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD ISSN: 2455-0620 Volume - 7, Issue - 8, Aug – 2021
Monthly, Peer-Reviewed, Refereed, Indexed Journal with IC Value: 86.87 Impact Factor: 6.719
Received Date: 28/07/2021 Acceptance Date: 13/08/2021 Publication Date: 15/08/2021

Script Overview: Gen CSV

JSON responses are powerful for conveying information in an organized fashion that is easily accessible through scripts. However, to import data into a spreadsheet software, to generate graphs, and make inferences about the collected response data, CSV files are far more suitable. CSV, comma separated values, is a file format that allows for storage of data in a fashion similar to a table, organized by columns (Shafranovich, 2005).

The gen-csv scripts parse through all JSON files, generating CSV files for each of the 8 groups, male and female. Below is a snapshot of one of these generated CSV files, for the 3_north_punjab male group. Each row represents one image. The first 8 lines of the FacePlusPlus response are shown below.

```
1 ,id,predicted_ethnicity,predicted_emotio
  n,predicted_gender,male_beauty_score,fem
  ale_beauty_score
2 0,0,N/A,surprise,Male,63.854,68.772
3 0,1,N/A,,,,
4 0,2,N/A,happiness,Male,53.472,54.405
5 0,3,N/A,neutral,Female,72.289,75.979
6 0,4,N/A,neutral,Male,45.51,47.265
7 0,5,N/A,happiness,Male,54.448,46.718
8 0,6,N/A,surprise,Male,42.722,40.952
```

Take for example, the entry on line 2, corresponding to image with id 0, in the 3_north_punjab male group. FacePlusPlus predicted the emotion of the subject in this image to be “surprise,” gender as “Male,” and male/female beauty scores as 63.854 and 68.772, respectively.

We have created 48 CSV files, 16 for each of the 3 cloud services being evaluated: FacePlusPlus, Microsoft Azure, and Google Cloud. Each of the 8 groups have 2 CSV files, corresponding to the male and female subgroups in each group. Each CSV file contains 100 entries/rows, corresponding to the 100 images in each subgroup.

Evaluation Criteria: FacePlusPlus

The features chosen to be analyzed and evaluated for FacePlusPlus were predicted-gender, male beauty score, female beauty score, and detection accuracy. Predicted gender and beauty scores are provided by the service. Detection accuracy is defined as the proportion of images in the dataset, in which the API/service was able to detect the presence of a face.

Thus the CSV files generated for FacePlusPlus contain the following columns: id, predicted_gender, male_beauty_score, and female_beauty_score.

Evaluation Criteria: Microsoft Azure

The features chosen to be analyzed and evaluated for Microsoft Azure were predicted-gender and detection accuracy. Detection accuracy is defined as the proportion of images in the dataset, in which the API/service was able to detect the presence of a face.

Thus the CSV files generated for Microsoft Azure contain the following columns: id, predicted_gender.

Evaluation Criteria: Google Cloud

The only feature we can evaluate for Google Cloud’s Face API service is detection accuracy. Detection accuracy is defined as the proportion of images in the dataset, in which the API/service was able to detect the presence of a face. This is because, as discussed in the introduction, Google’s service doesn’t predict gender for a given image.

Thus the CSV files generated for Google Cloud’s Face API include the following columns: id, detected. The detected column contains boolean values True and False corresponding to whether Google’s service successfully detected a face in the given image in our dataset.

Importing to Spreadsheets

CSV files are compact and robust, but they may be unreadable and inconvenient in comparison to the table-styled data in spreadsheets. Our objective is to import the generated CSV files into our spreadsheet in an organized manner that makes sense and allows for analysis of the results through quantitative and visual analysis (graphs).

We created 3 separate spreadsheets, 1 for each of the 3 services being assessed. Each spreadsheet had 9 sub-sheets, 8 of which contain table data for each of the 8 region groups, and the final one being a summary sheet with graphs and tables to compare the results of the services on the 8 unique regional groups with each other.

Compiling Graphical Representations

For FacePlusPlus, average beauty scores, gender prediction accuracy, and detection accuracy were graphed using bar charts for each of the 8 regional groups.

For instance, by graphing beauty scores for each region in a bar chart, we can compare the average beauty scores. Average beauty score should be equal in each region. However, if we observe certain regions with a significantly higher beauty score, we have reason to believe that the service is biased, favoring certain traits that are specific to a few regions.

