National College of Ireland

Connected Home Security System – An AI enabled surveillance system to detecting Unwanted Intrusions using Cloud Based system.

MSc Research Project
Cloud Computing

VIJAYAKUMAR KANNIAH
Student ID: x21188955

School of Computing
National College of Ireland

Supervisor: Sean Heeney

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	VIJAYAKUMAR KANNIAH
Student ID:	x21188955
Programme:	Cloud Computing **Year:** 2022
Module:	MSC Research Project
Supervisor:	Sean Heeney
Submission Due Date:	18/09/2023
Project Title:	Connected Home Security System – An AI enabled surveillance system to detecting Unwanted Intrusions using Cloud Based system.
Word Count:	**6874 Page Count: 20**

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: VIJAYAKUMAR KANNIAH

Date: 18/09/2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Connected Home Security System – An AI enabled surveillance system to detecting Unwanted Intrusions using Cloud Based system.

VIJAYAKUMAR KANNIAH
x21188955

Abstract

This paper presents a comprehensive facial recognition system that explores various deep learning and transfer learning models. Among them, the final model selection is based on three distinct approaches: Haar cascade, AdaBoost, and HOG+SVM. The selection criteria involve evaluating space consumption, memory utilization, and detection speed. This systematic model testing ensures an efficient and practical solution for real-time face detection and recognition. A key novelty of this system lies in its ability to not only identify known individuals but also detect and handle probable unknown people effectively. This feature enhances the system's security capabilities, protecting the house from potential attacks or unauthorized access attempts. By combining multiple detection methods, the system achieves robust and accurate identification of individuals, bolstering home security and surveillance. The research also leverages the advantages of deep learning and transfer learning techniques to optimize the models' performance. The training modules deployed on the AWS platform enable the system to continuously learn and adapt, generating the best model for each user over time. Overall, this paper showcases the innovative use of diverse approaches in facial recognition technology, considering both efficiency and security aspects. The proposed system not only excels in identifying known individuals but also effectively addresses the challenge of handling probable unknown people, making it a promising solution for improving home security and surveillance in a connected and intelligent environment.

Keywords: Residence Monitoring Solutions, Facial Identification, Data Storage Infrastructure, Advanced Neural Networks, Internet of Things (IoT), Cloud-Based Processing

1 Introduction

The study of facial recognition is vital because of the many real-world benefits it may provide. Applications such as access control, bank card identification, surveillance, and security monitoring are just a few examples. Galton is widely regarded as the pioneer who proposed using standard curves and outlier scores to classify people's facial features. The concept of remembering people by their faces may have originated with Galton. Modern face recognition algorithms have made it possible to identify people in a wide range of contexts, including those encountered in the real world. To protect their family from harm, homeowners may rely on the

Connected Home Security System, a state-of-the-art surveillance system with AI capabilities that makes use of cloud computing. This state-of-the-art surveillance system uses a combination of cloud computing, machine learning, and vision to maintain watch over the premises around the clock and spot any number of potential intrusions as they happen. Homeowners may receive alerts through email or a smartphone app, keeping them abreast of any potential threats in real time.

The system's effectiveness stems from its deployment of sensors, cameras, and other forms of intelligent hardware to study inhabitants' habits and adapt accordingly. This enables it to recognise out-of-the-ordinary occurrences, which may represent threats. This invention has the potential to revolutionize the home security sector by giving users cutting-edge defenses by fusing AI and cloud computing. Surveillance systems equipped with AI and face recognition are a powerful tool for bolstering security in a wide range of settings. These technologies' ability to detect people and potential threats in real time makes them invaluable for access control and monitoring, as well as for preserving the safety of homeowners and their families. The future of face recognition holds the promise of even larger potential for boosting people's capacity to communicate with one another and their sense of safety as research and technology continue to improve.

A. Problem Statement

In recent years, more sophisticated home security systems have been created as a result of the development of sensors, cameras, and smart locks that can identify and alert homeowners to suspicious activity. A capability to distinguish people from face images is crucial for security and surveillance applications such as visitor monitoring, intruder detection, and access management, as stated by **Chitnis et al. (2016)**. It's possible, though, that these algorithms won't always pick up on strangers. Taking use of the potential of interconnected home security systems, a solution must be developed that can identify unfamiliar guests.

B. Motivation

An approach to possibly identifying unknown visitors at a location secured by interconnected home security systems could utilize capabilities similar to Truecaller (Aprinia, 2022), an application that identifies unknown callers by comparing phone numbers against a crowdsourced contact database. The proposed system would leverage the photo capture abilities of networked home security cameras and match those images against a database to identify unrecognized individuals, much like Truecaller matches caller ID information to user profiles. To enable the image matching capabilities, the system could potentially utilize a cloud-based service like Amazon Web Services (AWS) to host the database and run the comparison algorithms. One way this may be achieved is by making use of cloud-based service providers like (AWS)Amazon Web Services.

C. Research Outcome

The goal of this project is to create a surveillance system using facial recognition that is affordable and flexible enough to be used in a variety of settings, including the home and the office. Using a cloud service, the system can discriminate between known and unknown individuals and predict the possibility that a known individual is present in a different linked house. The process involves locating and tracking individuals as they walk across the camera's vision in real time. Ijaradar and Xu's (2022) study found that while analysing an image, it is more important to focus on what is really visible within the frame. Users in a network can connect to one another and share data regardless of their physical location thanks to cloud computing. The system is a powerful tool that can be used in a number of settings to increase security and surveillance due to its low cost, straightforward installation procedure, and high degree of dependability. The study's findings show that the image analysis approach is effective, guaranteeing the accuracy and dependability of future face recognition jobs.

D. Research Questions

The primary objective that is motivating this inquiry is to develop a system that is analogous to Truecaller and is capable of recognising unknown individuals in an environment that already makes use of cloud-hosted and internet-connected home security equipment. In order to achieve this goal, we shall conduct research on the following questions:

RQ1: How can we build a system that can distinguish between unknown people and the likelihood that a known person would show up at someone's door?

RQ2: What methods and technologies are currently available for identifying persons who haven't shown identification at a place, especially in the context of linked home security systems deployed with AWS?

RQ3: Is it feasible to do a thorough analysis to ascertain the cost, the amount of processing power necessary, and the amount of storage space available? What type of computational complexity is required in the design of a system like this?

RQ4: When deciding on a edge device and cloud platform for the surveillance system, it is important to know which machine learning (ML) models have the highest performance in terms of real-time computation and memory utilisation.

2 Related Work

In order to detect potential threats, a connected surveillance system with AI system for connected home security analyses data collected from a wide range of sensors and devices, including cameras and motion detectors. The system may alert appropriate parties and take protective measures, such as sounding alarms and closing doors, in the case of an intrusion. A cloud-based architecture is utilised by the software, allowing for global accessibility and control, as well as scalability and flexibility. Homeowners may feel secure and relaxed in their own homes thanks to this cutting-edge security system's enhanced level of protection. The powerful artificial intelligence (AI) capabilities and smooth integration of smart devices allow for increased home security and continuous monitoring. This makes it a highly effective line of protection against threats and intrusions.

