TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON)
31 Oct - 3 Nov 2023. Chiang Mai, Thailand

Attendance System using Amazon Rekognition

Ravi Kishore Kodali, Aniket Panda and Lakshmi Boppana
Department of Electronics and Communication Engineering
National Institute of Technology, Warangal
Warangal, 506004, India
ravikkodali@gmail.com

***Abstract*—This work proposes a cloud-based attendance system that uses face recognition technology to authorize identity. The system uses the Amazon Web Services (AWS) Rekognition service and a serverless architecture. The proposed system provides a reliable and tamperproof solution to track attendance, eliminating the need for manual record keeping and minimizing human involvement. It also offers potential benefits, such as improved security and transparency in attendance management.**

***Index Terms*—Face Recognition, Serverless, Machine Learning, AWS AI Services**

I. Introduction

Managing attendance in schools, universities, and workplaces has always been critical and time-consuming. Traditional attendance management methods rely on manual record keeping and are prone to human error and malpractice [5], making them unreliable and inefficient. As institutions continue to grow in size, there is increasing demand for accuracy, security, and transparency in attendance management. This, in turn, calls for advanced technology-based solutions that can automate the process and ensure a secure system. This work proposes a cloud-based face recognition attendance system that utilizes Amazon Web Services (AWS) Rekognition service and a serverless paradigm to meet this demand. The cloud architecture for this system is covered in III-A and the results are available in Section IV.

A. Background

The field of face recognition has been an integral part of computer vision since its inception due to its practical significance and its interest among cognitive scientists. Although fingerprint and iris scanning methods are often more precise, face recognition has remained a major area of research due to its non-intrusive nature and its widespread use as a primary means of identifying individuals. Over time, face recognition technology has advanced towards becoming a universal biometric solution, as it requires minimal effort on the user's part, unlike other biometric options [3]. All facial identification systems can be broken down into facial detection and facial recognition [1]. Face detection is a general purpose activity and is agnostic of the training data supplied. Face identification is a classification-type activity that sorts detected faces into available training categories. There are several problems faced by biometric attendance systems being implemented in practice, which are covered in Section II. Even in face-recognition-based attendance systems, the difficulties in obtaining standardized data sets, larger training volumes, and standardizing image capture devices (testing data) are significant. This includes but is not limited to the difficulty in extraction of facial features with changes in skin tone, lighting, shadows, noise, and other artifacts. This system aims to be robust enough to withstand strong inconsistencies within the data. People's faces change with time, including but not limited to distinctions in facial hair, glasses, hair color, and Personal Protective Equipment (PPE), that alter extracted feature topography [4]. The system needs to be able to evolve with changing data to remain accurate. Figure 1 shows the components associated with facial authentication.

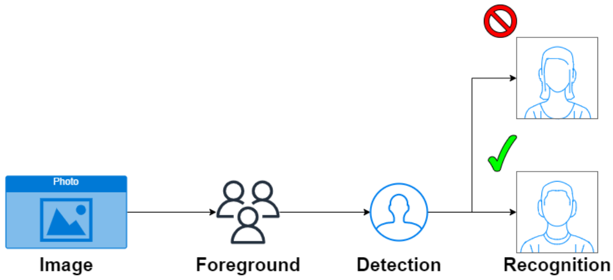


Fig. 1. Processes involved in facial authentication

B. Related Work

Kar [3] used an eigen-faces method where Haar cascade classifiers and the Paula-Viola Algorithm were used to extract characteristics and principal component analysis to classify them into saved training data (index of available faces). This system proved to be sensitive to the orientation of the face. Mittal [5] used a Haar selector, an Eigenfaces feature extractor, and an Adaboost estimator on OpenCV with video streaming as a method to capture validation data consisting of groups of individuals. Eigenface methods are noted to be very susceptible to rotations and tilts of the subjects' faces in Table I. Arsenovic [2] proposed a deep-convolution neural network (DCNN) comprising a CNN cascade to detect features, FaceNet CNN to detect face embedding, and a Support Vector Machine (SVM) to classify it into the preregistered index with a good accuracy of 95%. However, they required training data sets at the frontal, upward/downward/sideways tilts. The illumination and angular constraints while capturing training data, as well as the system's sensitivity to these, prove this to be an unfruitful strategy in deploying real-world applications