4. REGION OVERVIEW :

To understand results and identify patterns from collected data, we must understand the facial features, skin tone, complexion, and any other facial features that are characteristic to each region.

1 - North Kashmir

People in the Kashmir region have a lighter skin tone. They have minimal facial hair.


2 - North Ladakh

People in the Ladakh region have a slightly darker skin tone. They also wear traditional headdresses and have minimal facial hair.


3 - North Punjab

People in the Punjab region often wear colorful turbans and have heavy facial hair. Their skin tone is slightly darker than those in Kashmir. In certain parts, women also often wear turbans.


4 - Northwest Rajasthan

People in the Rajasthan region have darker skin tones. They wear colorful turbans and have medium facial hair.


5 - East Jharkhand

People in the Jharkhand region have dark skin tones. They have minimal facial hair.


6 - South Telangana

People in the Telangana region have darker skin tones and medium facial hair. Many women wear nose rings.


7 - Deepsouth Tamil Nadu

People in the Tamil Nadu region have very dark skin tones. They have minimal facial hair. Additionally, many women wear vibrant gold nose-rings.


8 - West Gujarat

People in the West Gujarat region tend to have lighter skin tones. Males tend to have medium facial hair.

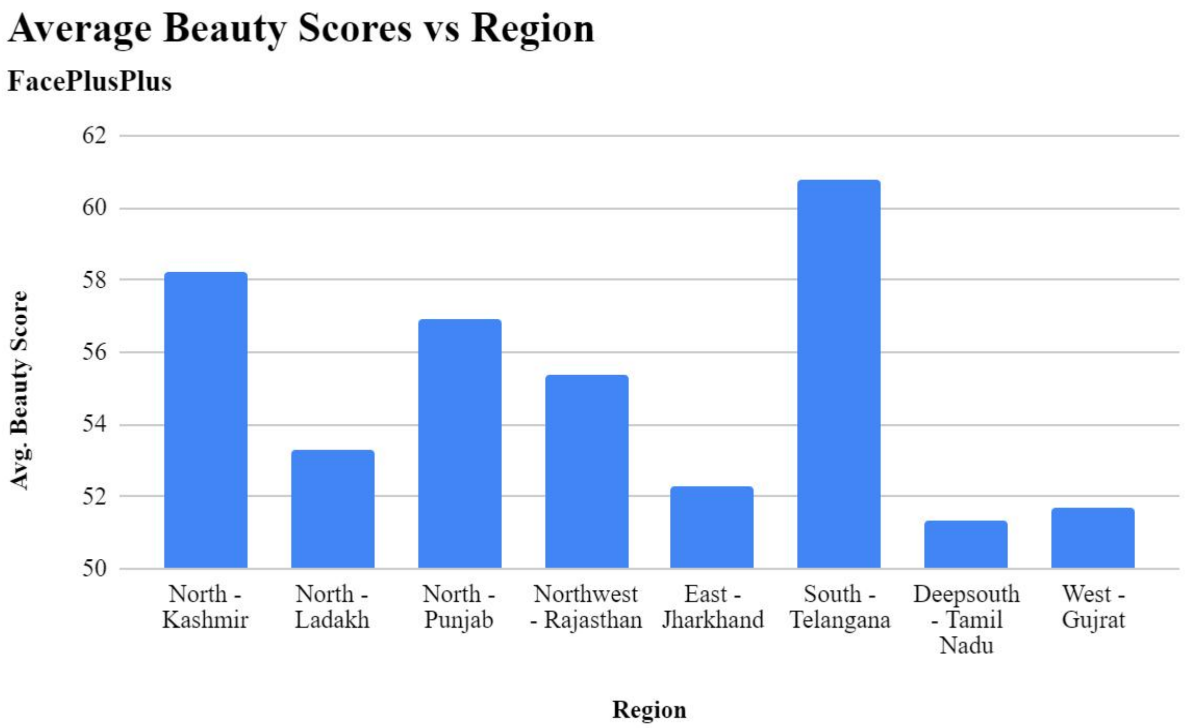

INTERNATIONAL JOURNAL FOR INNOVATIVE RESEARCH IN MULTIDISCIPLINARY FIELD ISSN: 2455-0620 Volume - 7, Issue - 8, Aug – 2021
Monthly, Peer-Reviewed, Refereed, Indexed Journal with IC Value: 86.87 Impact Factor: 6.719
Received Date: 28/07/2021 Acceptance Date: 13/08/2021 Publication Date: 15/08/2021