A. Deep Learning Based Video-Surveillance System

Recent research by Chen et al. (2021) has explored utilizing drones equipped with computer vision and deep learning for detecting and tracking people in aerial footage. Their proposed system used a drone with an onboard camera, a lightweight YOLOv3-based object detector, and the DeepSORT algorithm to follow multiple people simultaneously. The system demonstrated strong performance for multi-person tracking in real-time aerial videos. However, challenges remain in deploying such systems commercially due to factors like weather, lighting conditions, and flight time limitations. While showing promise, additional work is needed to improve the robustness and applicability of AI-powered drone tracking systems for real-world uses like security and surveillance.

Sung et al. (2021) presented a novel intelligent video surveillance system that makes use of deep learning to identify illegal conduct in real time. The technology proactively discovers illegal activity and overcomes the passive character of traditional surveillance systems by publishing videos and notice messages to the web. In their study on object tracking algorithms for smart city surveillance, Sharma et al. (2022) highlighted the considerable benefits gained by deep learning technologies in drone-based surveillance systems. The study was conducted by Sharma and colleagues. In their study on intelligent CCTV systems, Pandiaraja et al. (2023) emphasised how important it is to be able to identify anomalous events and analyse human behaviour. They suggested using machine learning and machine vision technologies in order to automate the identification of anomalies and minimise the amount of manual labour required, with the Grassmann approach showing to be reliable for face detection.

Abdelhafid Berroukham (2023) explored a variety of deep learning-based algorithms for identifying abnormalities in videos, and he categorised these methods according to the approach that they use. Deep learning demonstrated some encouraging findings in this area, and the publication supplied future researchers with assessment measures and datasets that are freely available to the public. For the purpose of human re-identification in high-tech video surveillance, Muazzam Maqsood (2023) suggested a deep learning architecture that was based on the adaptive feature refinement of an image. The model's performance was encouraging to look at from the standpoint of recognition accuracy. For the purpose of recognising human actions, Inzamam Mashood Nasir (2022) created a hybrid recognition approach that he named HAREDNet. The recognition accuracy of the system was far greater than that of any of the previously proposed methods, and this was accomplished across several datasets. Huu-Thanh Duong (2023) offered a comprehensive investigation of the most recent developments in the field of deep learning-based video anomaly detection techniques. A number of different tactics were categorised, and issues in the field of video surveillance were explored, along with potential solutions and prospective paths for future study.

B. Control system for intelligent smart homes based on the Internet of Things

In the year 2021 Gusti Made Ngurah Desnanjaya researched cited above recommends (Gusti Made Ngurah Desnanjaya, 2021) using Raspberry Pi and Telegram to set up a home security monitoring system. The system employs a passive infrared (PIR) sensor to spot motion, a Raspberry Pi camera to snap photos whenever the PIR sensor spots motion, and temperature and gas sensors to keep tabs on the ambiance. Telegram is used for communication between the Raspberry Pi and its users. Telegram's Mproto protocol enables devices to connect, allowing for the transfer of data and the receipt of alerts. The system recognises genuine users

and rejects invalid ones, and it responds to requests by providing the requested information or alerting the user. The Raspberry Pi-powered home security system provides continuous surveillance, detects and identifies potential intruders, and notifies residents of any significant changes in the local environment.

(Davin Bagas Adriano,2018) stated using Arduino Nano and NodeMCU ESP8266, the objective of this inquiry is to design and create an integrated home security and monitoring system. Door access is controlled by radio frequency identification readers, motion is detected with Passive Infrared (PIR) sensors, the environment is monitored with humidity and temperature sensors, rain and fire sensors, and lighting levels are tracked with light-dependent resistors. Light bulbs and solenoid valves are two examples of actuator components. The technology allows users to watch their houses from afar and control various equipment, such lights and solenoid valves, from anywhere with an internet connection and a smartphone running the appropriate software. The system's technique, built on the Internet of Things, allows for real-time monitoring and management, boosting home security and giving homeowners peace of mind. This investigation combines a Raspberry Pi 3 and an Arduino connected through USB to develop a biometrically-enabled home security system (Nico Surantha, 2018). The passive infrared (PIR) sensor is standard on the Arduino, whereas the camera is standard on the Raspberry Pi. The system uses the Histogram of Gradient (HoG) and Support Vector Machine (SVM) algorithms for object identification and human detection, respectively. The camera begins monitoring the area for possible dangers when the passive infrared (PIR) sensor detects motion, and the system employs a support vector machine (SVM) algorithm to make identifications. The security system recognises authorised inhabitants and alerts the owner instantly if an unauthorised individual is detected on the premises. With an average detection time of around two seconds and an accuracy of 89%, the proposed system is a viable alternative for bolstering the current state of home security.

(Masud, et al., 2020) With an emphasis on healthcare settings, This research aims to create a tree-based deep learning model for unrestricted automatic face recognition. Each piece of data that is input into the model is broken down into smaller pieces, and then a tree is built for each of those pieces using residual functions as the branching mechanism. The model achieves an accuracy of 98.65% on the FEI database, 99.19% on the ORL database, and 95.84% on the LFW database. Given its computational efficiency, the tree-based deep learning model is a good fit for usage in real-time applications.

Utilizing cloud technology in conjunction with an Android application, the researchers of this study propose an intelligent system for home automation (Taiwo et al., 2021). The system is designed with modularity in mind, encompassing various components that facilitate the control and monitoring of environmental conditions, electricity usage, and home security via motion-detecting cameras. A noteworthy feature of this system is its integration of machine learning, which can distinguish between regular residents and potential intruders, thereby minimizing unnecessary alarm triggers. The classification of images is accomplished using the support vector machine technique, ensuring that alerts are only sent to the user when warranted. This study illustrates the potential of machine learning to enhance home security and offers valuable insights for the development of a cloud-based intelligent home automation system. Furthermore, the paper elaborates on the advantages of leveraging machine learning to bolster residential safety. The practical implementation of the system involves the assembly of prototypes using relays, sensors, and ESP8266 and ESP32-CAM boards. Another aspect of this

paper introduces the concept of a Behavioral Modeling Intrusion Detection System (BMIDS) aimed at identifying security vulnerabilities in IoT devices, particularly within smart homes (Emmanuel et al., 2020). The proposed IDS employs an adaptive approach based on an Immunity-Inspired Algorithm, coupled with an Extreme Learning Machine (ELM) incorporating an Artificial Immune System (AIS-ELM) to effectively detect unusual patterns in IoT gateway traffic. To promptly inform homeowners of suspicious activities, the IDS may be integrated with a push notification system. The outcomes underscore the importance of smart home devices offering robust security measures that are energy-efficient.

