IoT83

From Data to Decisions

Quickest Way to Build AI-powered IoT and Machine Vision Application for OEMs.

Our Technical Know-Hows

TABLE OF CONTENTS

IoT83

INDUSTRY STATS

77%

of manufacturers consider machine vision important for meeting their business goals.

–IBM Report

51%

of the global machine vision market is covered by its industrial segment alone.

–Grand View Research

7% CAGR

Machine Vision market CAGR 2023-2030, with manufacturing as one of its fastest-growing segments.

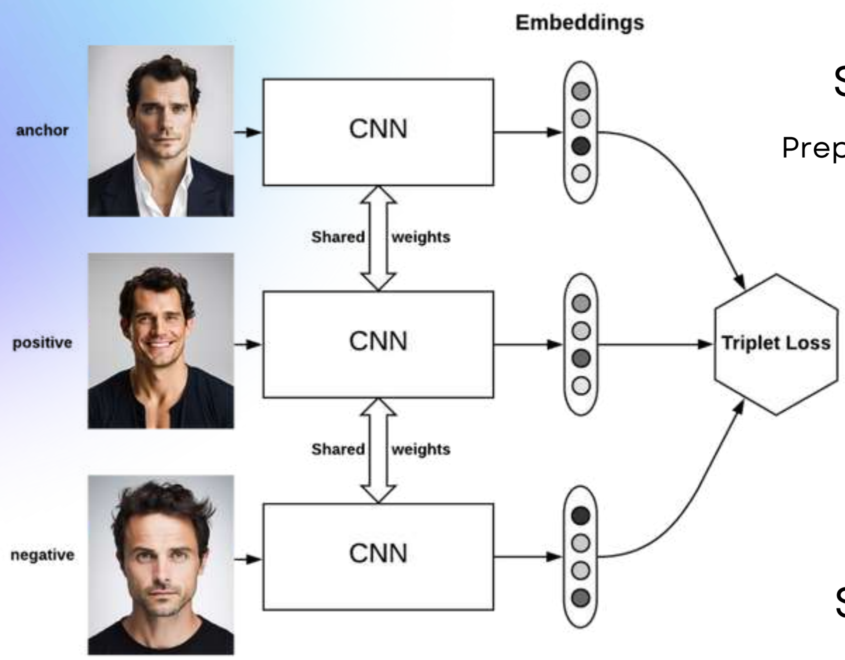
–Mordor Intelligence,

The Market Growth can be attributed to the rising need to implement machine vision solutions enabling industry smartification and integration of advanced Vision AI technologies using AI ML capabilities of IoT Solutions in a wide range of Industrial OEMs.

AI Vision Applications Industry Use Cases

These use cases can be broadly classified as:
- Face/Object Recognition
- Security Management
- Traffic Flow Analysis
- Medical Analysis
- Inventory Management
- Predictive Maintenance
- Intelligent Image/Video Analytics

MODEL CREATION



Our approach here is based on Face Classification via a Five-Stage Detection Process followed by multiple pipeline iterations and feedback modifications to find the face and its location with high accuracy.

Stage 1

Data Gathering

Stage 2

Preprocessing

Stage 3

MTCNN

Stage 4

FaceNet & Embedding

Stage 5

Model Creation and Hyperparameter Tuning

Introduction to AEP

The Internet of Things (IoT) transforms industries and how we interact with technology. As connected devices grow, businesses seek efficient ways to manage their IoT deployments. It is where Application Enablement Platforms (AEPs) come in. AEPs provide a comprehensive suite of tools and services that enable businesses to develop, deploy and manage IoT applications quickly.

What is an IoT Application Enablement Platform?

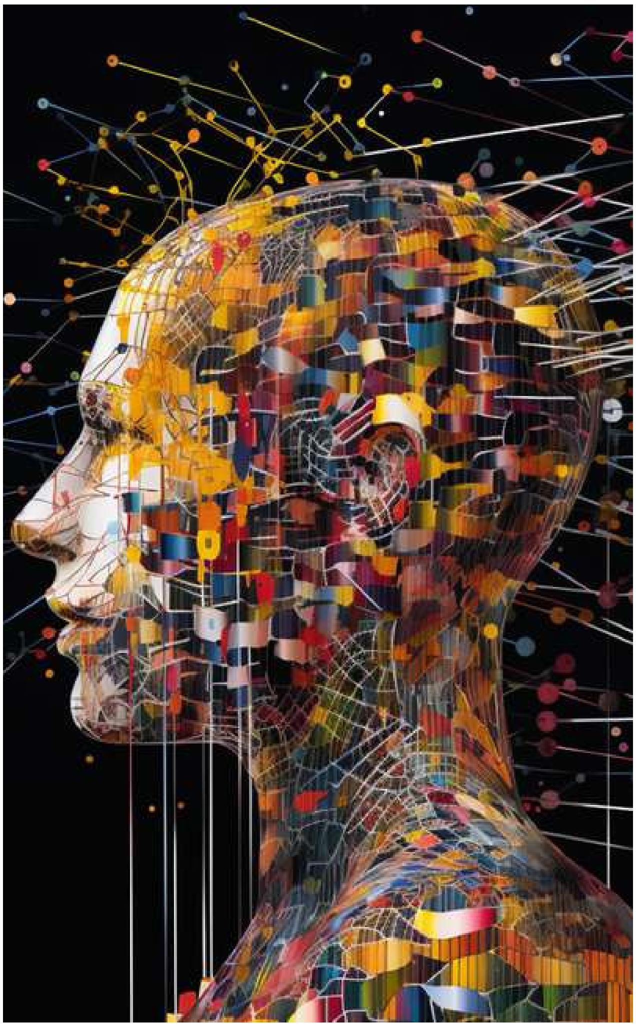
An IoT Application Enablement Platform (AEP) is a middleware platform that connects IoT devices, applications, and services. It provides tools and services that enable businesses to develop, deploy and manage IoT applications. The platform abstracts the complexity of IoT deployments and provides an easy-to-use interface for developers, allowing them to focus on building applications instead of addressing the infrastructure.

Key Features of an AEP

Industrial IoT Application Enablement Platform (AEP) offers various key features that enable businesses to develop and deploy IoT solutions. These features include device management, data analytics, and integration with third-party services. However, more than having an AEP alone is needed to overcome IoT development challenges.