979-8-3503-0219-6/23/$31.00 ©2023 IEEE

with bad data. Lukas [4] used the Discrete Wavelet Transform and Discrete Cosine Transform to extract details and a Radial Basis Function (Neural) Network (RBFN) to classify them into the preregistered index of faces. Although its system is the most lightweight offering, it requires huge training data and can be seen to suffer when deployed at scale due to its inability to distinguish between similar faces. Pattnaik [7] used You Only See Once (YOLO) and Dlib (as well as Local Binary Histogram Patterns (LBHP)) for detection and recognition and presented a comparative study. They also used the AWS Rekognition API (Application Programming Interface) by capturing images of individuals and tabulating said data. They used single-face detection (see II-D for disadvantages). Kodali [1] used the Multi-Task Cascade Nueral Network (MTCNN) to detect faces and have compared EfficientNet B4 (in operation, now referred to as Eff-Net) with ResNet-50 for detection in groups. The proposed system ran on a t2.micro Elastic Cloud Compute (EC2) instance on AWS. Although this also offers a cloud implementation, it does so on an EC2 instance. Attendance systems tend to have extremely spiky traffic (at the start of class hours), which EC2 instances do not handle well. They are more suitable for consistent traffic than for spiky traffic.

TABLE I
ACCURACY OF VARIOUS FACE DETECTION AND RECOGNITION MODELS

Detection	Recognition	Accuracy (%)
Eigenface(frontal)	PCA	95 [3]
Eigenface(profile)	PCA	0 [3]
CNN Cascade	FaceNet/SVM	95 [2]
DWT/DCT (level 2)	RBFN	82 [4]
YOLO	LBHP	89 [7]
YOLO	Dlib	99 [7]
Haar	LBHP	81 [7]
Haar	Dlib	95 [7]
AWS	AWS	100 [7]
Haar/Eigenface	Adaboost	85 [5]
MTCNN	Eff-Net	94 [1]

II. LITERATURE REVIEW

A. Argument against other bio-metrics

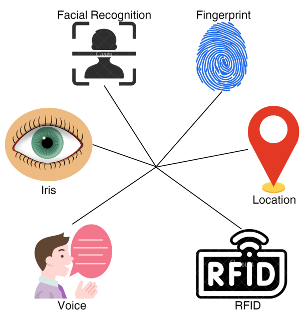


Fig. 2. All methods of bio-metric authentication

Figure 2 depicts a summary of various approaches used in biometric authentication. Arsenovic [2] proposed an RF-ID-based attendance system that utilized cards and reader devices to collect data on entrances and exits and transmit to a remote server via GP-RS and then store it in a database. Sharma [9] and Naen [6] also suggested similar RF-ID-based systems with varying degrees of complexity. Although the system is fairly robust, it cannot handle congested student traffic during class start and end times. It also does not verify students' identities and can be easily tricked by a malicious actor carrying someone else's RF-ID card. These deficits arise in addition to the expenses of otherwise single-use hardware. Thus, the proposed system should take advantage of the existing infrastructure to minimize the costs incurred in the production environment. The infrastructure should preferably be reuseable too and should not have performance bottlenecks with congested and spiky traffic. In addition, it should support a unique identifier that does not transfer to malicious actors as easily. Sultana [10] proposed a location-based attendance tracking system using GPS data extracted from an Android application. This can again be misused by bad agents, who can take advantage of approximate location data and report in the vicinity of classrooms. Continuous logging can take a toll on the device's battery life and prove impractical. The key takeaway is that any sort of system residing primarily on the student's device can prove to be a security risk. Rao [11] introduced an attendance system that relied on biometric authentication, specifically fingerprint verification using minutiae and pattern matching. This suffers from the same grievances as the RF-ID system, while also requiring more expensive hardware. The critical problem with specific-use hardware is that they tend to be neglected by institutions, and the entire system becomes unusable if even a single endpoint breaks. Thus, the capture unit has to be replaceable and the loss of a capture unit should not fail the entire system. Soewito [12] presented a smartphone-based attendance tracking system that integrated location and individual attributes, using fingerprint or voice recognition. Kadry [13] proposed an attendance system that utilized iris recognition through a wireless network. Although these biometrics perform well on paper, their effective utilization varies greatly according to the endpoint at which they are deployed. Modern iPhones do not support fingerprint scanners, while voice recognition models have varying success distinguishing between users, leading to loopholes for system abuse. Also, deploying the said application on the student's device is not preferable. Patil [14] applied face recognition for classroom attendance, using Eigenface for recognition, while Balcoh [15] used a similar approach for a face recognition-based attendance tracking system, achieving an overall recognition accuracy of 85%. Tharanga [16] used the principal component analysis (PCA) method for their attendance system, achieving an accuracy of 68%.