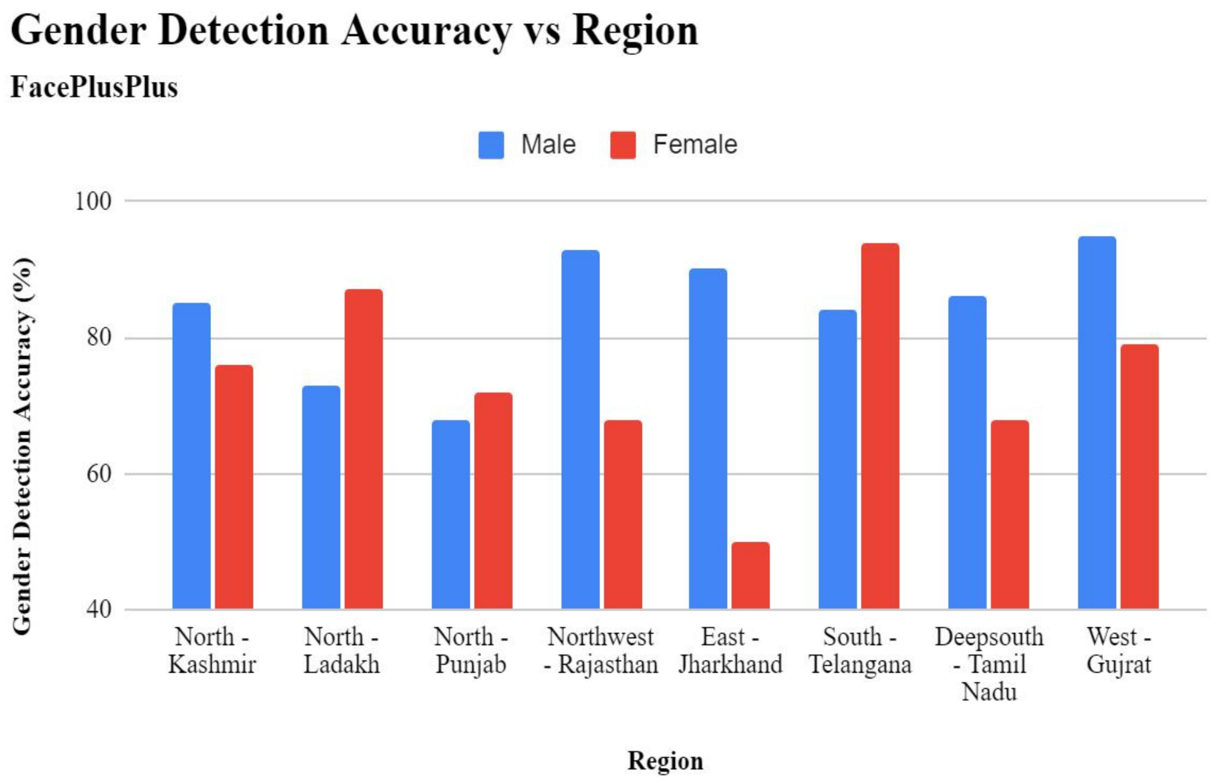
5. RESULTS :

