**Executive Summary - Identify Verification**

Name: Hewei Shen

Email: [shenh23@wfu.edu](mailto:shenh23@wfu.edu)

# Business problem

This project aims to find the optimal Facial Biometrics technology for the company to ensure that employees return to work. A recommendation will be given at the end of the report considering the performance, complexity, cost, and maintenance.

# methodology

We compared *facial\_recognition* (open-source) against *Amazon ReKognition API service* in two use cases.

In use case 1, we want to know how well the two packages recognize the same person from different photos. Specifically, the algorithm compares source images (headshots) to target images (aged headshots of the same people). A collection of similarity percentages is generated by each algorithm, based on which we determined the threshold of match versus non-match.

In use case 2, we want to know how well the two packages recognize known faces in a group photo. Specifically, the algorithm compares source images (headshots) to target images (group photos), matches the faces in the images to the headshots, and produces labeled images with names and similarity scores.

# key findings

## use case #1

|  |  |
| --- | --- |
| ***facial\_recognition (open-source)*** | ***Amazon ReKognition*** |
|  |  |
| **Expected match case A** |  |
|  |  |
| *facial\_recognition*  Similarity: **63.26%** | *Amazon Rekognition*  Similarity: **54.70%** |
| **Expected match case B** |  |
|  |  |
| *facial\_recognition*  Similarity: **56.22%** | *Amazon Rekognition*  Similarity: **97.82%** |
| **Expected match case C** |  |
|  |  |
| *facial\_recognition*  Similarity: **29.60%** | *Amazon Rekognition*  Similarity: **0.57%** |
| **Expected NON-match case** |  |
|  |  |
| *facial\_recognition*  Similarity: **17.50%** | *Amazon Rekognition*  Similarity: **0.26%** |
| **Model accuracy (threshold = 50)** |  |
| *facial\_recognition*  Accuracy: 0.90  Precision: 0.98  Recall: 0.88 | *Amazon Rekognition*  Accuracy: 0.91  Precision: 1.00  Recall: 0.88 |
|  |  |
| **Key findings** |  |
| *facial\_recognition*   * Similarity scores are more scatterly distributed, with the majority of expected match in between 50 and 70 * Can still detect some facial features in blurred or partially covered headshots * 2 cases that should be non-match are predicted as matches | *Amazon Rekognition*   * Similarity scores are heavily weighted in the two ends, with the majority of expected match in between 70 and 100 * Not good at detecting expected match in blurred or partially covered headshots * Has a precision of 1, which means all the predicted matches are actual matches |

## use case #2

|  |  |
| --- | --- |
| ***facial\_recognition (open-source)*** | ***Amazon ReKognition*** |
| **Case A – Mask on** |  |
|  |  |
| *facial\_recognition*  Not detected or misdetected | *Amazon Rekognition*  Correctly detected |
| **Case B – Hat on** |  |
|  |  |
| *facial\_recognition*  Misdetected | *Amazon Rekognition*  Correctly detected |
| **Case C - Side face** |  |
|  |  |
| *facial\_recognition*  Not detected or misdetected | *Amazon Rekognition*  Correctly detected |
| **Case D – Front Face** |  |
|  |  |
| *facial\_recognition*  Misdetected | *Amazon Rekognition*  Correctly detected |
| **Case E – Face not in database** |  |
|  |  |
| *facial\_recognition*  Misdetected | *Amazon Rekognition*  Correctly detected as no match |
| **Model accuracy (threshold = 50)** | **(threshold = 80)** |
| *facial\_recognition*  Number of faces detected: 26/30  Mis-detection: 0.31  Avg similarity score of success detections: 54.57  Avg similarity score of misdetections: 51.25  \* We set the thredshold as 50 because most of the correct detections have sililarity score of above 50 | *Amazon Rekognition*  Number of faces detected: 30/30  Misdetection: 0  Avg similarity score of success detections: 97.85  Avg similarity score of misdetections: /  \* We set the threshold as 80 because all the correct detections have sililarity score of above 80 |
| **Key findings** |  |
| *facial\_recognition*   * Not detected or misdected for some mask-on or hat-on people * Not detected or misdected for some side faces * One face was misdetected as someone else * One face that is not in database was misdetected as someone else | *Amazon Rekognition*   * High accuracy detection for mask-on and hat-on people, with average similarity score of 85 * High accuracy detection for side faces, with average similarity score of 98 * All the faces were correctly detected * Face that is not in database was correctly detected as no match |

# Recommendations

In terms of performance, Amazon Rekognition is more accurate than facial\_recognition in general. When recognizing known faces in a group photo. Amazon Rekognition has a significantly higher similarity score of above 90 for most of the match cases while facial\_recognition. When recognizing the same person from different photos, the gap between the two algorithms is minor.

In terms of cost and maintenance, Amazon Rekognition has an annual subscription fee while facial\_recognition is an open-source free algorithm. However, facial\_recognition will need in-house modification and maintenance to tailor to the needs of the company.

In terms of data infrastructure, Amazon Rekognition will be hosted on the cloud while facial\_recognition is run in the local environment, which means Amazon Rekognition will save more local storage and computing power, and is potentially better in cyber security. Meanwhile, Amazon Rekognition will become a part of our data infrastructure if we are using other Amazon products.

I recommend facial\_recognition because the real case will be more similar to use case 1, where there will be a camera installed outside the building scanning visitors’ faces and matching them to the database. In this case, facial\_recognition was able to deliver an accuracy and precision score similar to that of Amazon, while also being free to use.

# Appendix

## Use case 2 labeled photos - facial\_recognition







## Use case 2 labeled photos – Amazon rekognition





