Advanced Journal of Science, Technology and Engineering

Volume 2, Issue 1, 2022 (pp. 1-27)

www.abjournals.org

FACIAL RECOGNITION ATTENDANCE SCHEME ON CCTV CAMERAS USING OPEN COMPUTER VISION AND DEEP LEARNING: A CASE STUDY OF INTERNATIONAL UNIVERSITY OF EAST AFRICA (IUEA)

Edison Kagona

Department of Computer Science and Information Technology, International University of East Africa, Kampala, Uganda

E-mail: edisonmat15@gmail.com

Cite this article:

Edison Kagona (2022), Facial Recognition Attendance Scheme on CCTV Cameras Using Open Computer Vision and Deep Learning: A Case Study of International University of East Africa (IUEA). Advanced Journal of Science, Technology and Engineering 2(1), 1-27. DOI: 10.52589/AJSTE-HYVTCZ9E.

Manuscript History

Received: 8 May 2022

Accepted: 26 May 2022

Published: 9 Aug 2022

ABSTRACT: *There are plenty of home security cameras currently on the market, many allowing you to stream live footage over the internet and receive alerts whenever someone wanders past. This paper discusses the facial recognition attendance scheme on CCTV Cameras using OpenCV and Deep learning. Generally, Face Recognition is a method of identifying or verifying the identity of an individual by using their face. Various algorithms are there for face recognition but their accuracy might vary. In this paper, we discuss how we can do face recognition using deep learning. We used face embeddings to perform deep_metric_learning and the development steps of the scheme were; face detection, feature extraction, and lastly comparing faces. The legacy system used at IUEA Was firstly studied in more detail. More requirements for the proposed system were obtained and the system was developed. The interfaces for the new system were implemented using HTML, Bootstrap, and Django which is a high-level Python web framework. After the implementation, the new system was then tested, validated, and deployed for use.*

KEYWORDS: Facial Recognition, Detection, CCTV Cameras, OpenCV, Deep Learning.

INTRODUCTION

Background of the Problem

Face detection can be regarded as a specific case of object-class detection. In object-class detection, the task is to find the locations and sizes of all objects in an image that belongs to a given class. Examples include upper torsos, pedestrians, and cars. Face-detection algorithms focus on the detection of frontal human faces. It is analogous to image detection in which the image of a person is matched bit by bit. The image matches the image stored in the database. Any facial feature changes in the database will invalidate the matching process (M. Gowtham, 2021). A facial recognition system is a technology capable of identifying or verifying a person from a digital image or a video frame from a video source. There are multiple methods in which facial recognition systems work, but in general, they work by comparing selected facial features from a given image with faces within a database. It is also described as a Biometric Artificial Intelligence based application that can uniquely identify a person by analyzing patterns based on the person's facial textures and shape (Techopedia.com, 2019)

Over the past two decades, the problem of face recognition has attracted substantial attention from various disciplines and has witnessed impressive growth in basic and applied research, product development, and applications. Face-recognition systems have already been deployed at ports of entry at international airports in Australia and Portugal. Studies of how humans perceive faces have generated many interesting findings that can be used to help design practical systems. Besides applications related to identification and verification such as access control, law enforcement, ID (Sinha, 2010) and licensing, and surveillance, face recognition has also proven useful in applications such as human-computer interaction, virtual reality, database retrieval, multimedia, and computer entertainment (Sinha, 2010). During 1964 and 1965, Bledsoe, along with Helen Chan and Charles Bisson, worked on using the computer to recognise human faces (Bledsoe 1966a, 1966b; Bledsoe and Chan 1965). He was proud of this work, but little of the work was published because the funding was provided by an unnamed intelligence agency that did not allow much publicity. Based on the available references, it was revealed that Bledsoe's initial approach involved the manual marking of various landmarks on the face such as the eye centres, mouth, etc., and these were mathematically rotated by a computer to compensate for pose variation. The distances between landmarks were also automatically computed and compared between images to determine the identity (Bledsoe, 1965).

In the United States of America (USA) and many parts of the West, facial recognition systems including those from Apple, Google, Amazon, Microsoft, and Facebook have sparked intense consumer debate and are under heavy fire from privacy advocates for their ability to scan people's faces and identify them without their consent. Legal challenges, including high-profile lawsuits against these tech giants, make headlines almost daily and have already forced some companies to reverse course. In June 2019, Microsoft quietly deleted a massive online data set that contained more than 10 million images of 100,000 individuals that were used to train other companies' facial recognition systems. And later, also, Facebook faced heavy fines over its use of facial recognition in photo tagging and announced it would ask its users' permission to use the feature and delete their "face recognition template" if they did not opt-in. The use of biometrics and other personal data has also given rise to a host of new regulations in the West. For example, the EU's General Data Protection Regulation (GDPR) prevents the "processing" and sharing of biometric data without permission. In the United States (US), Illinois,

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Washington, and Texas have passed biometric privacy laws that impose restrictions on organizations that collect biometric information. In California, the California Consumer Privacy Act (CCPA) regulates biometric data, and the cities of Oakland and San Francisco have banned municipal authorities from using facial recognition (Petrock, 2019). Kenya's National Police Service (NPS) has launched a facial recognition system for CCTV cameras installed along major roads and highways as part of an upgrade of its Integrated Command and Control System (ICCS). The ICCS CCTV network, which includes thousands of cameras, already has implemented license plate recognition technology. The facial recognition system will be monitored by the Directorate of Criminal Investigations as part of the Critical Incident Management Suite (CIMS) and will issue alerts to law enforcement when a suspect is identified (Burt, 2018).

Most face recognition systems rely on face recognition algorithms to complete the following functional task as suggested by Shang-Hung Lin. (2000, p.2).

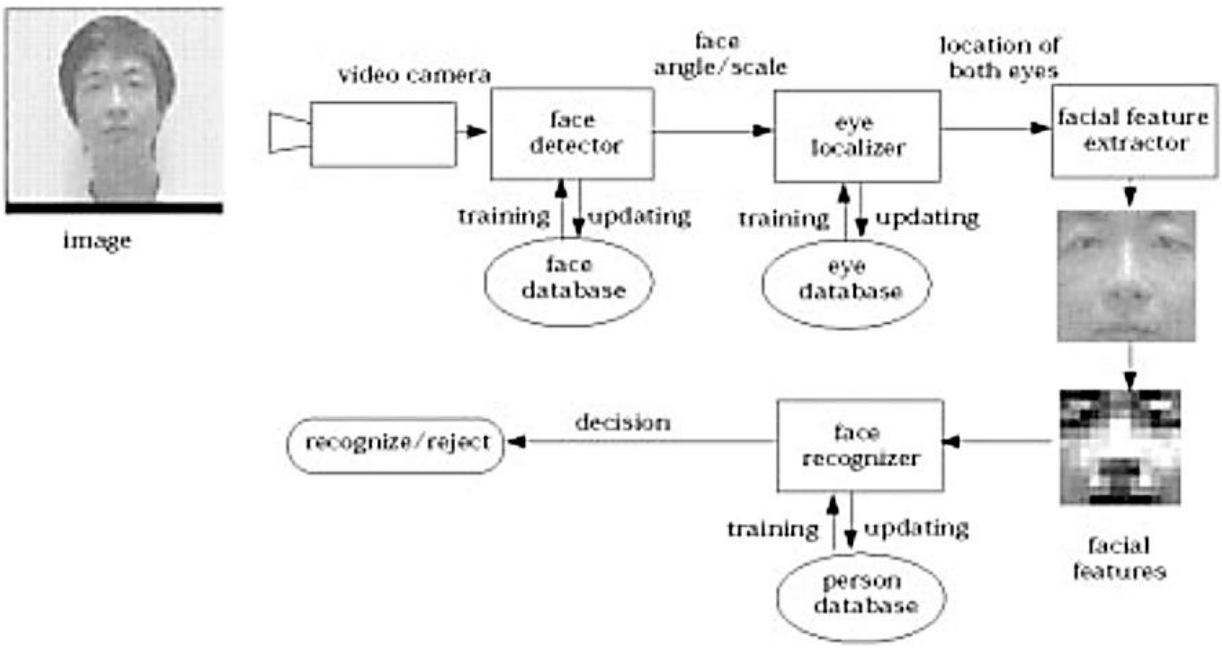

Figure 1: Face Recognition System Framework

Face recognition system framework (Hung, 2000)

The figure below shows a simplified diagram of the framework for face recognition from the study suggested by (Hung, 2000).