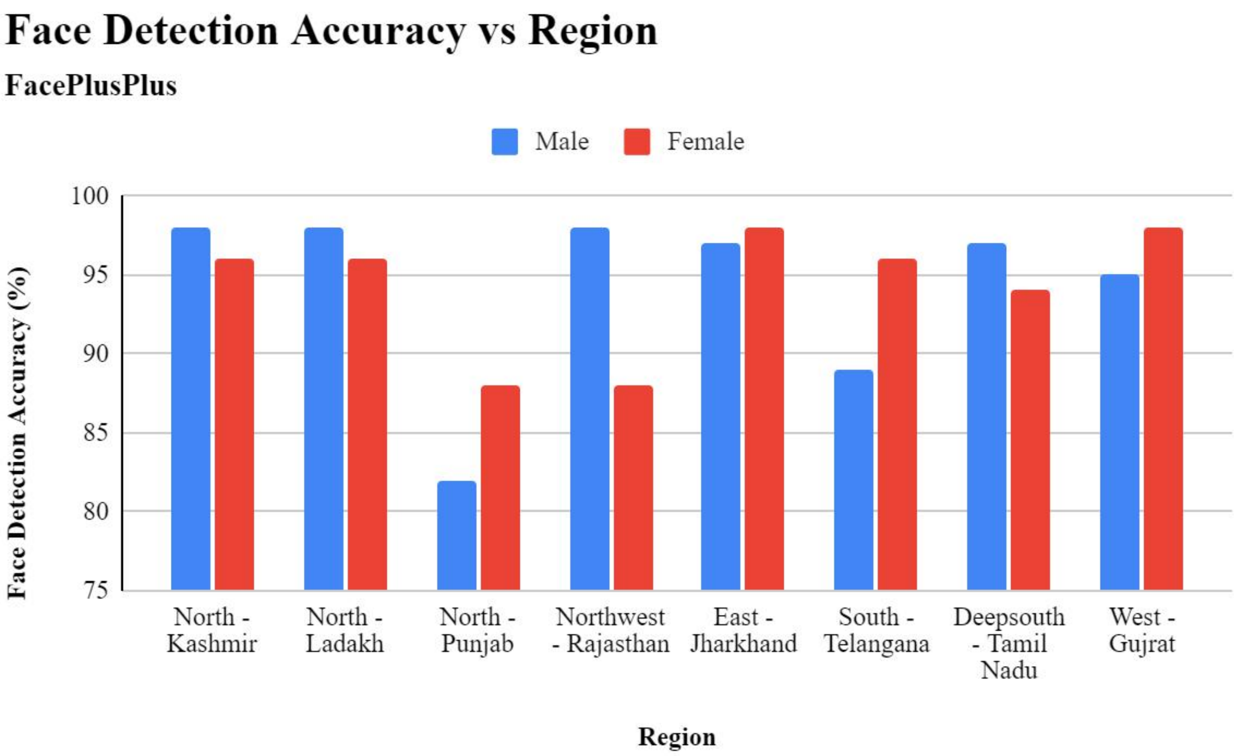
Below are the compiled results of our analysis. Some vertical axes have been scaled up (by starting at nonzero numbers) to magnify differences between regions.

FacePlusPlus

The three criteria which were studied include beauty scores, gender accuracy, and detection accuracy.