Every Industrial OEM requires an IoT Solutions partner who can help them scale their IoT project from MVP to a fully functional solution and customize it to their unique business needs. A reliable development partner can also help businesses overcome common hurdles in IoT development, such as security, compatibility, and reliability issues, and ensure that the IoT solution is built for success.

Abstract



Purpose: This white paper aims to present the development and optimization of an advanced Employee Recognition System utilizing Flex83's AI ML capabilities. The goal is to showcase a robust solution for accurate and real-time employee identification for security management and time tracking within organizational settings.

Methods: The solution is built upon a multi-stage pipeline integrating Multi-task Cascaded Convolutional Networks (MTCNN) for face detection and FaceNet for face embedding. Leveraging Flex83's AI ML capabilities, the system continuously evolves through a Feedback Network, adapting to user input and refining recognition accuracy.

Results: Integrating Flex83's AI ML capabilities significantly enhances the accuracy and efficiency of employee recognition. The MTCNN-based face detection and FaceNet-powered embedding contribute to precise identification, even in dynamic scenarios. Fueled by user feedback, the iterative model refinement process optimises ongoing performance.

Conclusion: By harnessing Flex83's AI ML capabilities, organizations can deploy a state-of-the-art Employee Recognition System beyond traditional methods. The solution's adaptability, accuracy, and real-time capabilities are powerful tools for enhancing security and optimizing employee management.

Overview

In recent years, rapid advancements in artificial intelligence have propelled biometric recognition, encompassing face recognition, voice recognition, fingerprint recognition, iris recognition, and more, to a pivotal position within the realm of AI. As current identification methods like smart cards and passwords yield limitations, the biometric sector is poised to capture a significant market share. With its evident advantages, biometrics has found widespread application in sectors ranging from finance, transportation, and e-commerce to law enforcement and enterprise management. Particularly, face recognition has swiftly evolved, showcasing substantial progress in accuracy and gaining global attention.

However, challenges like varying illumination, intricate backgrounds, and diverse facial angles have impeded optimal performance. Deep learning, particularly convolutional neural networks, has emerged as a powerful tool for processing structured facial images, combining innate facial knowledge with advanced image processing to drive accurate recognition.

The evolution of AI and ML has opened doors to revolutionary advancements in facial recognition. However, achieving high accuracy and real-time performance remains a challenge. Flex83's AI ML capabilities offer a powerful platform to overcome these challenges, empowering the system to process intricate tasks and deliver reliable results.

This paper draws from extensive research in facial recognition, machine learning algorithms, and real-time deployment techniques to establish a holistic solution that addresses the diverse needs of modern organizations. We will delve into the Face Detection training we conducted to develop a classification to identify a face and its location with high accuracy. The face recognition model based on a Multi-task Convolutional Neural Network (MTCNN) is combined with multiple face detection training and real-time classification deployment at the agent.

Problem Statement

In the context of modern organizational dynamics, the conventional manual methods employed for security management and tracking employee attendance have proven to be not only time-intensive but also riddled with inaccuracies.

While offering a semblance of efficiency, the prevalent Radio Frequency Identification (RFID) card system introduces its own set of challenges. Each employee is assigned a unique RFID card containing their identity information. However, this approach leaves room for potential card misplacement or unauthorized usage, ultimately leading to inaccurate intruder detection and attendance records. Additionally, while other biometric techniques such as fingerprint, iris, or voice recognition have emerged, they, too, present limitations that hinder their accuracy and efficacy.

For organizations that operate with a substantial workforce, ensuring the precision of attendance records becomes a formidable challenge. The inadequacies of existing methods jeopardise operational efficiency and cast doubt on the reliability of crucial processes such as payroll management and resource allocation.

Thus, a compelling organizational need arises for implementing a robust and dependable employee recognition system—one that is accurate, reliable, and capable of accommodating a large and diverse workforce. Addressing this need is essential to foster streamlined attendance management and fortify the foundation of organizational efficiency.

Our Approach

Our approach here is based on Face Classification via a Five-Stage Detection Process followed by multiple pipeline iterations and feedback modifications to find the face and its location with high accuracy.

Face Detection (Training)

Stage 1: Data Gathering

The initial phase of our process centred around the critical aspect of data gathering–specifically, acquiring facial images from every candidate or employee for subsequent face detection. To facilitate this, we developed a compact application, which was then disseminated to all users. This application prompted users to capture a short video encompassing various facial angles. The resultant videos were seamlessly uploaded to a designated S3 bucket using the integrated IoT83 storage service. A structured directory emerged after uploading, organizing each user's content based on their unique identifier (GUID: User's email) and corresponding video content.

Stage 2: Preprocessing

Following the comprehensive data-gathering phase, the subsequent step involved transforming video files into images and their subsequent classification into testing and training datasets. This process was expertly executed using the OpenCV library. Converting the video files into individual images allowed for enhanced versatility in subsequent processing. To enrich the dataset, an initial level of data augmentation was performed. Specifically, approximately 20% of the images from each individual's dataset were allocated for testing, while the remainder was earmarked for training purposes.

Stage 3: MTCNN

The third stage of our methodology revolves around the Multi-task Cascaded Convolutional Networks (MTCNN), a pivotal component designed to identify faces within the images and meticulously eliminate extraneous details (noise) that could confound subsequent analyses. This process is significant in refining the images to their core facial components while also extending its capacity to identify additional facial attributes such as ears and the nose. MTCNN stands as a robust tool for ensuring the focus is exclusively on the facial elements vital for accurate identification and recognition.

About MTCNN:

MTCNN stands as a cornerstone in our methodology, delivering state-of-the-art results in the realm of face detection. This remarkable algorithm is constructed upon a cascade of three Convolutional Neural Networks (CNNs), ingeniously interconnected to progressively enhance the precision of face detection. The three sequential stages of MTCNN unfold as follows:

Step 1: The Proposal Network (P-Net)

In the initial phase, a multi-scale approach is adopted, generating a pyramid of images in varying dimensions, ranging from the top left corner to the bottom right corner. The P-Net, or the Proposal Network, takes centre stage in this stage. It operates as a shallow yet highly effective fully connected CNN. P-Net is pivotal in identifying candidate windows within the images and computing their associated bounding box regression vector. A crucial step in this process is the application of Non-Maximum Suppression (NMS), expertly filtering the bounding boxes proposed by P-Net. Only the most promising candidates proceed to the subsequent stage, ensuring optimal precision and accuracy.