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

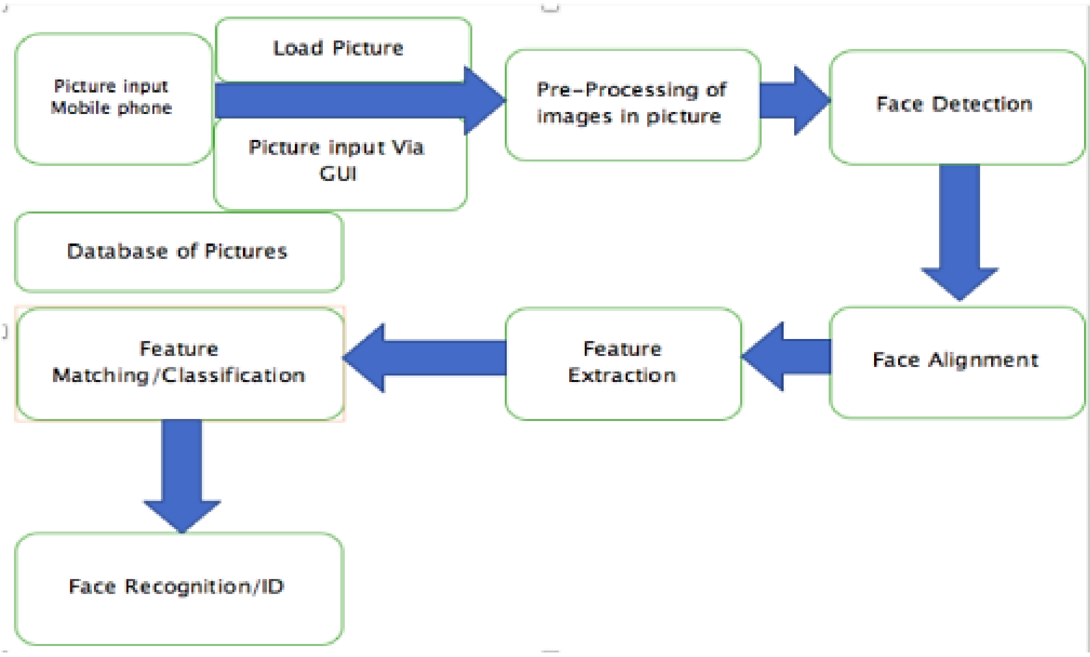

Figure 2: Simplified diagram from the framework for face recognition

From the figure, above, a Face Detection or face detector will detect any given face in the given image or input video. Face localization will detect where the faces are located in the given image/video, by use of bounding boxes. Face Alignment is when the system will find a face and align landmarks such as the nose, eyes, chin, and mouth for feature extraction. Feature extraction extracts key features such as the eyes, nose, and mouth to undergo tracking. Feature matching and classification match a face based on a trained data set of pictures from a database of about 200 pictures. Face recognition gives a positive or negative output of a recognized face based on feature matching and classification from a referenced facial image. Face detection is the process of locating a face in a digital image by any special computer software built for this purpose (Feruad, 2000). Face detection is "To detect a face in an image means to find its position in the image plane and its size or scale. "As figure 2 shows, the detection of a face in a digital image is a prerequisite to any further process in face recognition or any face processing software.

In the early years, face detection algorithms focused mainly on the frontal part of the human face (Srinivasan, 2016). However, in recent years (Cynganek, 2013), suggest that newer algorithms take into consideration different perspectives for face detection. Researchers have used such systems but the most challenge that has been faced is to make a system detect faces irrespective of different illumination conditions. This is based on a study by (Castrillon, 2011) on the Yale database which contains higher resolution images of 165 frontal faces. Face detection is often classified into different methods. To face the first major problem of the project (Detecting students' faces), a wide range of techniques have been researched. These several face detection techniques/methodologies have been proposed by many different researchers and are often classified into major categories of different approaches.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

i). Knowledge-based,

ii). Feature invariant

iii). Template matching

iv). Appearance-based approaches.

Knowledge-Based Method

This method uses human knowledge or human coding to model facial features based on the nature of the human face such as two eyes, a mouth, and a nose. This is very easy to apply the rules but very difficult to detect in the various background depending on the pose and illumination. Low detection accuracy with a small burden of calculation and short detection time. (Yang, 2002), to investigate this method, created a multiple resolution hierarchy of images by averaging and subsampling as in the figure below.

. (a) n = 1, original image. (b) n = 4. (c) n = 8. (d) n = 16. Original and corresponding low resolution images. Each square cell consists of pixels in which the intensity of each pixel is replaced by the average intensity of the pixels in that cell.

Figure 3: Knowledge-Based Method

They subdivided these resolution hierarchies into three levels with level 1 being the lowest resolution which only searches for face candidates and is further processed at finer resolutions. At level 2 they used the face candidate in level 1 alongside local histogram equalization followed by edge detection. At level three, the surviving face candidate region uses a set rule responding to facial features such as mouth and eyes. They conducted their experiment on 60 images. Their system located faces on 50 of these images and 28 images gave a false alarm, thus giving a success rate of 83.33% and a false alarm rate of 46.66%.

Feature-Based-Methods

This method uses algorithms to look for structural features regardless of the pose, viewpoint, or lighting conditions to find faces. Here, local features such as eyes, nose, and mouth are first of all extracted and their locations and local statistics (geometric and/or appearance) are fed into a structural classifier. A big challenge for feature extraction methods is feature "restoration", this is when the system tries to retrieve features that are invisible due to large variations, e.g. head Pose when we are matching' a frontal image with a profile image (Divyarajsinh N. Parmar1, 2013).

Template Matching Methods

This uses standard facial patterns stored for use to correlate an input image with the stored pattern to compute for detection. Appearance Base Methods; uses a set of training sets of

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

images to learn the templates and capture the representative of facial appearance. Furthermore, Yang et al. also carried out their experiments on a standard database set which is shown in the table below.

Table 1: Showing the detection rate results and false detection rates (Yang, 2002).

Test Sets for Face Detection

Data Set	Location	Description
MIT Test Set [154]	http://www.cs.cmu.edu/~har	Two sets of high and low resolution gray scale images with multiple faces in complex background.
CMU Test Set [128]	http://www.cs.cmu.edu/~har	130 gray scale images with a total of 507 frontal faces.
CMU Profile Face Test Set [141]	ftp://eyes.ius.cs.cmu.edu/usr20/ftp/testing_face_images.tar.gz	208 gray scale images with faces in profile views.
Kodak Data Set [94]	Eastman Kodak Corporation	Faces of multiple size, pose and under varying illumination in color images. Designed for face detection and recognition.

Table 2: Results of Two Image Test Sets experimented. Yang et al. (2002, p.54).

Experimental Results on Images from Test Set 1 (125 Images with 483 Faces) and Test Set 2 (23 Images with 136 Faces) (See Text for Details)

Method	Test Set 1		Test Set 2	
	Detection Rate	False Detections	Detection Rate	False Detections
Distribution based [154]	N/A	N/A	81.9%	13
Neural network [128]	92.5%	862	90.3%	42
Naive Bayes classifier [140]	93.0%	88	91.2%	12
Kullback relative information [24]	98.0%	12758	N/A	N/A
Support vector machine [107]	N/A	N/A	74.2%	20
Mixture of factor analyzers [175]	92.3%	82	89.4%	3
Fisher linear discriminant [175]	93.6%	74	91.5%	1
SNoW with primitive features [176]	94.2%	84	93.6%	3
SNoW with multi-scale features [176]	94.8%	78	94.1%	3
Inductive learning [38]	90%	N/A	N/A	N/A

As Table 2 summarizes, the experimental results show images of different training sets with different parameters of tuning which has a direct impact on the training performance. For example, dimensionality reduction is carried out to improve computation efficiency and detection efficacy, with image patterns projected to a lower-dimensional space to form a discriminant function for classification. Also, the training and execution time and the number of scanning windows in these experiments influenced the performance in some way. (Hjelmås, 2001), classifies face detection methodologies into two major categories. Image-based approaches, which is further sub-categorized into linear subspace methods, neural networks, and statistical approaches.

Image-Based Approaches

Most of the recent feature-based attempts in the same study (Hjelmås, 2001) have improved the ability to cope with variations, but are still limited to the head, shoulder, and part of the frontal faces. There is, therefore, a need for techniques to cope in hostile scenarios such as

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

detecting multiple faces in a cluttered scene, e.g., clutter-intensive background. Furthermore, this method ignores the basic knowledge of the face in general and uses face patterns from a given set of images. This is mostly known as the training stage in the detection method. From this training stage, the system may be able to detect similar face patterns from an input image. A decision of face existence by the system is now established based on a comparison of the distance between the pattern from the input image and the training image with a 2D intensity array extracted from the input image. Most image-based approaches use window-scanning techniques for face detection.