A group of researchers (Nargelekar et al., 2020) study's overarching objective is to provide real-time intrusion warnings to houses using a monitoring and detecting system that is both efficient and cost-effective. The device is able to detect intruders by comparing individual frames from a live video feed. This system provides real-time identification and reporting of possible threats, addressing a major deficiency of previous security measures. The focus is on providing homeowners with real-time notifications in the event that unauthorised persons obtain access to their property. An Enhanced Security System, the focus of this research, was developed to deter criminal activity in restricted areas. The system has laser security protection, tear gas cannons and a GSM module to notify authorities and launch tear gas if the security perimeter is violated. The system requires a MasterCard to operate and has an RFID-enabled administrative control panel. The technology is meant to provide maximum security in high- crime areas, such as those vulnerable to burglary and robbery (Anjum et al., 2022).

(Kenli et al., 2019) The study's goal is to propose a Distributed Intelligent Video Surveillance (DIVS) system that uses deep learning techniques and can function in an edge computing setting. To improve accuracy while decreasing network connection cost, the system employs a distributed deep learning training model and a multi-layer edge computing architecture. The DIVS system under discussion demonstrates the potential for improved precision in real-time surveillance applications. This article (Taiwo et al. (2022)) presents a sophisticated automation system for the house, one that can control appliances, keep tabs on the weather, and track who's coming and going. Deep learning is used in the process of identifying and categorising various motions. A camera, motion detector, and temperature/humidity sensor are just some of the sensors integrated into the system to boost safety and comfort. The proposed system allows for continuous monitoring and control via a cloud-based platform, with access granted via an Android app on a portable device.

C. Cloud based intruder detection system

With their wide open internet connections, smart home devices are easy targets for hackers and malware, and Nada Ratkovic (2022) looked into these dangers. Blockchain technology was mentioned in the article as a possible method for bolstering the security of IoT and smart home devices. Decentralised DNS entries, encrypted data storage, and protection from Distributed Denial of Service (DDoS) attacks are all within the capabilities of blockchain devices. Smart home appliances can keep personal information more secure in a globally interconnected society by using blockchain technology. D. Asir Antony Gnana Singh (2018) elaborated on the challenges of cloud security in the context of cloud computing. In order to enhance the precision of intrusion detection systems (IDS) used for cloud security, the article provided a cuckoo optimization-based approach for processing network traffic data. The technique

improved cloud security by making it simpler to spot potential dangers, making clouds more robust against attacks.

To protect cloud infrastructure from attacks, Wisam Elmasry (2021) created a unified Cloud-based Intrusion Detection System (CIDS). The proposed CIDS relied on five main modules: network monitoring, traffic flow capture, feature extraction, flow analysis, and intrusion detection. We deployed deep learning models inside of a refined bagging ensemble system to predict intrusions with high precision. The study's results showed that the CIDS effectively fends against cloud-based attacks. To boost the efficacy of facial recognition, Pengfei Hu (2018) developed a cloud-based face detection and resolution system. The model utilised cloud computing to improve its computation and storage capacity, and face IDs to better match and resolve facial images. The experimental results indicated that the proposed system could provide reliable face identification and resolution services.

In an effort to make machine learning more transparent, Martin Knoche (2023) set out to explain how face recognition algorithms arrive at their conclusions. In an effort to make better face recognition predictions, researchers looked at using confidence scores, explanation maps, and visualisation techniques. The study stressed how crucial it is to offer explanations for facial recognition systems and proposed ways for reaching this purpose.

Hexiao Yin (2022) looked at the possibility of using cloud computing to analyse public safety video footage for inspection. Researchers built a cloud-based public security video image detection system using Hadoop, CP-ABE encryption, and other server-side and backend technologies. The proposed approach demonstrated greater levels of accuracy and efficiency than more conventional detection techniques, suggesting it may be a beneficial strategy for the protection of the general public. Facial recognition systems, as underlined by Siqi Deng (2023), require a high degree of recognizability. A measure for measuring how easily a face can be recognised was developed via the study and incorporated into the final product. Error rates in both single-image and set-based face recognition were reduced because to the article's focus on improving the system's recognizability. The study shed light on novel avenues for improving facial recognition algorithms and reducing recognition errors.

F. Summary

Papers demonstrate that deep learning enhances surveillance through accurate intrusion detection in smart cities and the ability for real-time proactive criminal detection, as well as through the recognition and tracking of people utilising drone-based overhead imagery. Deep learning is what makes these advantages attainable. Improved object monitoring, anomaly detection, and human behaviour analysis are all achievable with the help of intelligent CCTV systems and adaptive feature refinement for re-identification. Internet of Things devices that use deep learning, such as Raspberry Pi and Arduino, are useful for home security since they can detect intruders, keep tabs on their surroundings, and provide alerts instantly. Methods such as cloud-based intrusion detection, blockchain-protected Internet of Things devices, and contextual face recognition algorithms all contribute to a higher level of security. These developments show promise for more advanced security systems that can safeguard buildings and citizens at a reasonable cost.

G. Research Contribution

It is feasible to construct sophisticated systems for object detection and facial recognition that can be hosted in the cloud by utilising Dlib and AWS deployment. These systems may be deployed everywhere there is an internet connection. Dlib is a robust library that can be used for both face and object identification, and Amazon Web Services (AWS) provides a wide variety of cloud services that can be utilised to install and maintain the system. It is feasible to develop a system that is scalable, dependable, and secure by integrating these technologies in such a way that it can be accessible from anywhere in the world that has an internet connection. Access control, client involvement, and security surveillance are some of the potential uses for a system like this one. In addition, using Amazon Web Services (AWS) makes it possible to easily integrate AWS with tools and other services, such as those used for data storage and analysis. Which will have folders for each guest and each user, which can be used to first detect a person and then identify the person after they have been found. Because the system may be connected for several users through the cloud, it may be simple for an unknown individual who is visiting to provide their identification.

3. Research Method & Specification

1. Research Method

To implement the solution, we can leverage cloud infrastructure and services:
First, Amazon EC2 free tier instances can be provisioned to provide the necessary compute resources without incurring additional costs. Amazon's image builder service simplifies selecting an appropriate machine learning AMI with Ubuntu 18.0 preconfigured.

Second, model training workflows can utilize FasTAPI, an open source Python library that facilitates quickly building deep learning interfaces. It provides high-level components to speed up creating standard UI elements as well as lower-level components that can be combined to develop new interfaces. FasTAPI is architected for flexibility, usability and performance, using Python and the Swagger framework to enable rapid prototyping.

Figure1 : Methodology of swagger development.