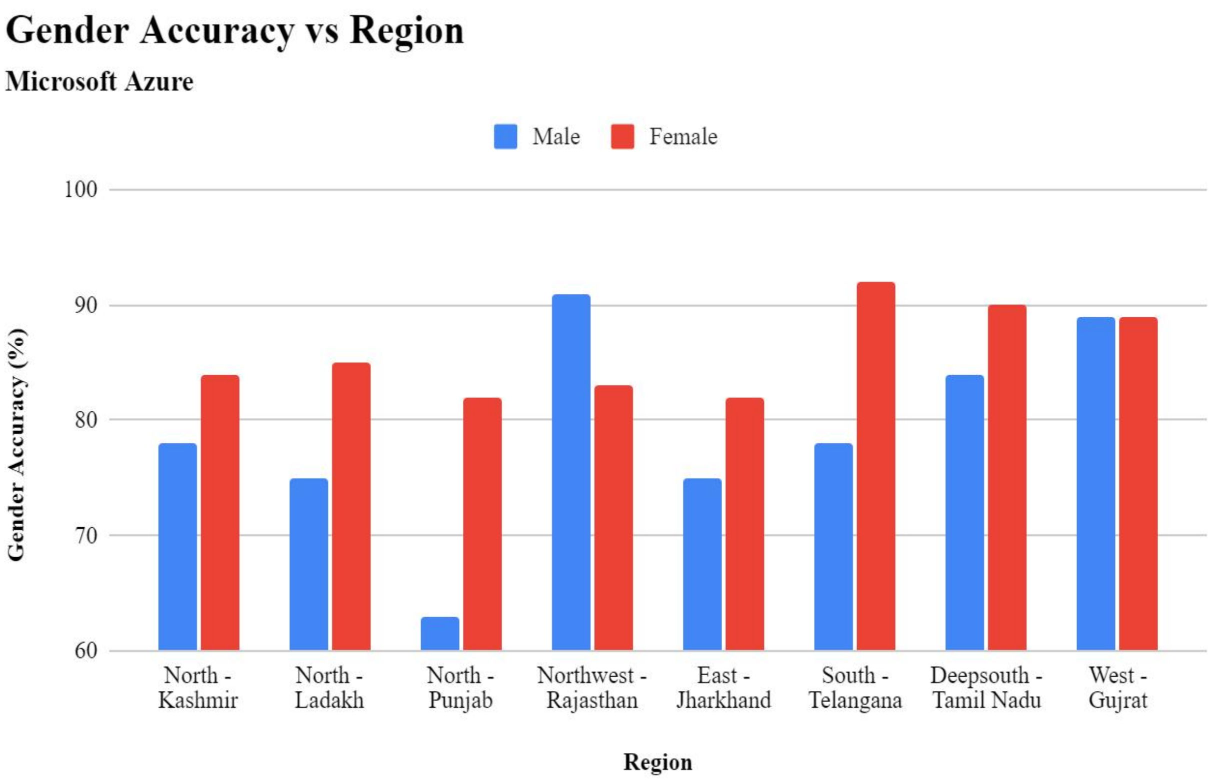


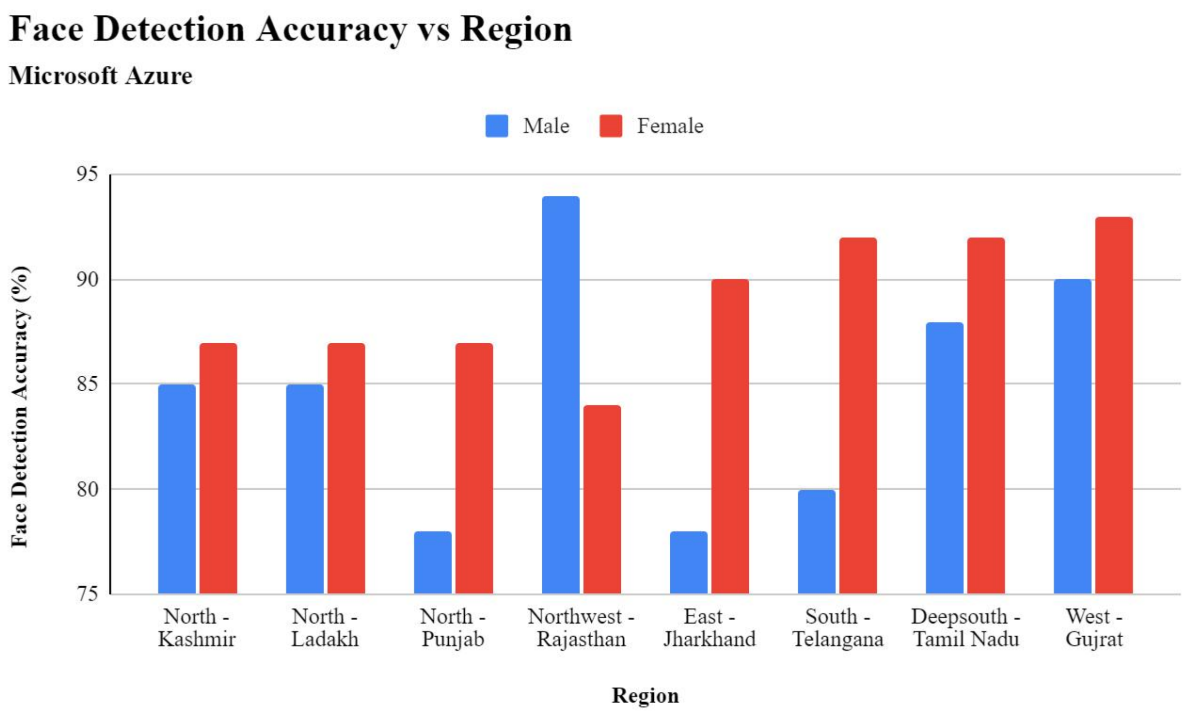




Microsoft Azure

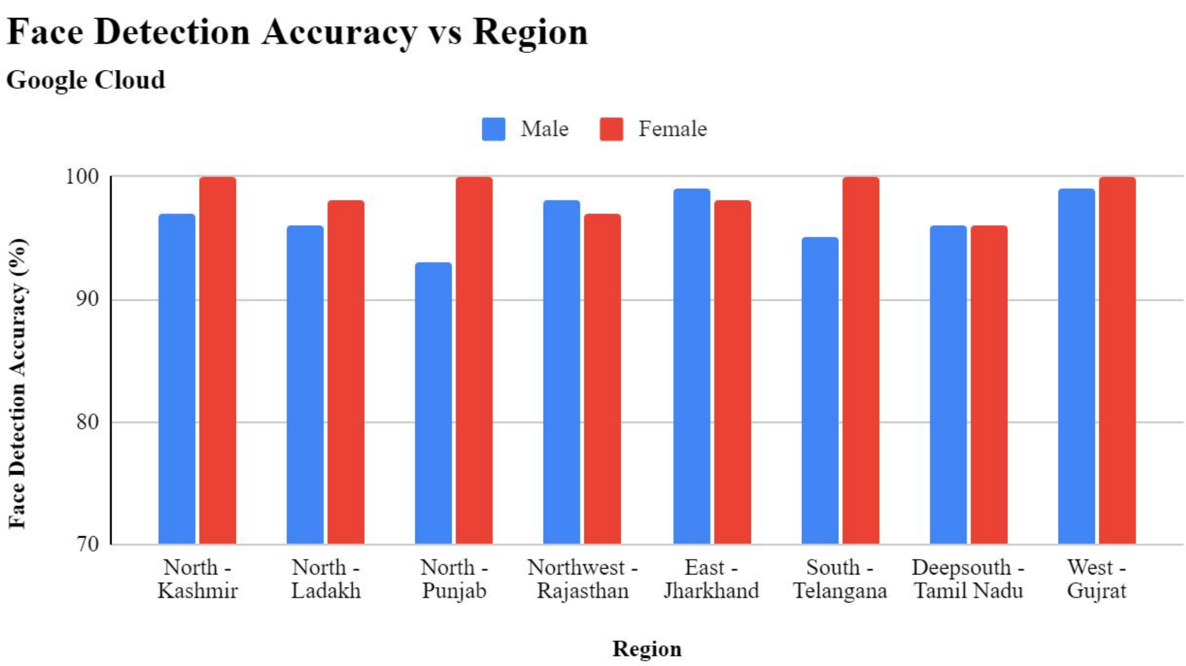
The two criteria studied include gender accuracy and detection accuracy.





Google Cloud

The criteria studied is detection accuracy.



6. ANALYSIS:

Results will be discussed and analyzed for each service. Any discussed results do not imply definite claims. Instead, they imply observable patterns and associations made from studying the data.

FacePlusPlus

The North Ladakh, East Jharkhand, Deepsouth Tamil Nadu, and West Gujrat regions had the lowest beauty scores across all eight regions. These regions shared a few characteristics. First, most of these regions, excluding the

West Gujarat region have dark skin tones. This would lead us to assume an association between darker skin tones and lower beauty scores. However, the South Telangana group, which had the highest average beauty score of any region, also had dark skin tones and the West Gujarat region had one of the lowest average beauty scores of any region, and light skin tones. We conclude that it is unclear whether skin tone has any effect on beauty scores determined by the FacePlusPlus service. However, the beauty scores are distributed unevenly, with a range of approximately 9.49. This brings us to a discussion of the ethical considerations of offering beauty scores as a feature. Beauty is highly subjective and varies by region (Brown, 2021). The existence of a global beauty score requires acknowledgement that certain features, that may be characteristic of a particular group, are objective traits of "beautiful" people (Bodini, 2019). Multiple beauty scores, for multiple regions or ethnic backgrounds, may be better fit to label beauty, as they take into account regional preferences that define "beauty." For example, FacePlusPlus could give ten beauty scores, each using a different model which has been trained on faces from ten different regions, similar to how FacePlusPlus currently returns a male and female beauty score.

Looking at the gender detection accuracy by region graph shows a disturbing pattern. On average, FacePlusPlus's service is much more likely to label a male's gender correctly, than for a female. Take for example the East Jharkhand region, where male gender detection accuracy is ~40 percentage points higher than female gender detection accuracy. In other regions however, such as the North Kashmir region, the male gender detection accuracy is ~10 percentage points higher than the female gender detection accuracy, a much smaller gap.