The window scanning algorithm searches for possible face locations at all scales. This method depends on window scanning algorithms. Other research carried out on this method depends on window scanning algorithms (Ryu, 2006). in their study experimented with the scanning window techniques discussed by (Hjelmås, 2001) in their system. They go further to experiment with their system, based on a combination of various classifiers for a more reliable result compared to a single classifier. They designed multiple face classifiers which can take different representations of face patterns. They used three classifiers, Gradient feature classifier which contains the integral information of pixel distribution that returns certain invariability among facial features. The second classifier is Texture Feature which extracts texture features by correlation (uses joint probability occurrence of the specified pixel), variance (measures the number of local variations in an image), and entropy (measures image disorder). The third classifier used here is Pixel Intensity Feature, which extracts pixel intensity features of the eye, nose, and mouth region for determining the face pattern. They further used the Coarse-To-Fine Classification approach with their classifications for computational efficiency. Based on 1056 images that were obtained from the AT&T, BioID, Stirling, and Yale dataset, they achieved the results presented in Table 2.4 and Table 2.5 Ryu et al. (2006, p.489) The first face classification of their experiment concerning the shift in both x and y direction achieved a detection rate of 80% when images are shifted within 10 pixels in the x-direction and 4 pixels in the y-direction. The second and third faces of their classification showed a detection rate of over 80% when 2 pixels were a shift in both x and y directions respectively.

Table 3: Results showing Exhaustive full scanning method and proposed scanning method (Ryu, 2006).

Test DB	Detection results						Reduction rates of # of scans
	Exhaustive full scanning method			Proposed scanning method			
	Detection rate	# of false	# of scans per image	Detection rate	# of false	# of scans per image	
IMM	96.2 %	28	755,418	95.7 %	8	72,273	90.4 %
Caltech	94.5 %	12	1,369,067	93.0 %	10	176,674	87.1 %
AR	95.7 %	22	1,128,541	95.0 %	6	142,136	87.3 %
WWW	80.7 %	46	4,312,203	83.7 %	12	513,152	88.1 %

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Table 4: Performance Comparison by different Researchers and Proposed System by Ryu et al (2006, p.490)

	Detection rate	# of false
Rowley method [3]	86.2%	23
Froba method [9]	87.8%	120
Feraud method [10]	86.0%	8
Proposed method (coarse-to-fine search)	86.6%	19
Proposed method (full search)	89.1%	32

As seen in table 4**,** their system achieved a detection rate between 93.0% and 95.7%. Rowley et al. 1998 in their study on Neural Network-Based face detection, experimented on their system which applies a set of neural network-based filters to an image and then uses an arbitrator to combine the outputs. They tested their system against two databases of images. The CMU database which was made of 130 images and the FERET database achieve a detection rate of 86.2% with 23 false detections. (Feraud, 2001) Also experimented on neural network-based face detection technique. They used a combination of different components in their system (motion filter, colour filter, pre-network filter, and large neural network). The pre-network filter is a single multilayer perceptron, with 300 inputs corresponding to the extracted sizes of the sub-windows, hidden with 20 neurons, and outputs a face/non-face for a total number of weights [reference]. These components, with a combination of neural networks, achieved an 86.0% detection rate with 8 false detections, based on a face database of 8000 images from Sussex Face Database and CMU Database which is further subdivided into different subsets of equal sizes corresponding to different views.

Table 5 and Table 6 (Feraud, 2001) below show the experimental results carried out by these researchers.

Table 5: Showing Results of Sussex Face Database (Feraud, 2001).

Results on Sussex Face Database

orientation (degree)	CGM1	CGM3	CGM5	Ensemble	Conditional ensemble	Conditional mixture
0	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %	100.0 %
10	62.5 %	100.0 %	87.5 %	100.0 %	100.0 %	100.0 %
20	50.0 %	100.0 %	87.5 %	87.5 %	100.0 %	100.0 %
30	12.5 %	100.0 %	62.5 %	62.5 %	100.0 %	100.0 %
40	0.0 %	100.0 %	50.0 %	12.5 %	62.5.0 %	87.5 %
50	0.0 %	75.0 %	0.0 %	0.0 %	37.5 %	62.5 %
60	0.0 %	37.5 %	0.0 %	0.0 %	0.0 %	37.5 %
70	0.0 %	37.5 %	0.0 %	0.0 %	0.0 %	25.0 %

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Table 6: Showing Results of CMU Test Set A (Feraud, 2001).

Results on the CMU Test Set A

Model	Detection rate	False alarms rate	False alarms
GM	84 %	1000 10^{-6}	≈ 20000
CGM1	77 %	5.43 10^{-6}	47
CGM3	85 %	6.3 10^{-6}	212
CGM5	85 %	1.36 10^{-6}	46
one SWN (Rowley95)	84 %	8.13 10^{-6}	179
Ensemble	74 %	0.71 10^{-6}	24
Conditional ensemble	82 %	0.77 10^{-6}	26
Conditional mixture	87 %	1.15 10^{-6}	39

Wang et al. (2016) in their study to support a neural network face detector used a multi-task convolutional neural network-based face detector, which relies directly on learning features from images instead of hand-crafted features. Hence their ability to differentiate faces from uncontrolled backgrounds or environments. The system experimented on the used Region Proposed Network which generates the candidate proposal and the CNN-Based detector for the final detection output. They experimented with this based on 183200 images from their database and used the AFLW dataset for validation. Their face detector system was evaluated on AFW, FDDB and Pascal faces datasets respectively, and achieved a 98.1% face detection rate. The authors did not reveal all the facts leading to the development of the system and I have limited time to implement this on OpenCV. 2.8 (Wang et al. 2016 p.479), shows the different comparisons of their system against another state of the arts. (Wang et al. 2016 p.480), discuss their system (Face Hunter) perform better than all other structured models. However, this cannot be independently verified as this system was commercialized. One cannot conclude if this was for marketing purposes or a complete solution to the problem as I have limited time to implement it.

Feature-based approaches

This category depends on extracted features that are not affected by variations in lighting conditions and pose. This according to these researchers, (Hjelmås, 2001), further clarifies that "visual features are organized into a more global concept of face and facial features using the information of face geometry ". This technique in my own opinion will be slightly difficult to use for images containing facial features from an uncontrolled background. This technique relies on feature analysis and feature derivation to gain the required knowledge about the face to be detected. The features extracted are skin colour, face shape, eyes, nose, and mouth. On the other hand, another study by (Mohamed, 2007, p.2) suggests that the "human skin colour is an effective feature used to detect faces, although different people have different skin colour, several studies have shown that the basic difference based on the intensity rather than their chrominance". The texture of the human skin can therefore be separated from different objects. Feature methods for face detection using features for face detection. Some users depend on the edges and then group the edges for face detection. Furthermore, (Sufyanu, 2016) suggest a

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

good extraction process will involve feature points chosen in terms of their reliability of automatic extraction and importance for face representation. Most geometric feature-based approaches use the Active Appearance Model (AAM). This allows localization of facial landmarks in different ways to extract the shape of facial features and movement of these features as expression evolves (Hjelmås, 2001), further placing the feature-based approach into sub-categories of; Low-level analysis (Edges, Gray-levels, Color, motion, and generalized measure).

Feature analysis (Feature searching and constellation analysis)

Active shape models (Snakes, Deformable templates, and Point distribution models (PDMS). Figure 4**:** shows the different approaches for Face detection as reported in a study by (Hjelmås, 2001), which can be compared with Figure 5 showing the same classification by (Macwan, 2014).