2. Face Detection Algorithm: Research Tool

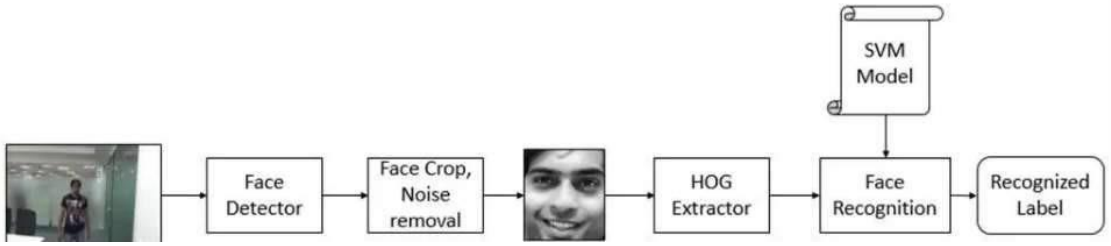
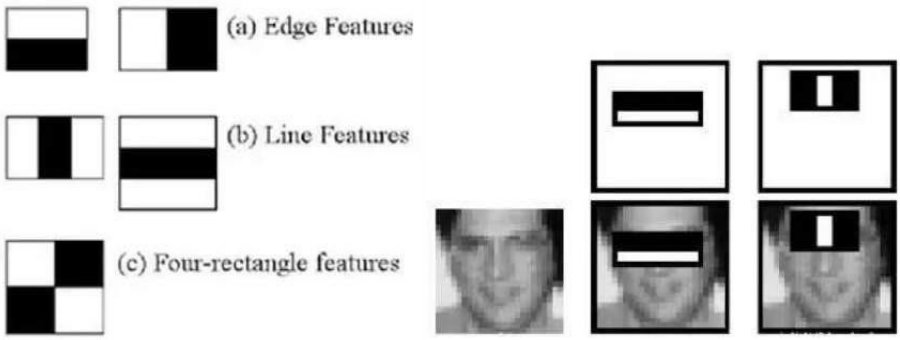


Figure 2: Implementation of Dlib on edge platform for Face Detection

Haar cascade-based classifier was chosen for inference on the edge device so that the inference is very lightweight and will require very less computation power so it can be deployed very easily on edge device. The identification of eyes, achieved using the Haar Cascade Classifier, is employed to construct a geometric face model, with the detection of the nose serving as a reaffirmation method in addition to the detection of the eyes. The HOG (Histogram of Oriented



Gradients) characteristics are taken from huge numbers of face photos and employed as part of the identification process, which is then applied to the photographs. Afterwards, the HOG characteristics are labelled together for a particular face/user, and a Support Vector Machine (SVM) model is trained to predict faces that are supplied into the system.

Figure 3: Illustrates the HAAR features denoted as a, b, c, and d, accompanied by corresponding images showcasing the areas where these features have been successfully matched.

A HOG descriptor is a type of feature descriptor commonly utilized in visual recognition tasks related to object identification, including applications in 3D object recognition. An illustrative example involves the utilization of helicopters to detect pedestrians walking alongside highways. The efficacy of a HOG descriptor relies on its ability to differentiate between distinct intensity gradient distributions or edge directions present within image elements.

In this approach, individual blocks of the image are assigned gradients, and these gradients are computed for each specific block in a sequential manner. These blocks can be envisioned as grid-arranged pixel compositions, wherein gradients are derived by altering the intensity change amplitudes and directions for each constituent pixel within the block. During the process of feature descriptor extraction, a set of feature descriptors is generated for each image, a set commonly referred to as a HOG. Each pixel within the image is associated with a gradient vector that functions as its descriptor. The techniques depicted in Figure 4 are employed to ascertain both the gradient magnitude and direction for every pixel.

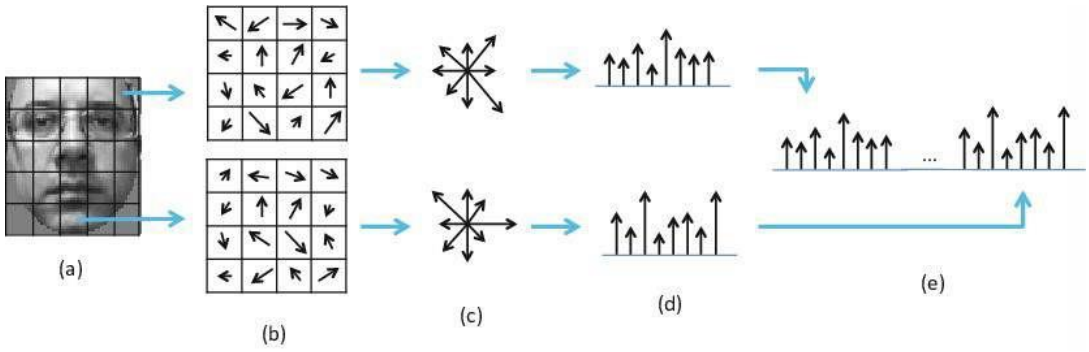


Figure 4: Gradient Components of the Face

In the provided illustration, the alterations in pixel intensity are denoted as Gx and Gy for the horizontal and vertical aspects of change, respectively. A facial image is exhibited within a window of dimensions 128 x 144 pixels, preserving the typical human face aspect ratio. The computation of descriptors occurs in pixel blocks measuring 8 x 8 pixels, which are then employed for analysis. To streamline this process, the descriptor values for each pixel within the 8 x 8 block are consolidated into nine bins, representing gradient angles and containing a value indicating the cumulative magnitude of pixels sharing the same angle. Furthermore, the histogram is standardized across an 8-by-8 block configuration, meaning that four such blocks are normalized in relation to each other, thus mitigating the impact of varying light conditions. This methodology contributes to reducing the potential accuracy loss caused by fluctuations in lighting. The Support Vector Machine (SVM) model is trained using numerous Histogram of Oriented Gradient (HOG) vectors derived from multiple facial images. These HOG vectors serve as the initial training data for the model, as outlined in Figure.

3. Evaluation

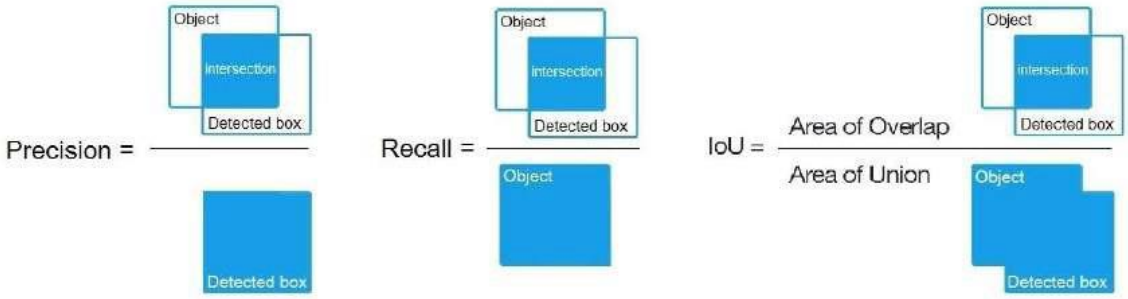


Figure 5: A face detection algorithm's evaluation

In object detection models, two important evaluation metrics are Intersection over Union (IOU) and recall. IOU, also called Jaccard index, quantifies the overlap between a predicted bounding box and the ground truth bounding box. It is calculated by taking the intersected area between the boxes divided by their union. A high IOU indicates the model accurately localized the object. An IOU threshold is used to determine true positive detections. Recall measures the model's ability to detect all positive object instances in the data by dividing the true positives by total ground truth objects. Higher recall means less missed detections. Precision evaluates detection quality by dividing true positives by all predicted boxes - higher precision means less false positives. Together, precision, recall and IOU provide insights into the accuracy, sensitivity, and localization capability of object detection models.