B. Argument against in-house algorithms

As can be seen from Table I in-house algorithms often produce poor results. This is a direct result of the costs involved in purchasing a large data set and then training a model on it [18]. For general-purpose tasks such as Human Face

Recognition, utilizing fully managed services from vendors is therefore in the interest of acquiring better accuracy. In practice, obtaining standard images for face recognition, that is, in ideal full frontal coverage [4], lighting conditions, equal image size [7], etc. is a difficult task. In-house algorithms tend to be very sensitive towards such outliers. They suffer a heavy blow to accuracy under adverse conditions, which makes them unsuitable. In real-world applications, users (students) can present training data in very diverse conditions, and capturing live data (lecturers) can prove arduous and have hazy artifacts and low quality. The proposed system should be resilient to poor data ingestion. From the same Table I, it can be seen that there is an acute discrepancy in the performance of DCNN algorithms compared to others. Improvement in Deep Learning has greatly improved the accuracy of facial recognition among other perception tasks. However, deep learning jobs can be very expensive to invoke and maintain [18]. This means that larger models are more suitable for our use, while their maintenance is difficult on a smaller scale. Higher accuracy is essential for our operation. Therefore, fully managed AI services are an important consideration for this task.

C. Argument against EDGE devices

Introducing specific hardware for attendance can be a costly matter for the institution as discussed in II-A. Especially since unit failure will require prompt service to prevent the entire system from being suspended. The attendance system needs to be highly reliable and available because if the system is only available selectively in some classrooms, the entire system will revert to manual attendance (owing to standardization preferences).

D. Cons - single-face detection and video streaming

Many employee attendance systems use single-face detection. Under review, only the work of Mittal [5] and Kodali [1] captures group photographs. This is imperative as having attendance requests from students unnecessarily populates the system in this context. This can also lead to malpractices, as students can record themselves and leave afterwards. Thus, it is better to give access to the infrastructure to the actor invested in providing more accurate readings. Group photos possess sufficient information to record attendance. It is considered a best practice for the system to ingest as little data as possible. However, there is one drawback with photographs that they fail to capture the liveliness of the subjects. What this entails is that it is hard for the system to detect whether there is a printed picture of the subject, any other malpractice or just the subject itself. But this can be considered to be a non-issue, as the lecturer can verify this in person, and video streaming solutions demand greater storage and bandwidth without offering any other advantage.

III. Proposed Methodology

A. Cloud Architecture

The cloud architecture, as shown in Figure 3, presents a serverless architecture using AWS Rekognition. This architecture divides the entire workflow into several lambda functions

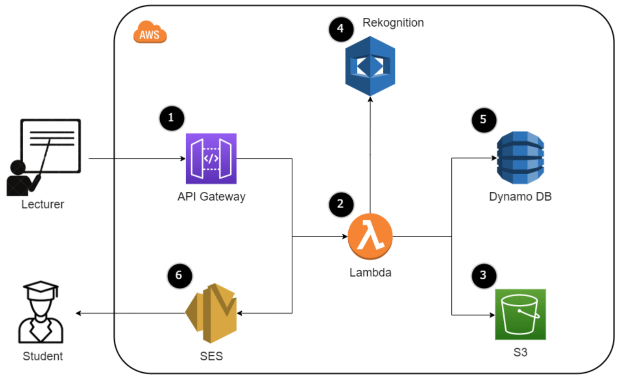


Fig. 3. Cloud Architecture of Proposed System

and uses collections from AWS Rekognition to assign unique face IDs. It stores the images and the results in a Simple Storage Service (S3) bucket and updates the database entries using the AWS DynamoDB service. The flow of information is as follows:

1) **Indexer**
   a) User uploads a file into the S3 bucket with prefix index.
   b) This triggers the lambda function Indexer.
   c) The lambda function sends a indexFaces() function call to the Rekognition API.
   d) The Rekognition API receives the call, and the S3 address of the indexer file and retrieves it.
   e) The lambda function then receives the response from the API after the face is indexed in Rekognition collections.
   f) The lambda function then populates the indices in the table.
2) **Recognition**
   a) The user uploads a file into the S3 bucket with prefix search.
   b) This triggers the Lambda function recognition.
   c) The lambda function sends a detectFaces() function call to the Rekognition API.
   d) The Rekognition API receives the call, and the S3 address of the indexer file and retrieves it. Returns a list of faces.
   e) The function then sends each detected face to be recognized. It populates the attendance list table.
   f) The lambda function then prints the bounding boxes with the roll number onto the images and uploads it onto the results prefix of the S3 bucket.
   g) The user then receives the printed index and the Excel call that includes all attendances from the results prefix of the S3 bucket.
3) **Cleaner**
   a) The lambda cleaner functions are triggered by the AWS Event Bridge schedule and automatically clean the S3 buckets.

b) The lambda function returns the successfully deleted 204 HTTP status code of the objects.

4) **Register**

a) The function is triggered by an API call on API Gateway, or as per event schedule. Extracts the values in the attendance list into a pandas data frame.

b) Extrapolates the necessary data, writes them onto the relevant database, and uploads the Excel sheet.

c) Sends relevant notifications to students through the simple email service (SES).

B. AWS Rekognition

AWS Rekognition is a fully managed AI service that offers a wide range of features including general object detection, face and person detection, facial recognition (through collections), Personal Protective Equipment (PPE) detection, and detection of face liveness (attention). It is based on the MXNet framework and is trained on very large data sets [23]. It offers a latency of 1.08 seconds when processing 3000 files in batches of 10 on N.Virginia servers [19]. One of its features, Amazon Rekognition Face Liveness is a tool that helps developers deter fraud in face-based identity verification [23]. In its operation, VGA (640x480) resolution or higher is recommended, greatly enhancing its ability in low-resolution imaging [22]. Furthermore, Amazon Rekognition can interpret emotional expressions such as happy, sad, or surprise, and demographic information such as gender or age from facial images. It can also detect facial occlusion when the eyes, nose, and / or mouth of a face are blocked by dark sunglasses, masks, or hands [22]. Compared to Google Vision, Rekognition has an advantage with its price factor and face detection at various angles [21]. In 2018, Amazon Rekognition was updated to be up to 80% more accurate in distinguishing between similar people, and up to 35% more accurate in recognizing dramatic changes in a person (hairstyle, hair color, facial hair, glasses, etc.) [20].

IV. RESULTS

A. Costs

The primary benefit of the system is how resource-sparse its deployment is. Given in Table II are the operating costs for this system. The prices are calculated for a standard batch size of 180 students registered for a total of 8 courses within a department. It is assumed that all 8 courses are conducted within the span of a day for all working days of a month, and recording attendance for each course involves 4 images to be captured to faithfully represent all present subjects in the situation. Each image used for both training as well as testing purposes is assumed to be 4 Megabytes in length.

The costs can be broken down into the costs incurred by the lambda functions for their invocations, the storage of all media, upload and transfer of said media, image recognition APIs, notifications, and system API calls by foreign systems. The indexer lambda function is called once per semester, for each and every student. This totals 180 invocations for the calculation of monthly costs and 360 for annual costs. All other costs are regular costs, so their monthly costs scale up to yearly costs. The lambda functions have the following configuration of architecture (x86), amount of ephemeral storage allocated (512 MB), invoke the mode (Buffered), and the number of requests (180) for Indexer. The recognition lambda function in Figure 3 is the architecture (x86), the amount of ephemeral storage allocated (512 MB), the invoke mode (buffered), the number of requests (7000 per month), time for which provisioned concurrency is enabled (100 hours), number of requests for Provisioned Concurrency (700 per month), Concurrency (3). The recognition lambda function is configured to have a total provisioned run time of 30 s. The average invocation/billing time was recorded at 11 seconds. However, for billing purposes, it is estimated that the function call is for 30 seconds. S3 storage is configured for standard storage (42 GB per month) and data transfer from across the Internet (42 GB per month). The Rekognition API is configured as follows Number of IndexFaces API calls exact (180), Number of SearchFacesByImage API calls per month (70 thousand per month). SES is configured for email messages sent from email clients (80 per month). The API Gateway costs were calculated by estimating HTTP API requests units (exact number), the average size of each request (34 KB), REST API request units (1 million), cache memory size (GB) (none), WebSocket message units (thousands), average message size (32 KB), requests (80 per day). All reported values from the AWS Cost Calculator are tabulated for the Mumbai region (ap-south-1) in Table II.

$$\text{projected costs (in USD)} = 336.46 \times n \times c/(180 \times 8) \quad (1)$$

where n represents the number of students and c represents the number of courses 180 is the number of students considered and 8 is the number of courses considered.