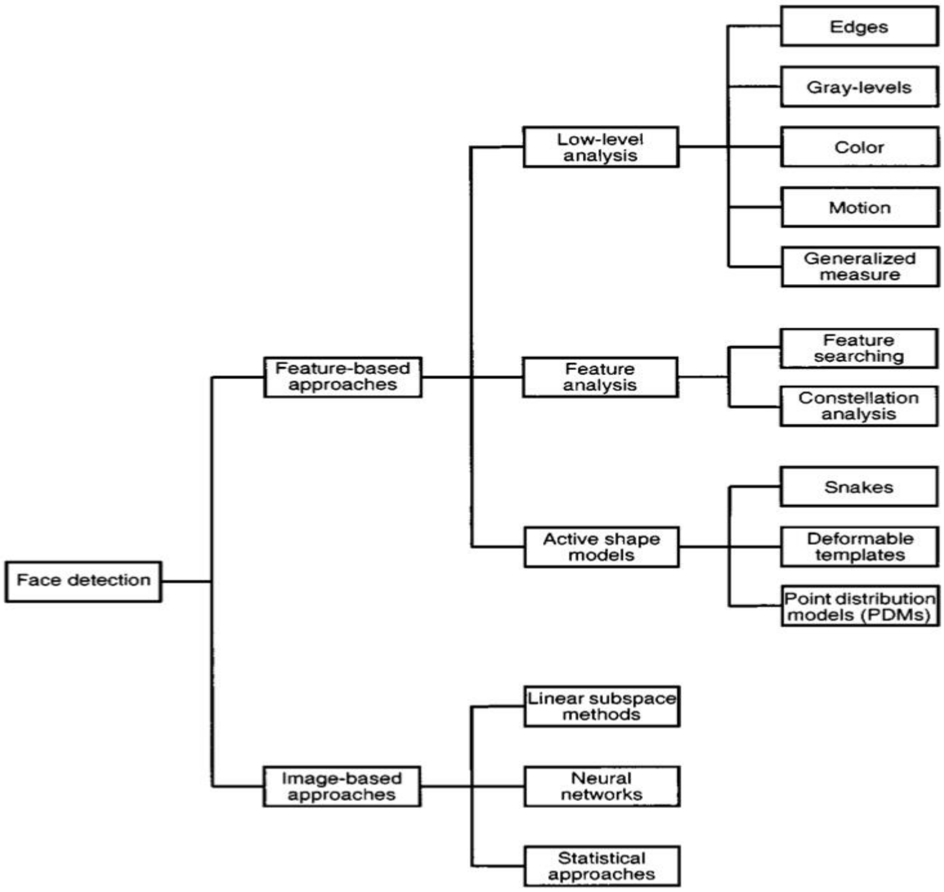

Figure 4: Face Detection, classified into different methodologies (Hjelmås, 2001).

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

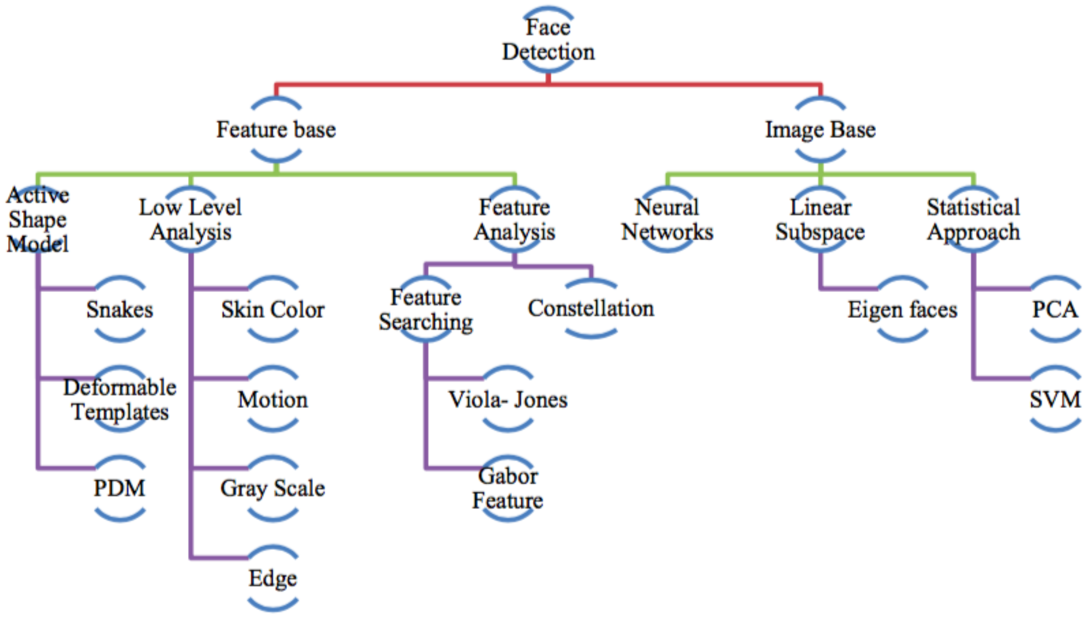

Figure 5: Various Face Detection Methodologies (Macwan, 2014)

(Hjelmås, 2001) In their study, show experiment based on an edge-detection based approach for face detection, on a set of 60 images of 9 faces, with complex backgrounds correctly detected 76% of faces with an average of two false alarms per image. (Nehru and Padmavathi, 2017), in their study, experimented with face detection based on the Viola-Jones algorithm in a dataset of dark and coloured men to support their statement which states “It is possible to detect various parts of the human body based on the facial features present”, like the eyes, nose, and mouth. In this case, systems as such will have to be trained properly to be able to distinguish features like the eyes, nose, mouth, etc., when a live dataset is used. The Viola-Jones algorithm to detect faces as seen in the images in Figure 2.7 shows dark and coloured skin faces detected accurately.

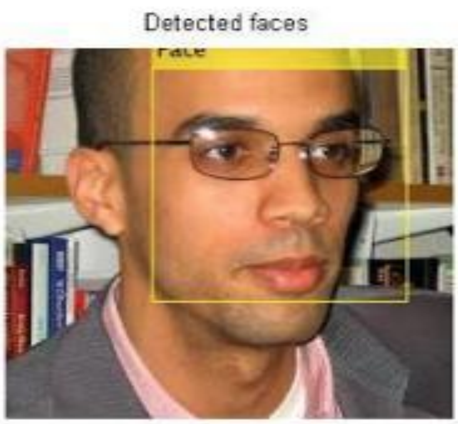

Figure 6: Face Detection in Dark and Colored Men (Nehru and Padmavathi, 2017).

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Also, in support of the claim made by (Nehru and Padmavathi, 2017), the research carried out by Viola-Jones to come up with the Viola-Jones algorithm in face detection, has had the most impact in the past decade. As suggested by (Mayank Chauhan, 2014), the Viola-Jones in face detection is widely used in genuine applications such as digital cameras and digital photo managing software. This claim is made based on a study by (Jones, 2017). Table 2.9 gives a summary of the results obtained by these experts, showing various numbers of false and positive detections based on the MIT and CMU database set with 130 images and 507 faces.

Detector \ False detections	10	31	50	65	78	95	110	167	422
Viola-Jones	78.3%	85.2%	88.8%	89.8%	90.1%	90.8%	91.1%	91.8%	93.7%
Rowley-Baluja-Kanade	83.2%	86.0%	-	-	-	89.2%	-	90.1%	89.9%
Schneiderman-Kanade	-	-	-	94.4%	-	-	-	-	-
Roth-Yang-Ahuja	-	-	-	-	(94.8%)	-	-	-	-

Figure 7: Various Detection rates by different algorithms showing positive and false detection rates (Jones, 2017).

(Wang, Modular Real-Time Face Detection System, 2015) States that" the process of searching a face is called face detection. Face detection is to search for faces with different expressions, sizes and angles in images in possession of complicated light and background and feeds back parameters of face". In their study, they tested face detection based on two modules which shows one module uses a combination of two algorithms (PCA with SVM) and the other module based on a real-time field-programmable gate array (FPGA). With these, they concluded with their results of face detection accuracy of 89%. Table 7 is a screenshot taken from this paper to show the experimental results of two units combined to investigate the accuracy of the system.

The number of testing	Correct detection	False detection	Accuracy
1000	890	110	89 %

Figure 8: Detection accuracy system by (Wang, 2015).

Another method is the Learning-based method which includes machine learning techniques that extract discriminative features from a trained dataset before detection. Some well-known classifiers used for face detection, based on a study by (Thai, 2011) are Canny, Principal Component Analysis (PCA), Support Vector Machine (SVM), and Artificial Neural Network (ANN). Although used for facial expression classification, the algorithms are, however, also used in the initial stage of their experiment, which is the detection phase. Their experiment achieved some results which are shown in Table 9. A screenshot from (Thai, 2011).

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Method	Classification Accuracy %
Rapid Facial Expression Classification Using Artificial Neural Networks [10]	73.3%
Facial Expression Classification Using Multi Artificial Neural Network [11]	83.0%
Proposal System (Canny_PCA_ANN)	85.7%

Figure 9: Comparing Different Algorithms on classification rates (Thai, 2011).

The overall objective of the Face detection part of this project will be to find out if any faces exist in the input image and if present, will return the location in bounding boxes and the extent of each face, counting the number of faces detected. It is a challenge to this project due to the variations in location, scale, pose orientation, facial expression, illumination or lighting condition, and various appearance features such as facial hair, makeup, etc. It will be difficult to achieve an excellent result. However, the performance of the system will be evaluated, taking into consideration the learning time, execution time, number of samples required for training, and the ratio between the detection rate and false detections. Table 10 below shows experiments from different researchers. They have used different sizes of the image dataset. Some have used a combination of different algorithms and applied other methods like colour filtering etc and different training sets to obtain their results. However, we can conclude the Viola-Jones algorithm which is on its own classifies images based on local features only and can still detect at very high accuracy and rapidly than pixel-based systems (Jones, 2017).