4 Implementation

The goal of this chapter is to demonstrate the end-to-end implementation workflow for this thesis. The code for this project is split into two parts, namely frontend part and the backend part. Frontend part consists of inference part which primarily is composed of package like streamlit which is forming core of UI and bare minimum deep learning framework made up of dlib for edge inference and tensorflow, keras for server based inference. Edge inference models can be run in edge devices and requires very less computation power. Whereas server based inference models requires some high end processors (if possible with GPUs) to give real time predictions. Below figure shows the complete architecture of the Intrusion detection system on the cloud as well as the edge devices. Api entry points refers to the web interfaces of cloud based system (backend) and edge device based system (frontend). Cognitive system for training used high speed and high compute aws EC2 instances whereas raspberry pi is a reflection of low compute edge devices which supports inference. Edge device should have a embedded camera module for capturing videos and monitoring premises continuously. Combined backend and frontend api entry points refers to as api gateway which will be accessible to users.

4.1 AWS account setup

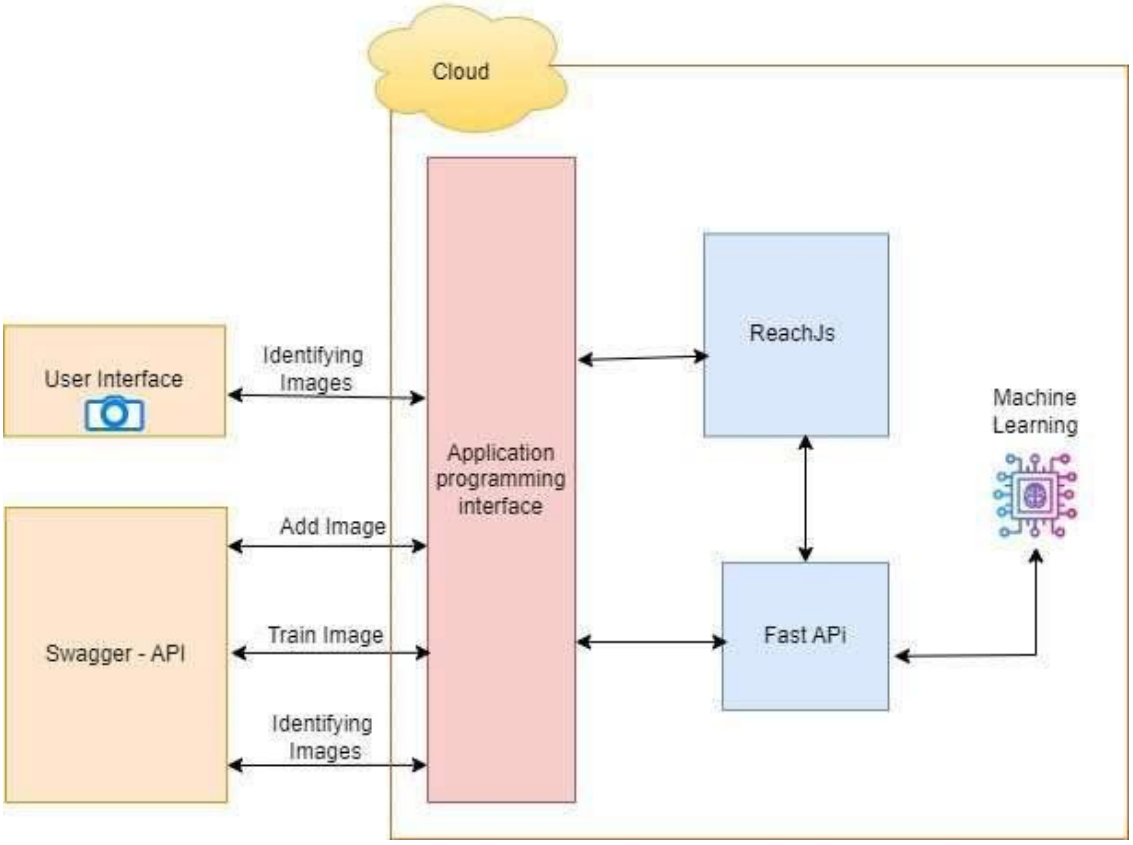


Figure 6: Architecture of the proposed system

For hosting the solution in cloud a AWS EC2 instances were found to be very good fit for this particular problem statement. AWS EC2 instances are called Amazon elastic compute cloud instances. In Amazon AWS various instances are available based on memory usage computational power and pricing. For this particular project since the more computational power instances are chargeable we have selected the free tier instances. After creating the Amazon account, we chose machine learning using the Amazon image builder service (AMI)
. Operating system selected was Ubuntu 18.0.

4.2 Training on Cloud

4.2.1 FastAPI

FastAPI is an open-source Python framework for building API applications. It aims to provide both high-level and low-level components for quickly creating performant and usable interfaces for deep learning systems. FastAPI enables rapidly building standard user interface elements for common machine learning tasks through high-level components. It also exposes lower-level abstractions that can be combined in novel ways to develop custom interfaces. This flexibility stems from a layered architecture that identifies common patterns across diverse UI/UX components. By leveraging Python's dynamic capabilities and the Swagger framework, FastAPI simplifies representing these abstractions.

4.3 Face Detection inference on the Edge

The haar cascade based classifier was chosen for making inference on the edge device so that it is very lightweight and will require very little computation power so it can be deployed very easily on edge device. The identification of eyes, achieved using the Haar Cascade Classifier, is employed to construct a geometric face model, with the detection of the nose serving as a reaffirmation method in addition to the detection of the eyes. The HOG (Histogram of Oriented Gradients) characteristics are taken from huge numbers of face photos and employed as part of the identification process, which is then applied to the photographs. Afterwards, the HOG characteristics are labelled together for a particular face/user, and a Support Vector Machine (SVM) model is trained to anticipate faces that are input into the system.

4.3.1 Applying HAAR Cascades

The framework for object detection known as the Haar classifier is built upon the research conducted by Paul Viola and Michael Jones, as presented in their paper titled "Rapid object identification using a boosted cascade of simple features." In their methodology, distinct classifiers are trained to recognize specific visual attributes, similar to those depicted in the provided image. However, the use of a solitary classifier does not yield a high level of accuracy. Instead, they employ a cascading arrangement where multiple such weak classifiers are combined. This is visually depicted in Figure 7. The ultimate robust classifier is formed by calculating a weighted sum of these weak classifiers. Through this approach, their framework attains an object detection accuracy of over 95%.