Step 2: Refinement via CNN and Facial Feature Localization (Refinement Network)

Building upon the outputs of the previous stage, Step 2 leverages another CNN to further refine the results. This refinement process entails the rejection of non-facial frames, effectively isolating facial features of interest. At the heart of this stage lies the creation of a bounding box representing the detected face, characterized by a four-element vector. An element vector is also meticulously generated, contributing to the localization of intricate facial features.

Step 3: Enhanced Localization with Output Network (O-Net)

The final phase, Step 3, introduces the Output Network (O-Net) to the cascade. This potent CNN takes input from the previous stage, where bounding boxes are resized to a standardized 48 x 48 pixels. The outcome of this stage significantly advances the precision of facial feature localization. Notably, O-Net produces three distinct outputs: the coordinates of the bounding box (out[0]), the coordinates of five essential facial landmarks (out[1]), and the confidence level assigned to each bounding box (out[2]). This multi-faceted output equips subsequent stages with invaluable information, optimizing the overall face detection process.

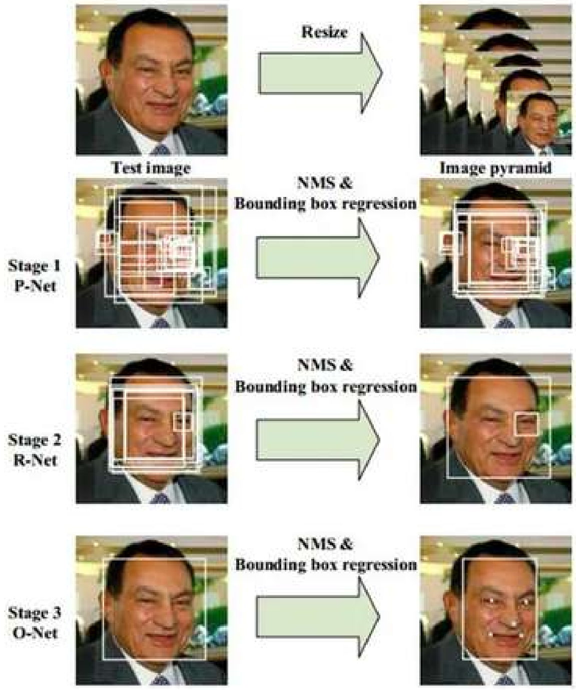


Image Credits IEEE Report

Upon successfully detecting a face in the input image using the Multi-task Cascaded Convolutional Networks (MTCNN), a pivotal transition occurs to process and organize the obtained facial data. This phase involves the creation of a structured dictionary, wherein each individual's name serves as the key. The data corresponding to the MTCNN's output—comprising facial feature coordinates, bounding box information, and confidence levels—is meticulously organized within this dictionary. The resultant dictionary is then meticulously stored as a .npz file, encapsulating the rich facial data in a compact and efficient format.

This process of assembling and archiving facial data is a crucial bridge between the intricate MTCNN-based face detection and subsequent stages of our methodology.

Organizations can lay the groundwork for subsequent processing, recognition, and analysis by effectively structuring the detected facial features and their associated attributes. This storage mechanism facilitates streamlined data management and preserves the invaluable information extracted during the face detection process, fostering a seamless transition toward subsequent stages of our innovative employee recognition system.

Stage 4: FaceNet and Embedding

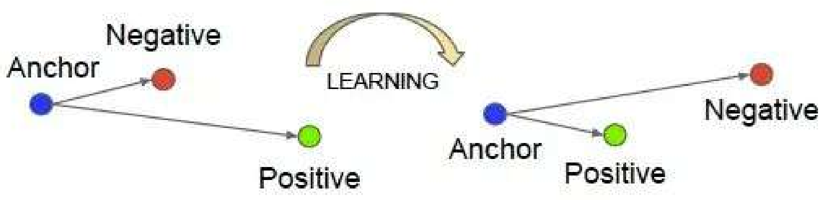
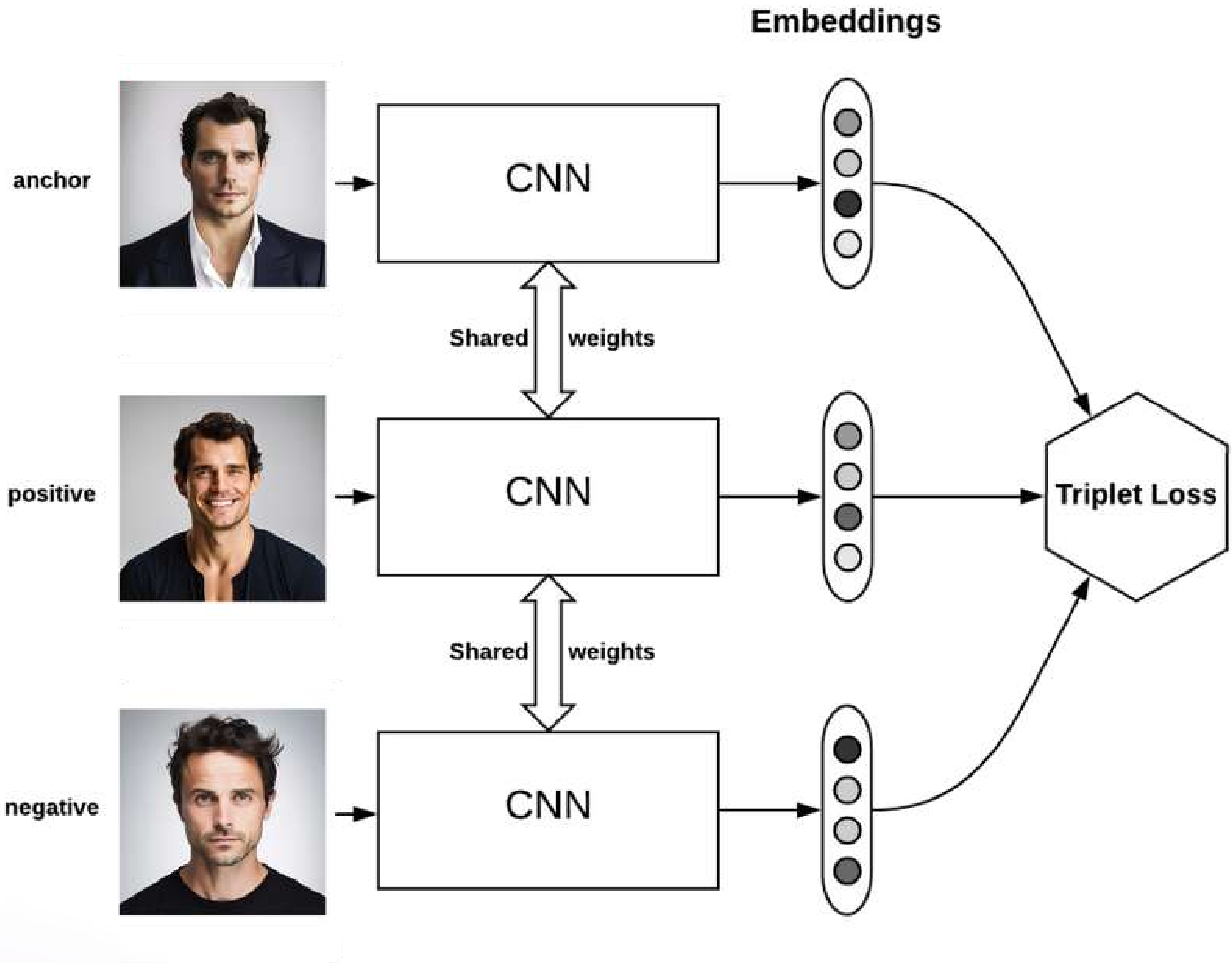