Table 7: Comparison of Results by Different Researchers Showing Face Detection Accuracy and False Detection (Baseer. K, 2015).

Author	Method	Detection Accuracy	#False Detection
(Yang, 2002)	Knowledge-Based-Method	83.33%	28
(Ryu, 2006)	Image-Based Method	89.1%	32
(Feruad, 2000)	neural network-based	86.0%	8
(Rowley, 1998)	Neural Network-Based	86.2%	23
(Wang, 2016)	CNN-Based	98.1%	

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

(Hjelmås, 2001)	Edge Detection-Based	76%	30
(Jones, 2017)	Viola-Jones	88.84%	103
(Wang, 2016)	PCA with SVM)	89%	110
(Thai, 2011)	Canny_PCA_ANN	85.7%	N/A

METHODOLOGY

System Study

To understand and gain the requirements of the new system, the current employee attendance of IUEA was studied. This was studied by interacting with some of the individuals working in the organization by asking them a series of informal questions to obtain knowledge and understanding of the entire structure of how attendance of employees is recorded and used for future purposes. The collected data was analysed using quantitative and qualitative methods of data analysis. Using the observational technique, we noticed that some employees can move out during the day since the records capture their attendance in the morning when reporting to work and evening when leaving work. The data collected was analysed using both structured and non-structured analysis approaches. Using the above techniques, we were able to understand the existing system and gain the users' opinions about the existing and the new system.

Existing system

IUEA's current system is by the use of ID cards that are swiped on the biometric system which captures the time an employee has arrived and the time the employee leaves work and manually writes down names in a book as another method of capturing the attendance of employees. These records are used every end of the week to audit the attendance of employees, hence capturing loopholes in the productivity of the organization and employees at large.

Existing system strengths and weaknesses

The existing system being used at IUEA is the widely used system to capture employee attendance to a percentage of 80% in most organizations. However, this system has strengths and weaknesses as described below.

1. **Strengths**

a) The existing system doesn't need a lot of computer literacy or technology exposure for it to operate since it involves some paperwork.

b) The existing system is easy to operate, train and maintain for both users and managers of the system

c) It easily captures records of the arrival and departure of employees before and after work since it involves the use of swiping identity cards that are tagged to the biometric system to capture the details of the employees.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

2. Weaknesses

a) Records kept on paper increases the chance of data loss due to bulk files requiring extra time to recover the records or to go through all the files and papers to find the needed information. Hence, slowing down decision-making at the farm.

b) There is less security of the company records since they are kept in a book/file which doesn't require authentication of user password to gain access to the company records and official confidential data.

c) Another weakness is that an employee can leave the ID card with a colleague that is present to capture that person. The time in and time out records in the case that person is absent unofficially and this creates disguised unproductivity in the company.

d) Inconsistency in data entry on both the paper-based attendance record capturing. This happens in a way that the paper-based system is operated by the human who is prone to doing mistakes while entering the data.

e) Difficulty in making backups of records since all records are on paper and kept in files which would require extra space and effort to go through all paper records and rewrite them on other paper, hence being costly and tiresome.

System Analysis

This phase mainly describes the functional and non-functional requirements for the employee attendance system at IUEA based on the results of the system study.

User requirements

1. **Administrator (Admin):** The Administrator is the person in charge of system management and below are the roles that will be done by the administrator.

    a) Has the highest privileges as is responsible to design the system.

    b) Admin registers management team in particular HR and provides them with a unique id.

    c) Admin is responsible to take images of employees and add them to the database.

    d) Admin can view and update the details of both management and employees.

    e) Admin can also view the attendance report.

2. **User:** The user will be the person in charge of system operation

    a) They can log into the system

    b) They can view and download the attendance report.

    c) They can add and remove employees.

    d) The design time schedules at work

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Functional requirements

Functional requirements are a set of requirements that the system will need to deliver or operate. The system provides the following functionalities;

a) Capture face images via webcam or external USB camera.

b) A professional HD Camera

c) Faces on an image must be detected.

d) The faces must be detected in bounding boxes.

e) Compute the total attendance based on detected faces.

f) Crop the total number of faces detected.

g) Resize the cropped faces to match faces the size required for recognition

h) Train faces for recognition.

i) Perform recognition for faces stored on the database.

j) Compute the recognition rate of the system.

k) Perform recognition one after the other for each face cropped by Face Detector.

l) Display the input image alongside the output image side by side, on the same plot.

m) Display the name of the output image above the image in the plot area.

Non-Functional Requirements

Non-functional requirements are a set of requirements with specific criteria to judge the operation of the system. They cover ease of use to the client, security, support availability, operational speed, and implementation considerations. More specifically:

i). The system should enable the user to find it very convenient to take photos.

ii). The system should be secure.

iii). The system should have a response time of 10 seconds.

iv). The system should be easy to install.

v). The system should be 100% efficient.

vi). The system should be fast and reliable.

vii). The system should be accurate enough to provide the right results with the right inputs.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

System Design

1. Face Detection

The first task that we perform is detecting faces in the image(photograph) or video stream. Since we now know the exact coordinates/location of the face, we extract this face for further processing.

2. Feature Extraction

Next, we crop out the face from the image and extract specific features from it. Here we use face embeddings to extract these features of the face. As we know a neural network takes an image of the face of the person as input and outputs a vector that represents the most important features of a face. In machine learning, this vector is nothing but called *embedding* and hence we call this vector ***face embedding.***

When we train the neural network, the network learns to output *similar vectors* for faces that look similar. Let us consider an example, if I have multiple images of faces within different timelapse, it's obvious that some features may change but not too much. So in this problem, the vectors associated with the faces are similar or we can say they are very close in the vector space. Up to this point, we came to know how this network works, let us see how to use this network on our data. Here we pass all the images in our data to this pre-trained network to get the respective embeddings and save these embeddings in a file for the next step.

3. Comparing faces

We have face embeddings for each face in our data saved in a file, the next step is to *recognize* a new image that is *not in our data*. Hence, the first step is to compute the face embedding for the image using the same network we used earlier and then compare this embedding with the rest of the embeddings that we have. We recognize the face if the generated embedding is closer or similar to any other embedding.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

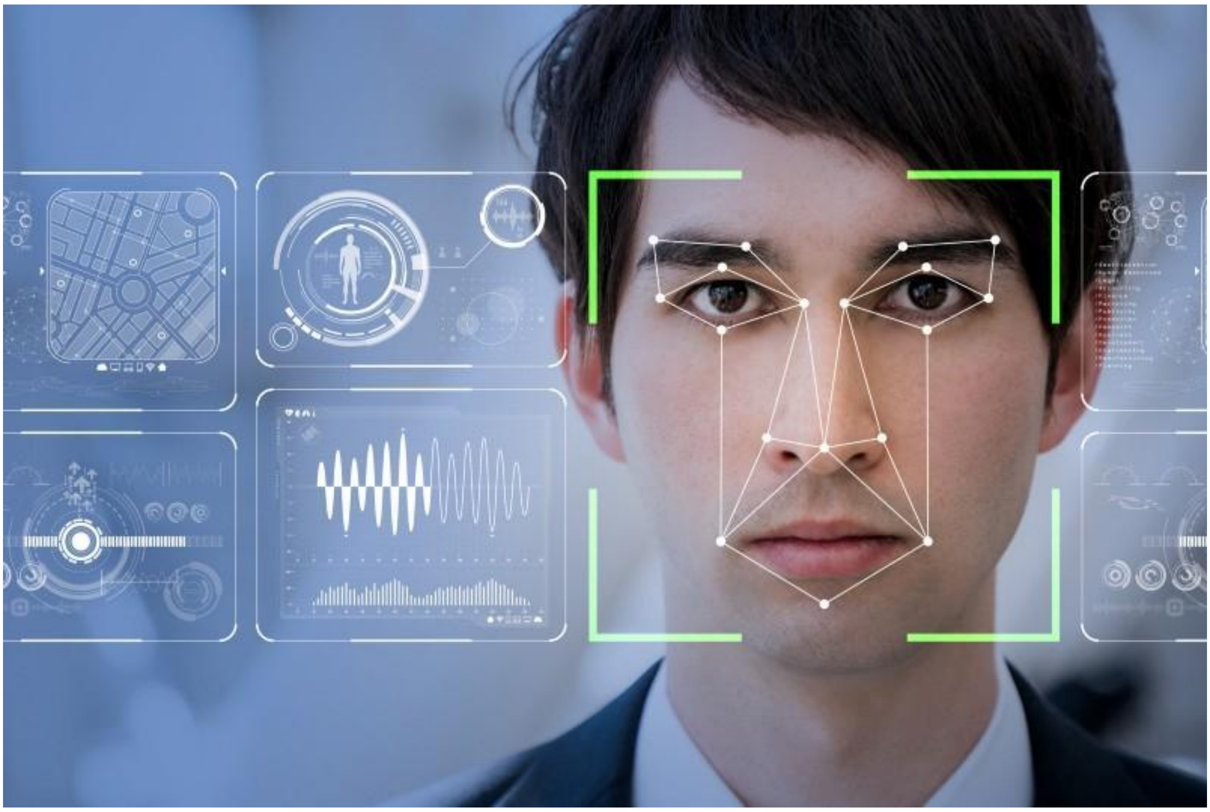

Figure 10: Image source: https://cdn-media-1.freecodecamp.org/images/1*fpDngO6lM5pDeIPOOezK1g.jpeg

The system design used for the project involves process and data flow concepts as described below.