Finally, analyzing facial detection accuracy (whether the service was able to find a face in the provided images), we see one group stand out for their low accuracy percentages: males of the North Punjab region. The North Punjab male group is characterized by thick facial hair and large turbans. This group has a detection accuracy of 82%, the lowest of any of the eight regions studied. To emphasize this, North Kashmir males have a detection accuracy of 98% and North Kashmir females have a detection accuracy of 96%. Based on the lower detection accuracy observed, we conjecture that the FacePlusPlus service is biased against certain regions who grow out their facial hair. As for discrepancies across gender, there seems to be no observable pattern. For example, the Northwest Rajasthan group has significantly higher detection accuracy percentages for males, while in the South Telangana region, females have significantly higher detection accuracy.

Microsoft Azure

Apart from the Northwest Rajasthan group, gender detection accuracy was consistently lower for males. Additionally, one group stood out with a significantly low gender accuracy percentage: males in the North Punjab region. Males in the North Punjab region had their gender correctly predicted just 63% of times, which is 12 percentage points behind the group with the next lowest gender detection accuracy, males in the East Jharkhand region whose gender was predicted correctly 75% of times, and a staggering 28 percentage points less than males in the Northwest Rajasthan region who had their gender correctly predicted 91% of times, the highest across any region and gender. Thus, extensive facial hair, observed in groups such as the North Punjab region, may lead to lower gender detection accuracy, when using Microsoft Azure's provided service (Givens, Beveridge, Draper, Grother, and Phillips, 2004).

Again, apart from the Northwest Rajasthan group, facial detection accuracy was consistently lower for males. This time, the difference was far more noticeable. In the North Punjab region, accuracy was ~10 percentage points less for males. In the East Jharkhand and South Telangana regions, accuracy was ~12 and ~13 percentage points lower for males, respectively. The exception of the Northwest Rajasthan, in having both higher gender and face detection accuracy percentages for males, whereas all other regions had lower gender and face detection accuracy for males, suggests that Microsoft Azure's service favors a certain attribute. The existence of this exception, though unexplained, suggests the need for improvement from Microsoft's side, to improve gender and face detection accuracy for males in the South Asian regions studied.

Google Cloud

Google Cloud performed phenomenally, and face detection accuracy was higher than 90% across every one of the eight regions. From the extent of our analysis, Google's service doesn't seem to show any bias across genders or regions. Scores are equally high for both genders in all eight studied regions.

7. CONCLUSION :

Each of the three studied services offers exciting features, as leading services in the facial recognition and computer vision markets. However, each service has its drawbacks. Through our analysis, we find patterns of bias in some of the features offered by these companies.

FacePlusPlus offers an ambitious feature: beauty scores. However, from our analysis and as expected, the feature is highly opinionated. Beauty scores vary greatly by region. Although the cause of this variance can't be narrowed down to a single factor, its existence proves the need for a new system that accounts for the subjective nature of beauty. Currently, FacePlusPlus gives two scores, male and female beauty scores, to label the beauty of a face in the image, if that face belonged to a male and female. Thus, a system that gives various ethnicity, nationality, or region based beauty scores, would be far more subjective in its attempt to quantify and label "beauty." Through our analysis, we observed that males in all eight groups, on average, were more likely to be gender labeled correctly. Additionally, groups with heavy facial hair, such as males in the North Punjab region had the lowest facial detection accuracy. To improve its service, FacePlusPlus needs to address these biases.

Microsoft Azure showed bias by gender in both features assessed: gender detection accuracy and facial detection accuracy. Gender detection accuracy and facial detection accuracy were consistently lower for males. Additionally, groups with thick facial hair had outrageously low gender and facial detection accuracy percentages. Males in the North Punjab region (thick facial hair), for instance, had their gender correctly predicted just 63% of times, a staggering 28 percentage points less than males in the Northwest Rajasthan region (medium facial hair) who had their gender correctly predicted 91% of times, the highest across any region and gender. To improve its service Microsoft Azure needs to address these biases.

Finally, Google Cloud performed phenomenally, with facial detection accuracy scores higher than 90% across all eight regions and genders. With little additional data, more inferences and patterns can't be made. We found no significant biases present in Google Cloud's services. That being said, Google's services didn't offer gender detection like Azure or beauty scores like FacePlusPlus.

8. ACKNOWLEDGMENTS :

I would like to thank Mrs. Ihwa Miao, my AP Computer Science Principles teacher at Evergreen Valley High School, who first showed me the TED talk that inspired this research paper. I would like to thank Joy Buolamwini for her paper and accompanying TED talk: "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification," which inspired and led to my own research. I'd also like to thank Professor Ethan Zuckerman, former director of the MIT Center for Civic Media and current professor at the University of Massachusetts Amherst for suggestions and advice that was integral to the paper.