Image Credits Google Inc. Report

Transitioning to the fourth stage, we delve into the integral realm of FaceNet and the pivotal process of embedding. Within this stage, the outputs generated by the third phase undergo a profound transformation through the utilization of a Convolutional Neural Network (CNN) model–FaceNet. This particular model is meticulously trained employing the intricate Triplet Loss Function, a sophisticated technique that mandates the presence of three distinct images: anchor, positive, and negative.



The foundational principle driving the triplet loss function is to ensure that the anchor image resides closer to the positive images than the negative ones. This intrinsic logic facilitates the creation of a discriminative feature space, enhancing the system's ability to accurately differentiate between individuals. FaceNet, with its advanced architecture and intricate training, adeptly extracts the most salient facial attributes, culminating in generating a 128-element vector—commonly referred to as embeddings.

These embeddings, which encapsulate the essence of an individual's facial features, play a pivotal role in our recognition system. Each image is meticulously assigned an embedding, intricately labeled with the person's unique identification (email), and systematically stored as a .npz file. This storage mechanism streamlines subsequent retrieval and processing and preserves the distinctive attributes essential for accurate identity prediction.

Stage 5: Model Creation and Hyperparameter Tuning

In the final stage of our progressive methodology, we embark on developing a classifier model that holds the key to predicting the identity of a given face. This sophisticated model leverages the embeddings generated in the preceding stages, infusing our recognition system with the ability to make informed and accurate predictions.

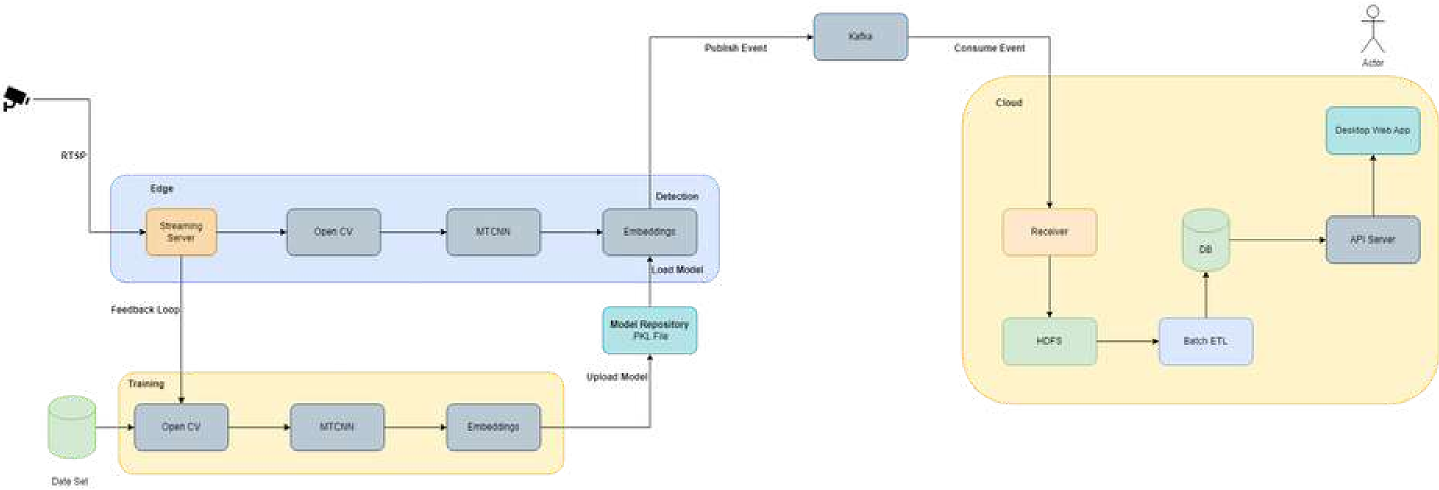
To ensure the optimal performance of this model, we did some hyperparameter tuning guided by the innovative IoT83 face-classification API, which integrates the robust GridSearchCV technique. This approach systematically explores a range of hyperparameter values, discerning the combination that yields the highest predictive accuracy.

Through this intricate amalgamation of cutting-edge model creation and strategic hyperparameter tuning, our recognition system is primed to deliver precise and reliable identity predictions. These stages embody the culmination of our innovative approach to employee recognition, empowering organizations with a transformative tool for streamlined workforce management and operational excellence.

Real Time Face Classification Deployment

The seamless transition from the development stages leads us to the next real-time face classification deployment phase—a transformative juncture in our FaceRec Methodology. Within this phase, a comprehensive approach is undertaken to ensure the smooth operationalization of our carefully crafted recognition system.

Deployment Diagram



Deployment Flow Diagram

Upon the successful deployment of the model to the designated agent, a sophisticated orchestration unfolds to drive the detection process. This sequence of events is briefly outlined as follows:

Capture and Conversion of Video: The deployment commences with the agent capturing video footage from the designated camera, operating at 50 frames per second (FPS). Employing the OpenCV library, the captured video is converted into a series of images—forming the basis for subsequent analysis.

Queue-based Image Transmission: The "Capture module" transmits batches of X images to the Multi-task Cascaded Convolutional Networks (MTCNN) queue.