Proposed Architectural Design

This shows the interaction between software (Internal) components and hardware (External) components with an interface to establish a framework to achieve system objectives. Both external and internal components have been considered. The internal component incorporates all the functionalities with a Graphical User Interface to allow the user to interact with the system.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

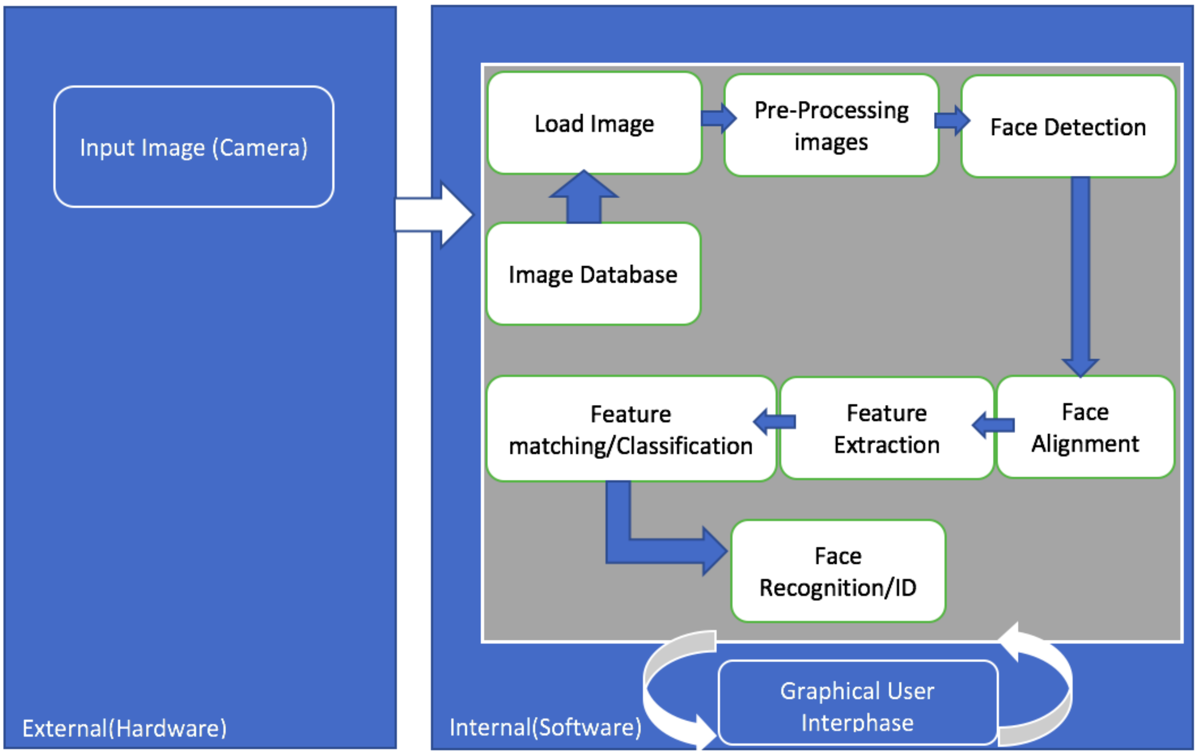

Figure 11: External and Internal components of the System.

Process modelling

The Process modelling section shows a sequence of activities and logical flow within the proposed system visually.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

a) Level 1 Data Flow Diagram

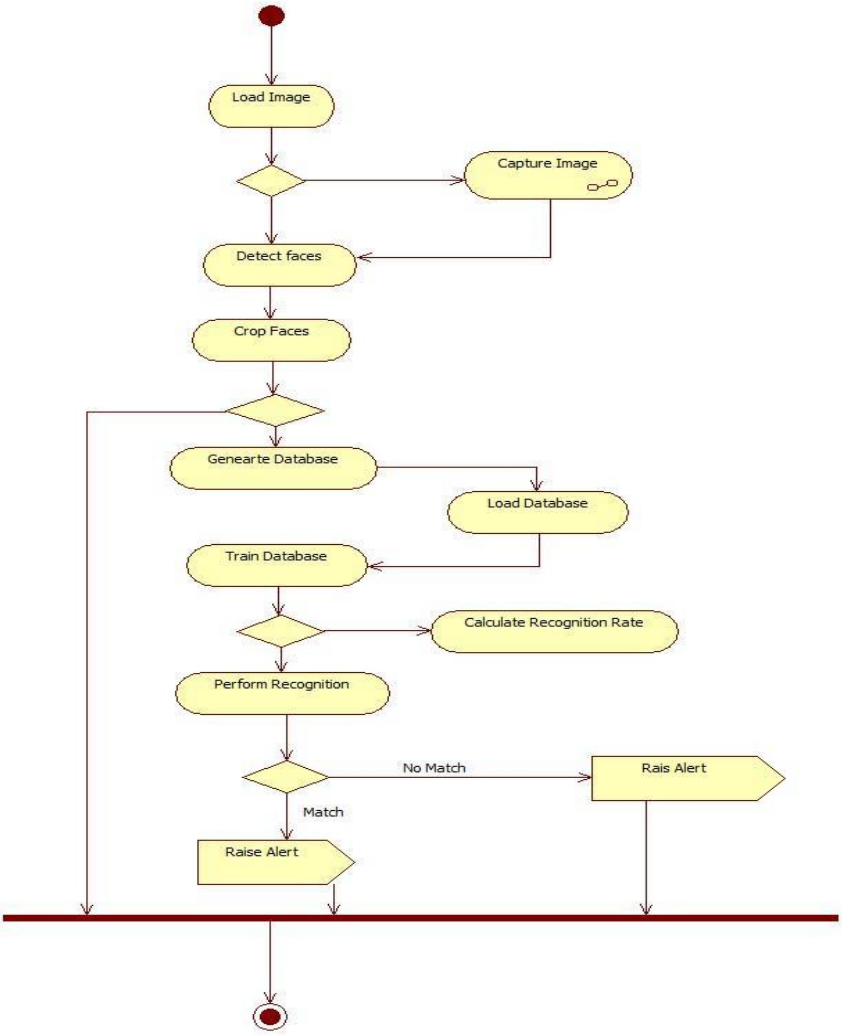

Figure 12: Level 1 Diagram for the system.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