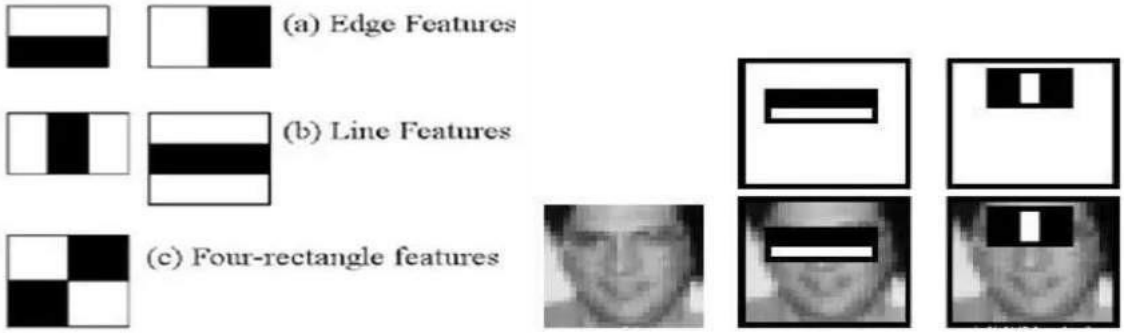


Figure 7: HAAR feature extractors are represented by a, b, c, and d. Furthermore, the pictures show the matched feature areas.

As previously described in the preceding section, Haar Feature-based Cascade Classifiers are used to identify the presence of a face. In most cases, the accuracy of face recognition is

Figure 8: Faces detected and cropped using Haar Features

strongly reliant on the quality and diversity of the sample photographs used in the analysis. The diversity of example photographs may be created by taking many images of the same face with different facial expressions on each of the images as depicted in Figure 8.

4.3.2 Applying Geometrical Face model

When creating a geometric face model, a pair of eyes is often regarded to be the first feature to be found inside the picture in order to construct a geometric face model. In an ideal situation,

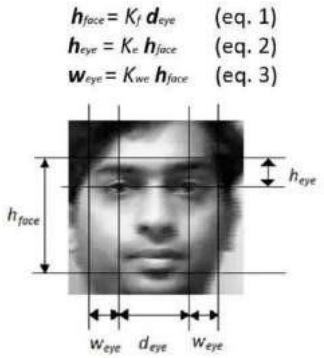


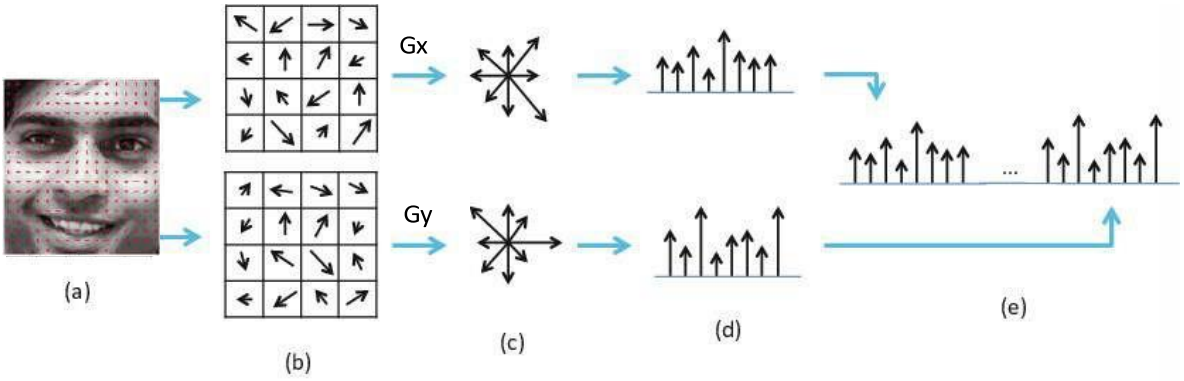
Figure 9: Geometric Face Model using eye

any of the attributes might be used as a starting point to create a face model, however, beginning with the position of the eyes gives a face model that is more accurate. A face model may be created in certain situations by determining where a person's nose is located. Although the eyes are normally viewed as the major beginning feature, in instances when the eyes are not found or are just partly obscured, the nose is often considered as a secondary starting feature.

4.3.3 Face Train finally using dlib and SVM - HOG features

A HOG descriptor is a type of feature descriptor commonly employed in visual recognition tasks related to object identification, including 3D object recognition. An illustrative application involves helicopters using HOG descriptors to detect pedestrians walking alongside highways. The effectiveness of a HOG descriptor relies on its capability to differentiate between elements within an image that exhibit distinct patterns of intensity gradients or edge orientations.

In this process, individual image blocks are allocated gradients. These gradients are computed for each specific block within the image. Conceptually, a block constitutes a segment of the image consisting of pixels organized in a grid. The gradients are generated by altering the magnitude and direction of intensity changes for each pixel encompassed within the block.



When a set of sample images containing a person's face is subjected to the feature descriptor extraction process, a series of feature descriptors, commonly referred to as HOGs, is produced for each image. These descriptors manifest as gradient vectors established for every pixel within the image.

The methods illustrated in above figure are utilized to ascertain both the magnitude and direction of gradients for each pixel:

In the provided illustration, the changes in pixel intensity's horizontal and vertical directions are denoted as Gx and Gy, respectively. This pertains to a face image displayed within a 128x144 pixel window, maintaining the typical aspect ratio of human faces. To compute descriptors, pixel blocks of 8x8 dimensions are employed. These descriptors are then utilized for computation.

For simplicity, the descriptor values for each pixel within the 8x8 block are categorized into nine bins, each representing a gradient angle. The value within each bin indicates the cumulative magnitude of pixels sharing the same angle. To mitigate lighting effects, the histogram is standardized across 8x8 blocks, involving the normalization of four adjacent blocks. This strategy helps minimize accuracy reduction caused by changes in lighting conditions.

The Support Vector Machine (SVM) model is trained using an extensive array of Histogram of Oriented Gradient (HOG) vectors derived from numerous faces. These vectors play a crucial role in the initial model training process, as depicted in below figure.

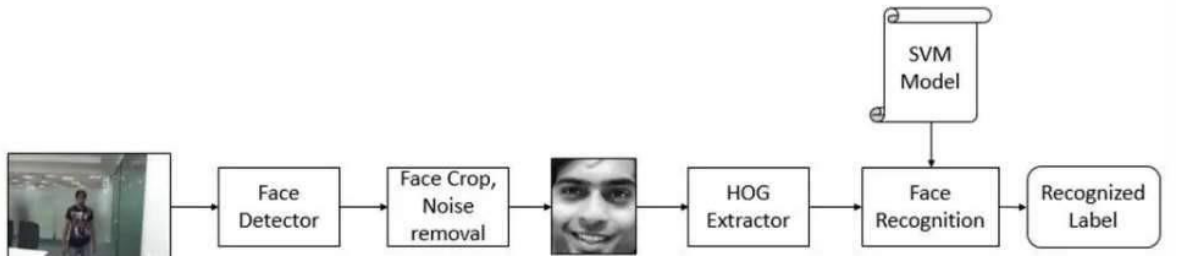
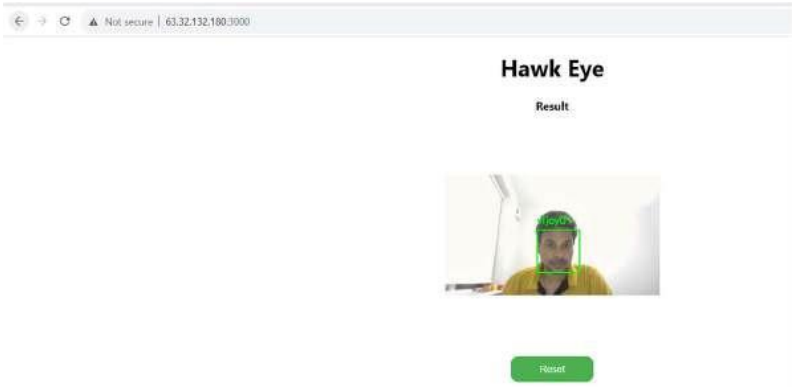