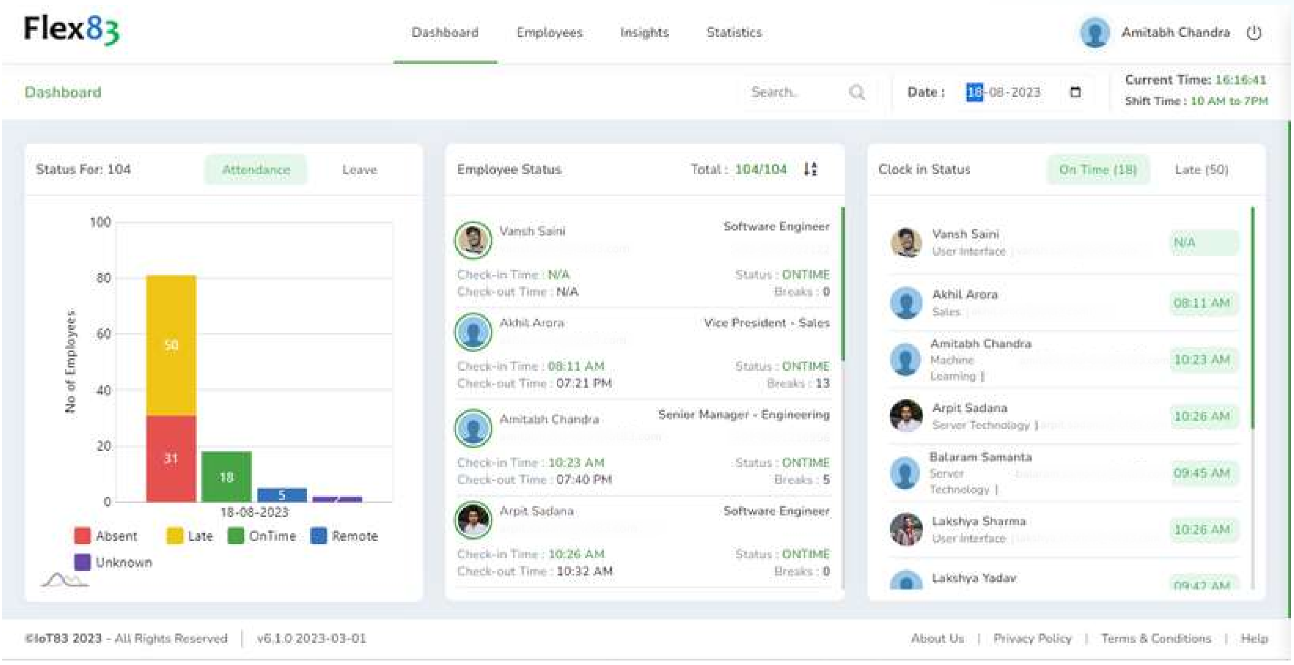
MTCNN-based Face Detection: Within the MTCNN module, the images are extracted from the designated queue, initiating the process of face detection. This step involves intricate decision-making, carefully calibrated through experimental refinement, guided by preset parameters such as min_face_size, scale_factor, and steps_threshold. A key criterion for successful detection is a confidence level exceeding 0.98, guaranteeing the reliability of the identified faces. Upon successful detection of 'Y number of face-detected objects,' they are put into the processing queue.

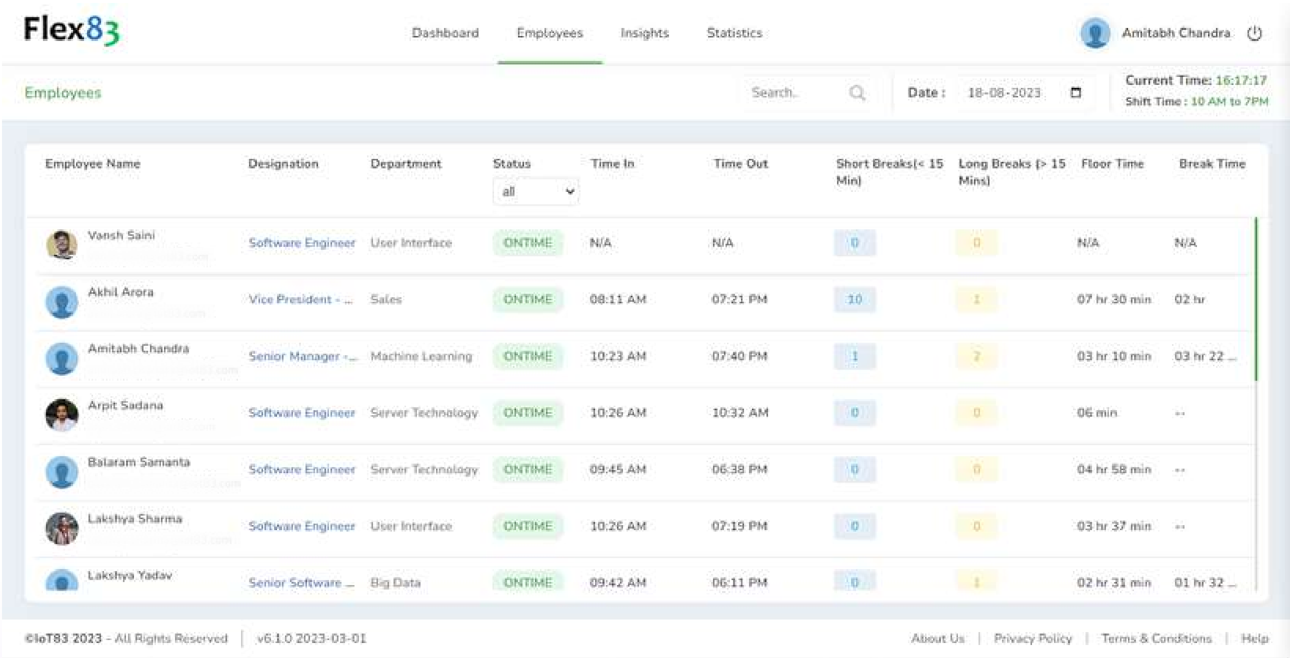
Embedding Generation and Detection Model Call: In the subsequent stage, the Processing module methodically processes each message extracted from the queue. This involves the creation of embeddings—representing the intricate facial attributes—and subsequent interaction with the detection model (using Support Vector Machines (SVM) in the current release). The detection model endeavours to match the embeddings with the identities of known individuals, enabling accurate person detection.