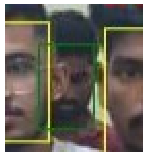
Equation 1 can be used to determine the estimates incurred for the particular use case, as the system can be seen to be not only highly configurable but also very linearly scalable to the size of operation and costs.

TABLE II
COST ANALYSIS OF PROPOSED SYSTEM (IN US$)

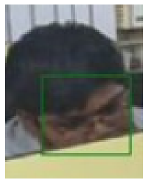
Functionality	Service	Upfront	Monthly	Annual
Indexer	Lambda	0	0	0.00
Recognition	Lambda	0	0.44	5.28
Storage	S3 Standard	0	1.12	13.44
Upload	Data Transfer	0	0	0.00
Face recognition	Rekognition	0	26.35	316.20
Notifications	SES	0	0.128	1.54
RESTful APIs	API Gateway	0	0.0026	0.03
Total	All	0	28.0406	336.46

B. Discussion

As can be seen in Figure 4, face detection has been largely very successful. All persons on the scene have been identified with significant confidence and have thresholds greater than 99. 8%. Even those with their backs turned and barely visible



(a)



(b)

(c)

Fig. 4. Results of the proposed system in adverse conditions. All faces were detected. (a) The face of this subject lacked sufficient embedding information for identification, as key features like nose and eyes are not clearly visible (b) The face of this subject lacked a complete vertical description, without the fuller segmentation of the identity of the nose and mouth could not be resolved. (c) The training data of this subject (index) consisted only of a full frontal image (as in the others), but they were able to recognize a profile with 99.9991% accuracy

faces have been detected. This is consistent across multiple scenes and test conditions. Recognition can be seen to have struggled with some subjects. Subject (a) in Figure 4 did not contain enough detail for successful recognition. Subject (b) has the bottom half of their face hidden, which prevents proper extraction of features. The dense population of the image does not impede the system from recognizing a modest number of faces with high confidence. The lowest reported confidence threshold was 99.99919891357422 of the subject (c), which is a side profile and has a beard artifact in contrast to the frontal training data provided. This implies that longitudinal cross sections and side profiles can be sued to extract face embedding for accurate recognition.

The data shown in Figure 4 are believed to be a faithful rendition of the worst-case real-world data. The results obtained are believed to be within the range of success for implementing this system on a wide scale.

Fig. 5. Sample of training data in diverse conditions

V. Conclusions

This work proposes a cloud-based face recognition attendance system that uses the AWS Rekognition service and a serverless architecture. The system offers a reliable and tamperproof solution for tracking attendance, eliminating the need for manual record keeping and minimizing human involvement. The proposed system provides a highly scalable and available solution for institutions such as schools and universities. Using AWS Rekognition, the system provides high accuracy and real-time performance. The cloud architecture and serverless paradigm enable the system to handle large workloads with ease, while also reducing costs and complexity. The proposed system also offers potential benefits, such as improved security and transparency in attendance management. Future work could focus on further improving the system's accuracy, enhancing user experience, and investigating the potential for integration with other technologies such as facial recognition-based access control systems. This work shows that a cloud-based face recognition attendance system can offer a robust, scalable, and automated solution for attendance management in various institutions.

References

[1] R. K. Kodali and R.V. Hemadri, ”Attendance Management System,” 2021 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2021, pp. 1-5, doi: 10.1109/ICCCI50826.2021.9402659.

[2] M. Arsenovic, S. Sladojevic, A. Anderla, and D. Stefanovic, “FaceTime - Deep learning-based face recognition attendance system,” in the IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY), Subotica, Serbia: IEEE, Sep. 2017, pp. 000053–000058. doi: 10.1109/SISY.2017.8080587.

[3] N. Kar, M. K. Debbarma, A. Saha and D. R. Pal, 'Study of Implementing an Automated Attendance System Using Face Recognition Technique', IJCCE, pp. 100–103, 2012, doi: 10.7763/IJCCE.2012.V1.28.