a) Level 2 Data Flow Diagram

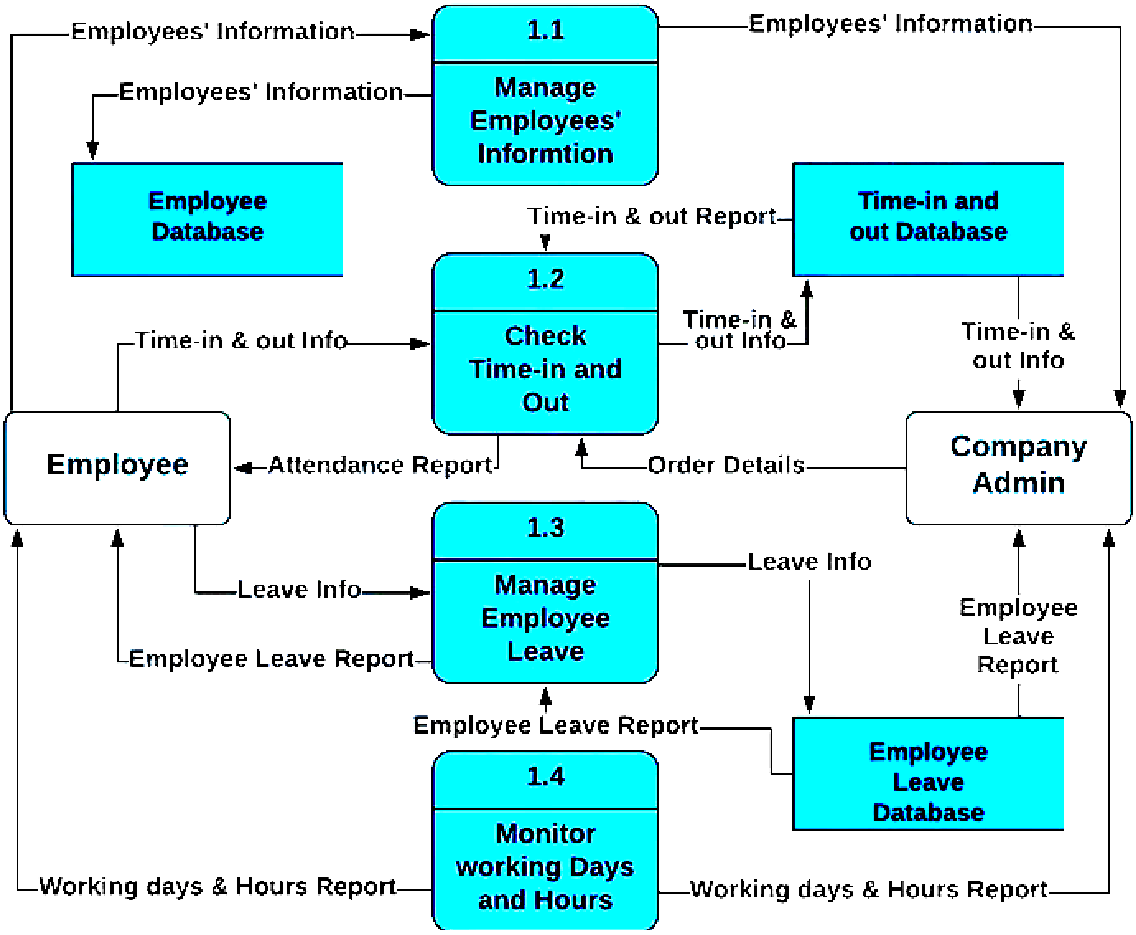

Figure 13: Data Flow Diagram for the system

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Entity Relationship Diagram (ERD)

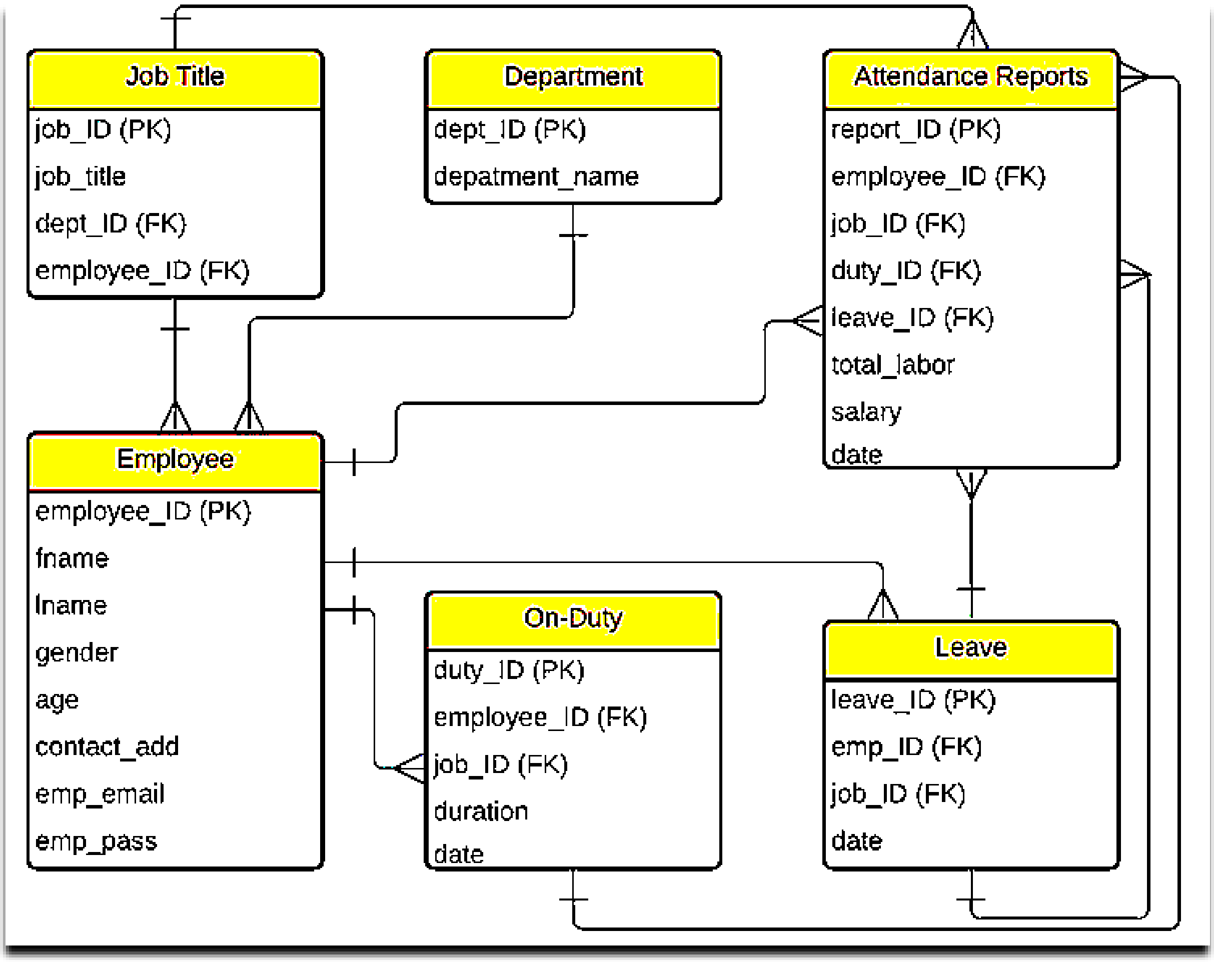

Figure 14: Entity Relationship Diagram for the system

System Implementation, Testing, and Validation

The various outputs for this design are functional specification, detailed design, user interface specification, data model, and archetype implementation plan as explained below.

The system provides a couple of functions for its users following their roles. The system prompts the users/Administrator for their usernames and passwords for authentication and verification purposes, the users provide them, and then the system verifies the user. The user (Administrator) is then able to use the system for inputting the employees' data hence saved to the database.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Sample Screen Shots Captured

1. Login page

This is where access, authentication, and authorization of users by the use of specific credentials i.e. usernames and passwords

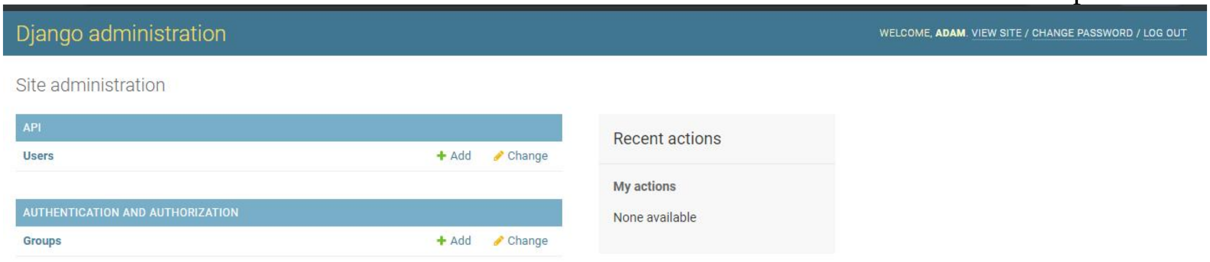

Figure 15: Administrator Panel for User Login Credentials Management.

2. Adding employees to the system and training the classifier

This is also done by the Administrator who captures the details of an employee. For example, face image, employee role, and department, and these details are stored in the database for future use in capturing the attendance of employees. On this same page, is where the classifier is trained to be able to detect and match faces stored in the database to capture the attendance of employees.