Data Packaging and Cloud Transmission: Upon identification, a comprehensive JSON payload is formulated, incorporating parameters such as name, GUID, timestamp, and event details. This crafted JSON is then transmitted to the cloud infrastructure through the efficient Kafka communication protocol.

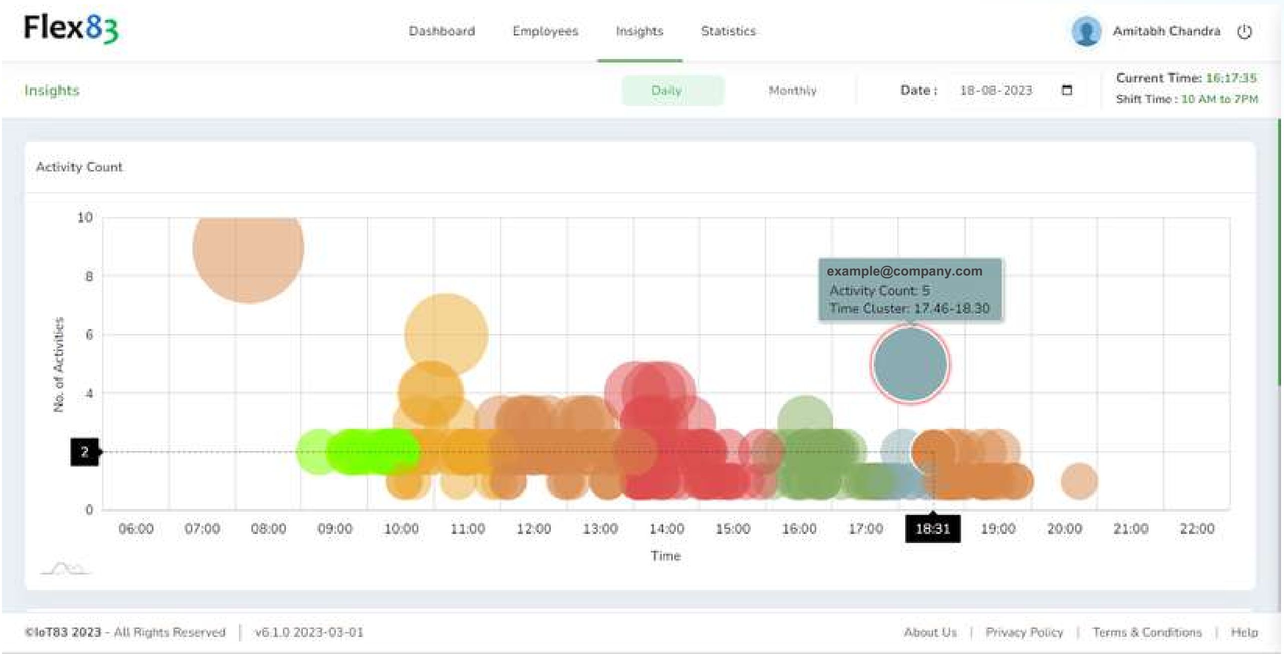
Cloud-based Processing and Insights: The cloud infrastructure efficiently processes transmitted data. The insights garnered from this cloud-based processing bolster managerial efficacy and enhance operational insights.

Dashboard Insights





Dashboard Insights



Results & Discussions

First Version of the Pipeline

Challenge

Our journey commenced by acquiring user-generated images converted from videos. However, initial model predictions were disappointingly inaccurate. This was due to several issues we encountered:

- **Imbalanced Data**: Variability in video lengths, ranging from 2 to 6 seconds, resulting in an imbalance in the number of images per person.
- **Lack of Augmentation**: Data augmentation was initially neglected, leading to limited model performance.
- **Image Quality Variation**: Disparities in image quality emerged between high-quality videos and lower-resolution images captured from suspended cameras, often resulting in blurred frames.

Solution

To address these challenges, we pursued a multi-faceted solution:

- **Imbalanced Data:** We collected additional data to equalize image counts, gathering approximately 160 images per person. These were divided into 100 for training, 30 for testing, and 40 for augmentation.
- **Augmentation:** Leveraging 40 images, we executed techniques such as cropping, rotation, and blurring. Care was taken to prevent excessive blurring that could hinder facial recognition.
- **Image Quality:** We ensured proper alignment of cameras for improved image quality and proximity.

Result

These interventions yielded a substantial 20% increase in model accuracy.

Second Version of the Pipeline

Challenge

While initial enhancements were promising, many false positives persisted, especially as individuals moved out of focus within 2-3 seconds of camera exposure. With our camera at 50FPS, we could capture enough images of a person/persons moving from the camera.

Solution

To mitigate this, we developed an algorithm for precision enhancement:

- A window of X frames was created.
- For a person to be considered a match, they needed to be detected Y times within the X-frame window.
- Individuals successfully matched had X frames to prove Y times of detection.

Result

This approach mitigated false positives, offering better accuracy as individuals stayed within the frame for a longer duration.

Third Version of the Pipeline

Challenge

Although false positives were reduced, challenges remained in detecting multiple people within a single frame. To address this, we addressed false face bounding boxes generated by MTCNN.

Solution

- Frames were collected with detection names, bounding box coordinates, and probabilities.
- We corrected mislabeled persons and adjusted inaccurate bounding boxes.
- Augmentation and retraining improved model performance.
- To address the false face issue, we engaged in distance analysis using MTCNN's key points (eyes, nose, mouth). We also tuned MTCNN parameters for optimal rejection of false faces. We tried through various distance analyses to find the correct face (like the minimum distance between the left eye, the left side of mouth w.r.t. nose similarly with the right eye, right side of mouth, and the difference between left and right eye and so on). We got good results, but the performance was still affected.
- We played with configuration parameters of the MTCNN like min_face_size, scale_factor, and steps_threshold. We got the optimized value, which mostly rejected false faces without reducing much performance.

Result

After this, false positive were reduced, as well as the detection of multi-person in a frame was increased.