[4] S. Lukas, A. R. Mitra, R. I. Desanti, and D. Krisnadi, 'Student attendance system in classroom using face recognition technique', in 2016 International Conference on Information and Communication Technology Convergence (ICTC), Jeju: IEEE, October 2016, pp. 1032–1035. doi: 10.1109/ICTC.2016.7763360.

[5] A. Mittal, F.S. Khan, P. Kumar, and T. Choudhury, “Cloud-based intelligent attendance system through video streaming,” in 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), Bangalore: IEEE, Aug. 2017, pp. 1352–1357. doi: 10.1109/SmartTechCon.2017.8358587.

[6] M. F. Naen, M. H. Muhamad Adnan, N. A. Yazi, and C. K. Nee, 'Development of Attendance Monitoring System with Artificial Intelligence Optimization in Cloud', International Journal of Artificial Intelligence, vol. 8, no. 2, pp. 88-98, December 2021, doi: 10.36079/lamintang.ijai-0802.315.

[7] P. Pattnaik and K. K. Mohanty, 'AI-Based Techniques for Real-Time Face Recognition-Based Attendance System- A Comparative Study', in the 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA) 2020, Coimbatore, India: IEEE, Nov. 2020, pp. 1034–1039. doi: 10.1109/ICECA49313.2020.9297643.

[8] E. Strubell, A. Ganesh, and A. McCallum, 'Energy and Policy Considerations for Deep Learning in NLP' 2019, doi: 10.48550/ARXIV.1906.02243.

[9] Sharma, Saumya, S. L. Shimi and S. Chatterji. ”Radio frequency identification (RFID) based employee monitoring system (EMS).” International Journal of Current Engineering and Technology 4.5 (2014): 3441-3444.

[10] Sultana, Shermin, Asma Enayet and Ishrat Jahan Mouri. ”A SMART, LOCATION BASED TIME AND ATTENDANCE TRACKING SYSTEM USING ANDROID APPLICATION.”International Journal of Computer Science, Engineering, and Information Technology (IJCSEIT) 5.1 (2015).

[11] Rao, Seema, and K.J. Satoa. ”An attendance monitoring system using biometrics authentication.” International Journal of Advanced Research in Computer Science and Software Engineering 3.4 (2013).

[12] Soewito, Benfano, and others ”Smart mobile attendance system using voice recognition and fingerprint on smartphone.” Intelligent Technology and Its Applications (ISITIA), 2016 International Seminar on. IEEE, 2016.

[13] Kadry, Seifedine, and Mohamad Smaili. ”Wireless attendance management system based on iris recognition.” Scientific Research and Essays 5.12 (2013): 1428-1435.

[14] Patil, Ajinkya, and Mrudang Shukla. ”Implementation of Classroom Attendance System Based On Face Recognition In Class.” International Journal of Advances in Engineering & Technology 7.3 (2014): 974.

[15] Balcoh, Naveed Khan, and others ”Algorithm for efficient attendance management: Face recognition based approach.” IJCSI International Journal of Computer Science Issues 9.4 (2012): 146-150.

[16] Tharanga, JG Roshan and others ”SMART ATTENDANCE USING REAL TIME FACE RECOGNITION (SMART-FR).” Department of Electronic and Computer Engineering, Sri Lanka Institute of Information Technology (SLIIT), Malabe, Sri Lanka.

[17] N.K.Jayant, S.Borra.“Attendance Management System Using Hybrid Face Recognition Techniques', Conference on Advances in Signal Processing (CASP), Cummins College of Engineering for Women, Pune, 2016.

[18] E. Strubell, A. Ganesh, and A. McCallum, “Energy and Policy Considerations for Deep Learning in NLP,” 2019, doi: 10.48550/ARXIV.1906.02243.

[19] https://blog.filestack.com/thoughts-and-knowledge/comparing-google-vision-microsoft-cognitive-amazon-rekognition-clarifai/

[20] https://aws.amazon.com/about-aws/whats-new/2018/04/amazon-rekognition-improves-accuracy-of-real-time-face-recognition-and-verification/

[21] https://webuters.medium.com/google-cloud-vision-vs-amazon-rekoginition-which-is-better-c7cb4780d82a

[22] https://docs.aws.amazon.com/rekognition/latest/dg/what-is.html

[23] https://aws.amazon.com/rekognition/image-features/