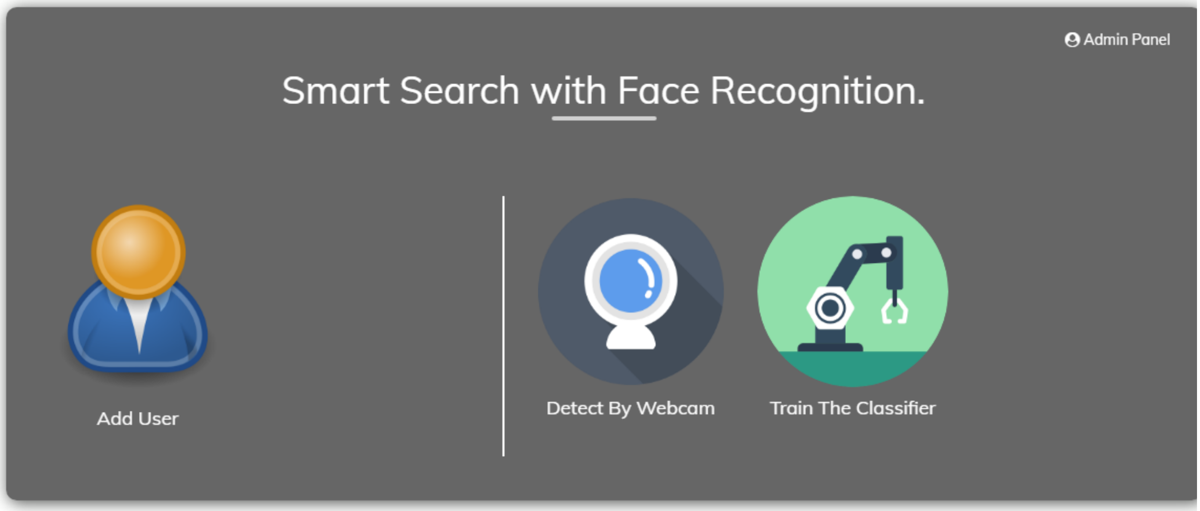

Figure 16: Registration of employees.

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Advanced Journal of Science, Technology and Engineering
Volume 2, Issue 1, 2022 (pp. 1-27)
www.abjournals.org

3. **Attendance report sample page**

This will capture the time in and out and the number of times an employee reports to work. This will be consolidated in one single report showing the dates and time captured

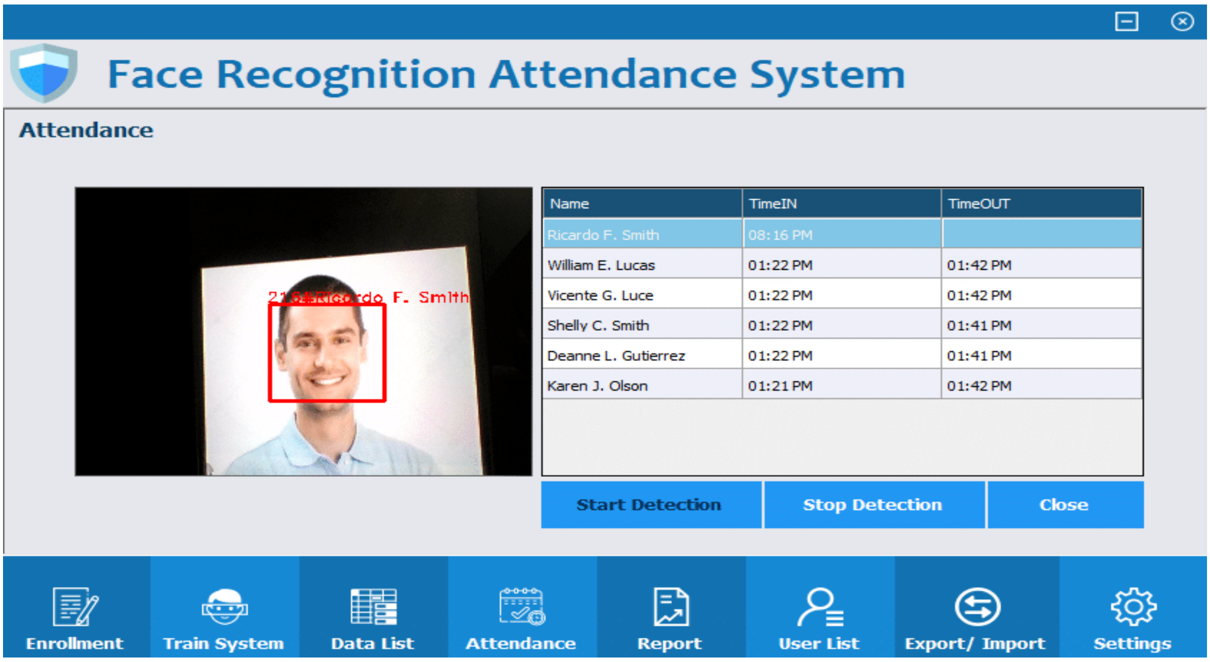

Figure 17: 3. Attendance report screen.

System Testing and Validation Results

1. Login page

Table 8: Login Page test results.

Description	**Action**	**Expected Response**	**Observed Response**
Log in with the right credentials	Authenticate, Authorize and grant access	Successfully logged in response on the screen	Worked as expected
Log in with the wrong credentials	No access granted	An incorrect username/password response is returned.	Worked as expected

2. Registration of employees and capturing images

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

Table 9: Registration of employees' results.

Description	Action	Expected Response	Observed Response
Entering employee data into the system	Datasets are matched in the database	Training the classifier to be able to match the data and return the right matches to record the presence or absence of employees.	Worked as expected.

3. View attendance employee report

Table 10: Employee report results

Description	Action	Expected Response	Observed Response
Access and view reports	Administrator views employee attendance report	System to display report successfully with a download option.	Worked as expected.

System Requirements

For the system to perform as expected, the following hardware and software requirements had to be in place.

1. Hardware Requirements.

Hardware	System Requirements (Minimum)
Processor	Intel Core i5, 2.6 GHz
Memory	4 GB RAM
Disk Space	300 GB
USB Port	2.0
AI IP Camera	5MP, optimal detection 5m

2. Software Requirements.

Software	System Requirements (Minimum)
Operating System	Microsoft Windows 7 or higher/ Mac OS 10.10 or higher
Microsoft Visual Studio	2008
Browser	Chrome, Internet Explorer, Mozilla Firefox

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

CONCLUSION, FUTURE WORK, AND IMPROVEMENTS

In the Artificial Intelligence field, Computer Vision is one of the most interesting and challenging tasks. Computer Vision acts as a bridge between Computer Software and visualizations allowing computer software to understand and learn about the visualizations in the *surroundings*. OpenCV is an open-source library. It is supported by different programming languages such as R, Python, etc. It runs undoubtedly on most platforms such as Windows, Linux, and macOS. This entire project has been developed from the requirements to a complete system alongside evaluation and testing. The system developed has achieved its aim and objectives. However, though some challenges were encountered during implementation, they were addressed and implemented, and also, future work and strategies on how to improve the system are further in this section.

Future Work and Improvement

a) A login time-out functionality would be implemented on the system for security purposes and alteration of data by unauthorized individuals.

b) The system will be deployed as a standalone which could be used by other organizations. This will now be done using the MATLAB App builder.

c) The system can have cloud-based functionality so that it can be hosted on the cloud for easy access anywhere at any time.

d) Data confidentiality is very important. So, data will be updated accordingly every year. Each employee will have the right to be informed about the use of their faces for a face recognition attendance system. This must be in line with the government laws on ethical issues and data protection laws and rights. The employees will have to consent to the images used for attendance.

REFERENCES

Bernette, C. (2019, March). *economictimes.indiatimes.com.* Retrieved from economictimes.indiatimes.com: https://economictimes.indiatimes.com/definition/authentication

Bledsoe, W. (1965). Retrieved from en.wikipedia.org: https://en.wikipedia.org/wiki/Facial_recognition_system

Burt, C. (2018, September 24). Retrieved from www.biometricupdate.com: https://www.biometricupdate.com/201809/kenyan-police-launch-facial-recognition-on-urban-cctv-network

Castrillon. (2011). Retrieved from www.nform.nu.: http://inform.nu/Articles/Vol3/v3n1p01-07.pdf

Divyarajsinh N. Parmar1, B. B. (2013). *Face Recognition Methods & Applications.* www.ijcta.com.

Jones, V. (2017, March 19). Retrieved from http://www.computerscijournal.org/: http://www.computerscijournal.org/vol10no1/comparison-of-viola-jones-and-kanade-lucas-tomasi-face-detection-algorithms/

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E

M. Gowtham, U. M. (2021). *Automatic Attendance Scheme Using Face Recognition System.* International Journal of Advanced Research in Computer Science Engineering and Information Technology.

Merriam-Webstar. (2019, September 17). *www.merriam-webster.com.* Retrieved from www.merriam-webster.com: https://www.merriam-webster.com/dictionary/attendance

Nehru and Padmavathi. (2017). *eeexplore.ieee.org.* Retrieved from www.eeexplore.ieee.org: https://ieeexplore.ieee.org/document/8014571

Petrock, V. (2019). *Biometric Marketing 2019.* USA.

Sinha, P. (2010). *Face recognition by computers and humans.* IEEE Computer Society.

Techopedia.com. (2019). Retrieved from www.techopedia.com/: https://www.techopedia.com/definition/32071/facial-recognition

Wang, e. a. (2015). *Modular Real-Time Face Detection System.*

Article DOI: 10.52589/AJSTE-HYVTCZ9E
DOI URL: https://doi.org/10.52589/AJSTE-HYVTCZ9E