Fourth Version of the Pipeline (Feedback Network)

To further refine the model's performance, we introduced a Feedback Network, which follows these key steps:

- **Data Storage:** The Agent now systematically stores images, along with comprehensive metadata, including timestamps, detected individuals, and probabilities. This organized repository employs a specific path and directory structure.
- **Cloud-Driven Upload:** A separate, slower process within the Agent is dedicated to uploading this data to the Cloud for subsequent processing.
- **UI-Empowered Evaluation:** Our UI boasts multiple APIs designed for image assessment. Upon presentation, the UI displays images, along with their current bounding boxes and detected individuals. Users are empowered to make decisions based on the following conditions:
    - If the label and bound box are both correct, no alterations are needed.
    - If the label is incorrect, but the bound box is accurate, users can correct the label.
    - If both the label and bound box are incorrect, an integrated **Labelimg** software enables users to create precise bounding boxes and assign labels (person GUID or email).
- **Handling Partial Correctness:** In scenarios involving multiple faces within a single frame, where some bounding boxes are accurately tagged, but others are not, the UI allows the same flexibility for refining annotations.
- **API-Enabled Feedback:** The UI utilizes the provided APIs to transmit the necessary information. Both the image and its associated metadata are earmarked as candidates for creating a new model.
- **Automated Deployment:** This new model is automatically generated and deployed in the Agent's environment, ensuring continuous enhancements to the recognition system's accuracy and efficacy.

Fifth Version of the Pipeline (Agent Architecture Tuning)

Challenge:

In this iteration, we encountered a significant challenge regarding the MTCNN module, specifically its performance on a non-GPU system. It took an average of 0.9 to 1.1 seconds per image for processing. With 50 frames-per-second input from two cameras, we confronted issues in maintaining a steady operational state. After several hours of operation, the system often crashed, losing crucial detection data.

Solution:

To overcome this hurdle and ensure uninterrupted system operation, we devised an innovative solution that involved the implementation of multiple MTCNN queues. The process unfolded as follows:

- Message Distribution: The Camera_opencv Module was tasked with posting messages to the various MTCNN queues.
- Load Balancing: To manage the distribution of data across these queues efficiently, a compact load balancer module was introduced. This module dynamically assessed the workload of each queue and directed incoming data to the least busy one. Given that frames containing multiple individuals require more processing time compared to single-person frames, the load balancer ensured equitable distribution.
- Temporal Order Preservation: Recognizing the importance of maintaining chronological sequence, especially with multiple MTCNN queues in play, the Camera_opencv Module embedded each frame with metadata detailing its arrival time.
- Streamlined Data Flow: After processing within each MTCNN queue, the resulting objects were promptly dispatched to the Processing Queues, streamlining the overall data flow.
- Metadata-Driven Processing: The Processing Queues incorporated the metadata timestamp during the stages of embedding and detection. This timestamp-based approach ensured that the detection results were sent to Kafka in the correct chronological order, enabling coherent and accurate data representation.

Result:

Through these enhancements, we successfully addressed the challenge posed by the MTCNN processing time, mitigated operational disruptions, and enhanced our system's resilience and efficiency.

Blur and Deblur Classification

The images in motion could turn blurry, yet with some images, despite being blurred, we can detect the person if we use the deblur technique.

To facilitate blur classification, we organized it into three distinct categories:

Highly Blurred Image:
- Action: Reject the image for classification.
- Explanation: Images exhibiting severe blurriness are unsuitable for classification.

Moderately Blurred Image:
- Action: Consider it for deblurring using techniques.
- Explanation: These images, while blurred, hold potential for improvement through deblurring techniques. They are candidates for face classification.

Sharp Image:
- Action: Utilize the image as-is for face classification.
- Explanation: Images in this category are clear and require no further processing for accurate classification.

Our efforts in blur classification included the use of various techniques:

- **Laplacian Operator:** We employed this method to identify edges and calculate the variance of the Laplacian operation. A lower variance value signifies increased image blurriness.
- **FFT (Fast Fourier Transform):** Our use of FFT allowed us to analyze frequency distributions within the image. A higher frequency count indicates improved image quality, while fewer high frequencies suggest image blurring.

Although we explored several other methods, we encountered challenges in classification accuracy. The primary issue stemmed from using the entire image as input. Even when only the face displayed blurriness, the background and floor components remained sharp, causing algorithms to assign a high score and treat the image as non-blurred. To address this, we implemented the following improvement measures:

- **Isolating the Extracted Face:** We modified our approach to take only the extracted face from MTCNN as input, thereby improving classification accuracy.
- **SVD (Single Value Decomposition) for Blur Classification:** By employing SVD factorization, we determined the degree of blur and classified images as blurred or not based on a defined threshold. This method yielded an impressive 80% accuracy rate in classification.

While pursuing deblurring techniques, we have explored various methods, some of which are standard and readily available online. These techniques include:

1. **Weiner Filter:** This method minimises the mean square error between the image and its estimate. To enhance the image quality, we also applied contrast stretching.
2. **Blind Deconvolution:** Blind deconvolution is a viable option when no prior information is available about the distortion present in the image
3. **Richardson-Lucy:** We employed an iterative process using the Richardson-Lucy algorithm to recover the original image from its blurred counterpart.
4. **Nafnet:** Nafnet represents a state-of-the-art image restoration model that operates without nonlinear activation functions. While it has shown promising results, ongoing work is dedicated to further enhancing its performance.

Our research endeavouring in the fields of blur classification and deblurring techniques remains ongoing. We are actively exploring the development of Convolutional Neural Network (CNN) models for this purpose. From our learnings, we anticipate providing more comprehensive findings in the upcoming version of our work.

Conclusion

Accuracy

The journey undertaken in developing and exploring the Employee Recognition System powered by Flex83's AI ML capabilities concluded with a 90% accuracy in Face Detection.

To increase the accuracy beyond 90%, we are presently working on several key strategies and techniques:

1. **Feedback Network Improvements:** We are in the process of implementing advanced feedback networks that will enable the model to learn from its mistakes and successes. By analyzing these detections and incorporating regular feedback, we aim to fine-tune our AI algorithms and increase the overall accuracy.
2. **Blur Deblur Techniques:** In the real world, environmental factors can introduce blurriness to images, making face detection challenging. To address this issue, we have introduced blur deblurring techniques. These methods will help our system improve its ability to detect faces even in images with varying degrees of blur, ultimately boosting the accuracy of the recognition system.
3. **Image Denoising:** Image noise can also pose a significant challenge to face detection accuracy. To tackle this problem, we are already working towards image denoising techniques that will allow the AI system to better distinguish facial features from noisy backgrounds, enhancing the overall precision of the detection algorithms.