Figure: Represents a face detection model that was developed and deployed on an edge device using dlib.

5 Evaluation

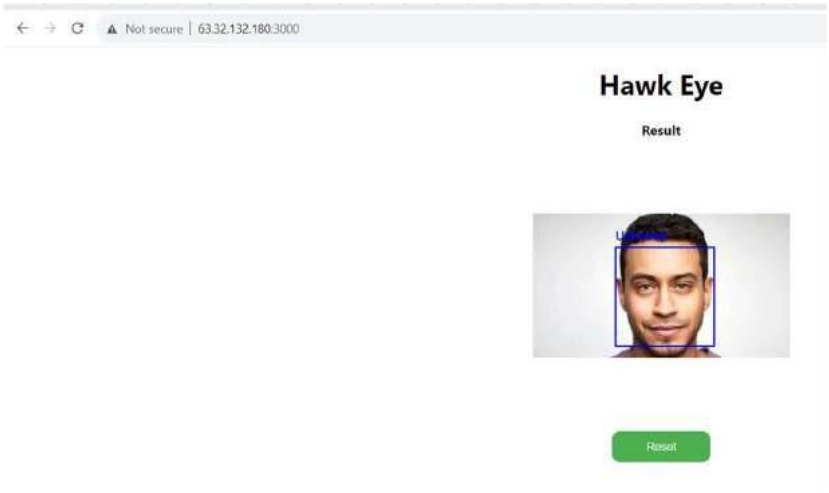
5.1 Experiment 1 - Comparing Detection models for edge deployment

Two small database of images were created, one from Unsplash and other from google, to test performance of models working on different varieties of images, small and big images. Following figure shows performance on above mentioned dataset, where dlib was performing considerably better than other models.



5.1 Experiment 2 – Testing Image is not in the system.

We took random face images from google to validate the trained models. As we can see in the below figure. The test results turned the person as UNKNOWN. Below figure shows the sample output for various samples.



5.1 Experiment 3 – Testing Image without face.

We tried to test the image without face and the trained models does not predict any faces. It just prompts the reset button options.

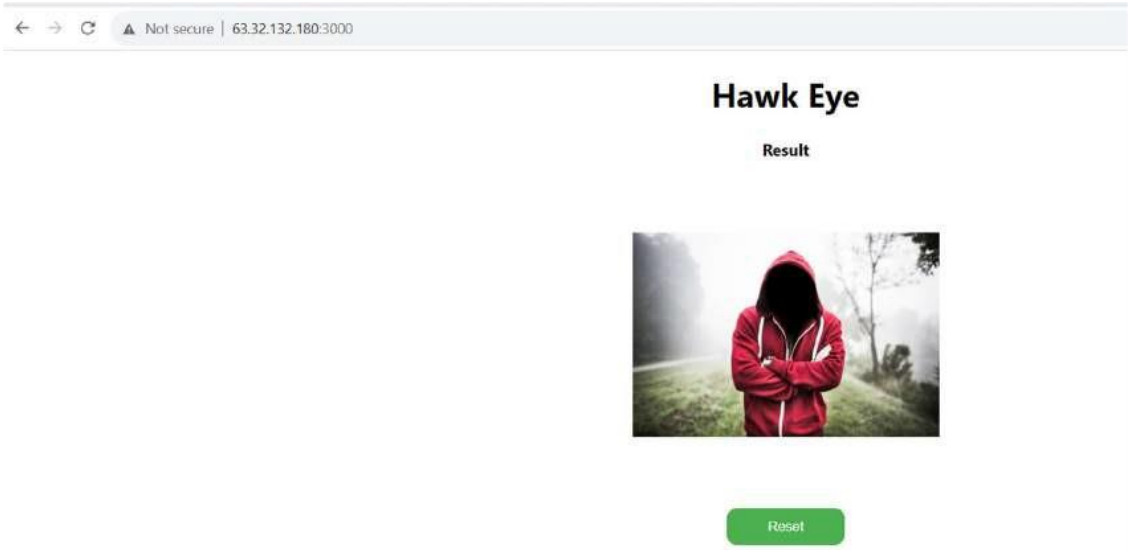


Figure: Testing model on images without face

5.2 Training Dlib based face recognition model.

Based on the images of seven famous personalities collected from Labelled Face Database, DLIB based face detection model was trained and accuracies were determined.

Figure: Represents snippets of faces from the utilized database

Following are the results from the training phase achieved in aws ec2 cloud using the backend api system created for this project

	precision	recall	f1-score	support
Ariel Sharon	0.88	0.54	0.67	13
Colin Powell	0.83	0.87	0.85	60
Donald Rumsfeld	0.89	0.63	0.74	27
George W Bush	0.83	0.98	0.90	146
Gerhard Schroeder	0.95	0.80	0.87	25
Hugo Chavez	1.00	0.53	0.70	15
Tony Blair	0.97	0.81	0.88	36
accuracy			0.86	322
macro avg	0.91	0.74	0.80	322
weighted avg	0.87	0.86	0.85	322