REFERENCES:

1. Cowgill, Bo, et al. "Biased Programmers? Or Biased Data? A Field Experiment in Operationalizing AI Ethics." *SSRN Electronic Journal*, Elsevier BV, 2020. *Crossref*, doi:10.2139/ssrn.3615404.
2. Choraś, Michał. "Intelligent Computing for Automated Biometrics, Criminal and Forensic Applications." *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2007, pp. 1170–81, http://dx.doi.org/10.1007/978-3-540-74171-8_119.
3. L. A. Elrefaei, A. Alharthi, H. Alamoudi, S. Almutairi and F. Al-rammah, "Real-time face detection and tracking on mobile phones for criminal detection," *2017 2nd International Conference on Anti-Cyber Crimes (ICACC)*, 2017, pp. 75-80, doi:10.1109/Anti-Cybercrime.2017.7905267.
4. King, Thomas C., et al. "Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions." *Science and Engineering Ethics*, no. 1, Springer Science and Business Media LLC, Feb. 2019, pp. 89–120. *Crossref*, doi:10.1007/s11948-018-00081-0
5. Buckley, Ben, and Matt Hunter. "Say Cheese! Privacy and Facial Recognition." *Computer Law & Security Review*, no. 6, Elsevier BV, Dec. 2011, pp. 637–40. *Crossref*, doi:10.1016/j.clsr.2011.09.011.
6. Yugo Hayashi and Kosuke Wakabayashi. 2017. Can AI become Reliable Source to Support Human Decision Making in a Court Scene? In *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (*CSCW '17 Companion*). Association for Computing Machinery, New York, NY, USA, 195–198. doi:https://doi.org/10.1145/3022198.3026338
7. "Face Detection - Face++ Cognitive Services." *Face++ - Face++ Cognitive Services*, https://www.faceplusplus.com/face-detection/. Accessed 5 Aug. 2021.
8. Jacobs, Harrison, and Pat Ralph. "China Facial Recognition Tech Company: Megvii, FacePlusPlus." *Business Insider*, Business Insider, 8 July 2018, https://www.businessinsider.com/china-facial-recognition-tech-company-megvii-faceplusplus-2018-5.
9. "Facial Recognition | Microsoft Azure." *Cloud Computing Services | Microsoft Azure*, https://azure.microsoft.com/en-us/services/cognitive-services/face/. Accessed 5 Aug. 2021.
10. "Detect Faces | Cloud Vision API | Google Cloud." *Google Cloud*, https://cloud.google.com/vision/docs/detecting-faces. Accessed 5 Aug. 2021.

11. Lyons, Kim. “Google AI Tool Will No Longer Use Gendered Labels like ‘Woman’ or ‘Man’ in Photos of People - The Verge.” *The Verge*, The Verge, 20 Feb. 2020, https://www.theverge.com/2020/2/20/21145356/google-ai-images-gender-bias-labels-people.
12. Zhuang, Ziqing, et al. “Facial Anthropometric Differences among Gender, Ethnicity, and Age Groups.” *The Annals of Occupational Hygiene*, Oxford University Press (OUP), Mar. 2010. *Crossref*, doi:10.1093/annhyg/meq007.
13. Nurseitov, Nurzhan, et al. "Comparison of JSON and XML data interchange formats: a case study." *Caine* 9 (2009): 157-162.
14. Shafranovich, Yakov. "Common format and MIME type for comma-separated values (CSV) files." (2005).
15. Brown, Jessica. “What Being Beautiful Means in 25 Countries around the World | Indy100 | Indy100.” *Indy100 | News | Politics | YouTube and TikTok News*, indy100, 21 July 2021, https://www.indy100.com/news/beauty-around-the-world-countries-b1887790.
16. Bodini, Matteo. “Will the Machine Like Your Image?Automatic Assessment of Beauty in Images with Machine Learning Techniques †.” *Inventions*, no. 3, MDPI AG, June 2019, p. 34. *Crossref*, doi:10.3390/inventions4030034.
17. G. Givens, J. R. Beveridge, B. A. Draper, P. Grother and P. J. Phillips, "How features of the human face affect recognition: a statistical comparison of three face recognition algorithms," Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., 2004, pp. II-II, doi: 10.1109/CVPR.2004.1315189.