Safety Management

By harnessing the prowess of artificial intelligence and machine learning, organizations can seamlessly transcend the limitations of traditional recognition methods. The employee recognition technology also plays a crucial role in securing a safe working environment with the below real-life examples:

1. **Employee Presence Detection in Hazard Zones:** In the context of workplace safety, this technology enables precise employee presence detection within hazardous zones. The application can accurately identify and track employees as they enter or exit these areas using advanced AI algorithms and sensors. In the event of unauthorized access or the presence of an employee without the required safety training or equipment, real-time alerts are triggered, allowing for swift intervention and accident prevention.
2. **Employee Safety Gear Identification:** Another critical application of employee recognition technology is identifying safety gear. Through AI and Machine Learning, we are developing a model to visually inspect employees to ensure they wear the necessary safety gear, such as helmets, goggles, or vests, in designated work areas of an Industrial OEM. Any instances of non-compliance will be immediately flagged, notifying supervisors, thus reinforcing a culture of safety and reducing the risk of workplace accidents.
3. **Intruder Detection and False Positive Feedback Loop:** The Flex83 Face Recognition App incorporates an invaluable feature: the ability to identify unknown individuals who may intrude on the workplace premises. This feature is enhanced by implementing a secondary feedback system designed to minimize false positives. When the AI/ML system detects a potential intruder, it exercises caution and employs multiple checks to verify the identification. In cases where a false positive is determined, this information is incorporated into a feedback loop. This iterative process fine-tunes the machine learning model, reducing the occurrence of false alarms and ensuring that genuine security threats are promptly addressed, thus bolstering workplace safety and security.

Cameras & Sensors

For camera and sensor placement and technical settings, achieving an ideal setup is essential to ensure accurate results in face detection. Our journey towards this ideal configuration involved experimenting with several key factors:

1. **Physical Alignment of Cameras:** We recognized that the physical positioning of cameras significantly impacts face detection accuracy. To optimize this, we conducted rigorous experiments to determine the ideal camera placement. This involved considering factors such as camera height, angle, and coverage area. By fine-tuning these physical aspects, we ensured that the cameras captured facial images consistently and effectively.
2. **Sensors with Adjustable Aperture, Shutter, and Focus:** The flexibility of camera sensors is crucial in adapting to different lighting conditions and environmental factors. We invested in camera sensors equipped with adjustable aperture, shutter speed, and focus settings. This allowed us to customize the camera's behavior to suit specific scenarios, such as low-light conditions or rapidly changing focal points. These adjustments significantly contributed to the system's ability to capture clear and accurate facial images.
3. **Environmental Optimization:** In addition to camera settings, we also focused on optimizing the environment in which the cameras were placed. This involved controlling factors like lighting, background noise, and temperature. For instance, we installed proper lighting fixtures to ensure consistent illumination for facial recognition, reducing shadows and enhancing image quality.
4. **Sensor Redundancy:** We incorporated sensor redundancy in our system to enhance reliability. This means strategically placing multiple cameras and sensors to minimize blind spots and increase the system's overall coverage. Redundancy ensures that even if one sensor fails or is obstructed, the system can rely on others to maintain accurate face detection.

5. **Integration with AI Algorithms:** Beyond hardware adjustments, we integrated advanced AI algorithms that could adapt to changing conditions and self-calibrate based on the data collected. These algorithms continuously learn from the environment and user feedback, making real-time adjustments to optimize face detection accuracy.

The enhanced accuracy and efficiency achieved through MTCNN-based face detection and FaceNet embeddings form the cornerstone of this project. The ability to seamlessly adapt to dynamic scenarios, alongside the iterative refinement driven by user feedback, ensures a solution that meets and surpasses modern workforce and security management demands.

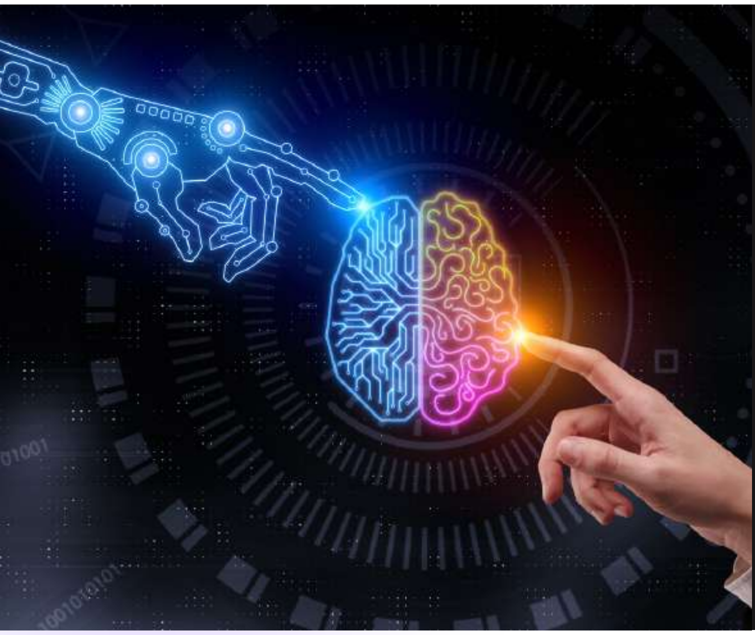
Flex83's Face Recognition System stands as a testament to the harmonious blend of cutting-edge technology and visionary innovation. By embracing this solution, organizations can unlock the potential to usher in a new era of precision, security, and efficiency in employee recognition—a transformation that will inevitably permeate every facet of the organizational ecosystem.

Author Contributions

A. Chandra led the development of the technological aspects, encompassing challenges, solutions, and strategic approaches. He authored the Results & Discussions section and carefully verified technical accuracy throughout the paper. ***A. Rana*** edited and formatted the paper, incorporating valuable sections from research materials enhancing the readability and structure.

References

Robust face recognition based on multi-task convolutional neural network
Face Recognition with FaceNet and MTCNN
Research on MTCNN Face Recognition System
Recent Advances in Deep Learning Techniques for Face Recognition
FaceNet: A Unified Embedding for Face Recognition and Clustering (Google Inc Report)
FACIAL ATTENDANCE SYSTEM USING MTCNN AND FEATURE MAPPING
Face Identification Using Data Augmentation Based on the Combination of DCGANs and Basic Manipulations
Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks



By embracing the boundless potential of digital interconnectivity, we can experience unprecedented growth and efficiency for the industrial OEMs.

Lee House, CEO at IoT83

HAVE QUESTIONS TO ASK?

CONTACT US FOR INQUIRIES

Phone

+1 650 996 6650

Email

contact-sales@IoT83.com

Website

www.IoT83.com

IoT83

Connect Automate Optimize