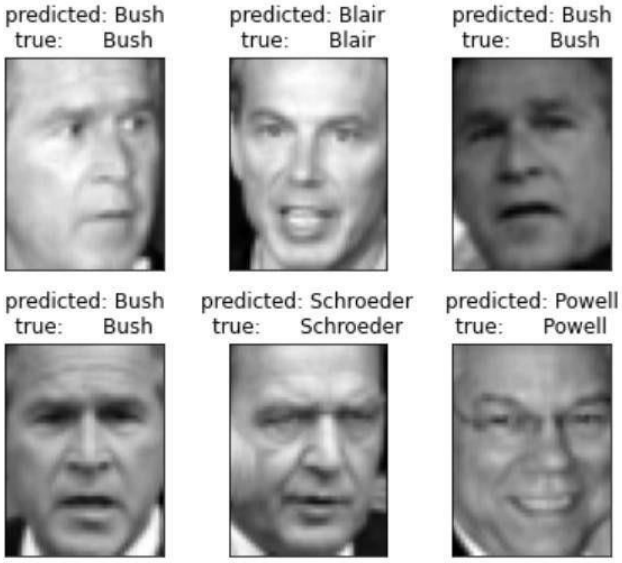


Figure: Predicted Results for the different images in the testing scenarios

6 Conclusion

This paper highlights the many methodologies and algorithms that have been developed for face identification. In every study, the face detection approach should be chosen depending on the unique needs of the application under consideration. None of the present options are the best answer for all applications in all scenarios. HOG-like features are digital picture properties that are used to identify objects in digital photographs. They were named by their perceptual similarity to Haar wavelets, and they were used in the development of the first real-time face detector. A HOG-like feature may be created by evaluating successive rectangular parts in the same location in a detection window, summarising the pixel intensities in each sector, and then computing the difference between these sums. The difference between the two is then used to categorise the different parts of a photograph. The major advantage of a HOG-like feature over most other features is its ability to be computed quickly. A successful face detection system must have both accuracy and speed in its face detection algorithm. In general, there is a trade- off between the two. Apart from the usual analysis of the algorithms present, which is issue with most of the industrial projects as people can develop algorithms but they stay in labs. A webapp based on cloud and edge device was created as a part of this thesis which can make this project practical and ready to be used by industry players. Intrusion detection systems like this can play a huge role in maintaining optimum security in any area where manpower is scarce in todays time like covid. Also in retrospect this system can also be used to identify high potential clients in any high profile banks and luxury shops which can alert the management if any high potential client walks inside the premises, so that the best salesman / marketing professional/ support staff can be sent to assist him. A ease of doing business definitely can result in increase in long term economy and good relationship with such clients.

6.1 Future Work

In this project emphasis was put on to analyze algorithms that can detect and recognize faces with great accuracy, but computing power of edge device was a limiting factor. As many state-

of-the-art modern face detection and recognition systems requires high precision GPUs. Recently Quantization of deep learning models have started gaining a lot of traction which reduces the compute precision of models from 32 bit to 8 bit which leads to massive increase in computation speed and great reduction of memory footprint of models by up to 4 times. Thus realm of Quantization and relevant techniques can help bring out the next generation of algorithm ready to deployed even in edge devices

References

Arun Francis G, Manikandan S,Sangeeth Thanga Dharsan .A, Shastiayappan R, Ravi K,T.Thiruneelakandan, Cloud Based Intruder Security System https://www.researchgate.net/profile/Arun-Francis/publication/357768507_Cloud_Based_Intruder_Security_System/links/61de9a1c4e4aff4a64360c52/Cloud-Based-Intruder-Security-System.pdf

Chang-Soo Sung , Joo Yeon Park, Design of an intelligent video surveillance system for crime prevention: applying deep learning technology https://link.springer.com/article/10.1007/s11042-021-10809-z

Emmanuel Dare Alalade , , "Intrusion Detection System in Smart Home Network Using Artificial Immune System and Extreme Learning Machine Hybrid Approach," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), New Orleans, LA, USA, 2020, pp. 1-2, doi: 10.1109/WF-IoT48130.2020.9221151 https://ieeexplore.ieee.org/document/9221151

Jianguo Chen; Kenli Li; Qingying Deng; Keqin Li; Philip S. Yu, Distributed Deep Learning Model for Intelligent Video Surveillance Systems with Edge Computing https://ieeexplore.ieee.org/abstract/document/8681645

Kamel Boudjita , Naeem Ramzan Human detection based on deep learning YOLO-v2 for real-time UAV applications https://www.tandfonline.com/doi/abs/10.1080/0952813X.2021.1907793

Mehedi Masud , Ghulam Muhammad , Hesham Alhumyani , Sultan S Alshamrani , Omar Cheikhrouhou , Saleh Ibrahim , M. Shamim Hossain , Deep learning-based intelligent face recognition in IoT-cloud environment https://www.sciencedirect.com/science/article/abs/pii/S0140366419312988

Mohamed Abd Elaziz, Mohammed A.A. Al-qaness , Abdelghani Dahou , Rehab Ali Ibrahim , Ahmed A. Abd El-Latif, Intrusion detection approach for cloud and IoT environments using deep learning and Capuchin Search Algorithm https://www.sciencedirect.com/science/article/abs/pii/S0965997822003039

NicoSuranthaa,WingkyR.Wicaksonob , Design of Smart Home Security System using Object Recognition and PIR Sensor https://reader.elsevier.com/reader/sd/pii/S1877050918314881?token=FCC9158E318D8BE14F7E8BE5A92AF50FBCA3B756A29C6543451CDC5DDFE5D8C2A7EA30C2B123B7BDFAA75A3F65FCB440&originRegion=eu-west-1&originCreation=20230408142946

Olutosin Taiwo, Absalom E. Ezugwu, Internet of Things-Based Intelligent Smart Home Control System https://www.hindawi.com/journals/scn/2021/9928254/

Olutosin Taiwo, Absalom E. Ezugwu, Olaide N. Oyelade, Mubarak S. Almutairi Enhanced Intelligent Smart Home Control and Security System Based on Deep Learning Model https://www.hindawi.com/journals/wcmc/2022/9307961/

P. Pandiaraja; Saarumathi R; Parashakthi M; Logapriya R, An Analysis of Abnormal Event Detection and Person Identification from Surveillance Cameras using Motion Vectors with Deep Learning https://ieeexplore.ieee.org/abstract/document/10085466

P.R. Nargelekar, J.S. Morbale, Md. Atif Khan, Shrishti Singh, Souvik De, Smart Intrusion Detection System https://www.researchgate.net/profile/Prajakta-

Naregalkar/publication/362537642_Smart_Intrusion_Detection_System/links/62ef6658505511283e98fd9f/Smart-Intrusion-Detection-System.pdf

Preeti Nagrath, Narina Thakur, Rachna Jain, Dharmender Saini, Nitika Sharma & Jude Hemanth, Understanding New Age of Intelligent Video Surveillance and Deeper Analysis on Deep Learning Techniques for Object Tracking https://link.springer.com/chapter/10.1007/978-3-030-89554-9_2

Rashmiranjan Nayak; Mohini Mohan Behera; Umesh Chandra Pati; Santos Kumar Das, Video-based Real-time Intrusion Detection System using Deep-Learning for Smart City Applications https://ieeexplore.ieee.org/abstract/document/9117960

S. Brindha; N.Nazeeya Anjum; R.G.Sharath Kumar; S.Adithya Subramani; V. Rithika; N. Jayapreetha; D. Swetha, Augmented Security System for Commercial buildings by Manipulating Object Detection and Admin Panel https://ieeexplore.ieee.org/abstract/document/9767997

Chen, Chunling, Ziyue Zheng, Tongyu Xu, Shuang Guo, Shuai Feng, Weixiang Yao, and Yubin Lan. 2023. "YOLO-Based UAV Technology: A Review of the Research and Its Applications" Drones 7, no. 3: 190. https://doi.org/10.3390/drones7030190