Vol 5 No 2 (2023) E-ISSN: 2682-860X

International Journal on Robotics, Automation and Sciences

Development of AI-Enabled Contactless Visitor Access Monitoring System

Whei Chung Yuen, Gin Chong Lee*, and Hock Kheng Sim

***Abstract* - This research focuses on developing an AI-enabled Contactless Visitors Access Monitoring System. The monitoring system integrated a facial recognition system with a real-time database. Visitors registered themselves through an online registration form. This research developed and compared two different facial recognition systems. The first facial recognition system integrated the dlib model with the face recognition library, while the second integrated the FaceNet model with the Haar Cascade Classifier. Twenty facial images were collected. This research found out that the facial recognition system with FaceNet has higher accuracy of 82% while the has 76% of accuracy. The value of EER (Equal Error Rate) obtained for FaceNet is at 51% with an allowed threshold of 0.52. This research found that the accuracy of the facial recognition system could be affected by different conditions, such as the visitors' facial features, the distance between the camera and the face, and the illumination condition of the test environment. The number of images does not affect the speed and the accuracy of the facial recognition system in this research due to the small number of images.**

Keywords— Contactless, Access Monitoring, Face Recognition, Artificial Intelligence, Biometric.

I. INTRODUCTION

Facial recognition systems are widely used biometric technology that is more complex than fingerprint scanners. They rely on the unique features and structures of a person's face and are highly accurate, secure, and fast compared to other authentication methods. These systems find applications in high-security areas and have gained popularity in devices like phones and laptops, with Apple's FaceID being a prominent example.

The development of facial recognition technology dates back to the 1960s when semi-automatic programming was used to input facial characteristics manually [1]. Over the years, advancements such as image manipulation and eigenfaces based on PCA have improved efficiency and accuracy. The introduction of deep learning in 2011 has further propelled facial recognition technology by utilizing artificial neural networks for data analysis [2].

This study focuses on deep learning-based facial recognition systems for contactless visitor access monitoring. These systems allow computers or AI to independently analyze and make decisions, reducing

*Corresponding author. Email: ginchong.lee@uow.edu.my, ORCID: 0009-0006-7664-5851

Whei Chung Yuen is with Department of Engineering, School of Engineering, Computing and Built Environment, UOW Malaysia KDU Penang University College, 32, Jalan Anson, 10400 Georgetown, Pulau Pinang. (e-mail: 0202217@student.uow.edu.my).

Gin Chong Lee is with Department of Engineering, School of Engineering, Computing and Built Environment, UOW Malaysia KDU Penang University College, 32, Jalan Anson, 10400 Georgetown, Pulau Pinang. (e-mail: ginchong.lee@uow.edu.my).

Hock Kheng Sim is with Department of Engineering, School of Engineering, Computing and Built Environment, UOW Malaysia KDU Penang University College, 32, Jalan Anson, 10400 Georgetown, Pulau Pinang. (e-mail: hockkheng.sim@uow.edu.my).

International Journal on Robotics, Automation and Sciences (2023) 5,2:1-13

https://doi.org/10.33093/ijoras.2023.5.2.1

Manuscript received: 12 May 2023 | Revised: 31st July 2023 | Accepted: 1st September 2023 | Published: 30 September 2023

Published by MMU PRESS. URL: http://journals.mmupress.com/ijoras

physical contact and the spread of diseases like Covid-19. While some contactless access devices exist, many still rely on methods like RFID or fingerprints that can transmit bacteria. Implementing a comprehensive contactless visitor access monitoring system would mitigate these risks and prevent the transmission of bacteria and viruses.

A. Research Aim

This research aims to develop an AI-Enabled Contactless Visitor Access Monitoring System. This system is fully automated, and a facial recognition system will authorize registered personnel. The facial recognition system will be implemented with a deep learning algorithm. An online form will be implemented to allow visitors to register themselves. This monitoring system will utilize the real-time database to store the registered personnel's data and images. At the same time, this research would study and analyze two different facial recognition algorithms. The objectives of the project are summarized as below:

(a) To develop an AI-Enabled Contactless Visitor Access Monitoring System
(b) To implement deep learning facial recognition into the monitoring system
(c) To study the accuracy of 2 facial recognition algorithms

B. Project Scope

This research is divided into three main parts. Firstly, an online registration form using Jotform is created for visitors to register themselves. They provide their personal details and facial image, with their consent obtained for data usage. An admin verifies the details before storing them in the database.

Secondly, two facial recognition systems are studied. The first system combines the OpenCV library with the dlib model and is integrated into the monitoring system which is the Facial Recognition Library. The second system combines the FaceNet model with the Haar Cascade Classifier and serves as a comparison for accuracy. A confusion matrix is chosen to evaluate both systems.

Lastly, Firebase is chosen as the real-time database for this research. The collected data is stored in the database, and when authorized personnel successfully access the system, the database is updated with the total number of registered personnel and the access time.

C. Problem Statement

In the transition from a pandemic to an endemic phase, it is crucial to continue taking measures to prevent the spread of Covid-19. While mask-wearing and social distancing may no longer be mandated, it is still important to implement prudent measures for safety. In Malaysia, some condominiums and commercial buildings have security guards who collect visitor information without practicing social distancing or wearing masks, increasing the risk of transmission. To address this, a contactless visitor access monitoring system should be implemented, eliminating the need for physical interaction. However, many existing systems rely on fingerprints for identification, which requires individuals to touch the device, potentially spreading bacteria and viruses if not properly sanitized. These pathogens can be resistant to medication and may lead to the development of immunological conditions. It is advisable to minimize touching such devices or surfaces to reduce the risk of contamination.

Human involvement in security and monitoring systems introduces the possibility of errors and mistakes. In Malaysia, receptionists work fixed hours while security employees work long shifts, often around 12 hours. Fatigue and decreased alertness during night shifts can compromise building security and allow unauthorized individuals access to restricted areas. Implementing an automated visitor tracking system can help mitigate these issues. Computers can operate continuously without breaks, ensuring constant surveillance and minimizing risks to tenant safety. By reducing reliance on human presence, the automated system offers enhanced security and peace of mind.In some condominiums and commercial buildings in Malaysia, security guards physically approach visitors to collect their information, often without practicing social distancing or wearing masks. This increases the risk of Covid-19 transmission. To address this, a contactless visitor access monitoring system should be implemented, eliminating the need for physical interaction with staff member.

II. Literature Review

A. Registration Platform

(a) Websites

HTML, which stands for Hypertext Markup Language, is a markup language used to create and design websites. It is not a programming language as it lacks conditional statements and iterative looping structures. HTML can be easily used with a web browser and text editor, without the need for an interpreter or IDE. It is commonly used by professionals to create websites.

To make a website accessible on the internet, hosting and a domain are required. Free hosting options with limited features are available, but creating a local host can be challenging and less secure. Local servers need to be constantly powered on for the website to remain online. However, finding an available domain and ensuring security can be difficult.

Virtual hosting is a preferable option as it offers better security and eliminates the need for server configuration. However, data collected on a virtual hosting site may be shared with a third-party hosting provider. There are free hosting website builders like Wix.com, Weebly, and GoDaddy, but they have limited options unless paid for.

Additionally, connecting a website created with these builders to an external database can be problematic due to differences in programming languages and potential syncing errors [5].

(b) Online Form Builder

For the registration platform in this research, an online form is a practical solution. There are several free online form builders available. Google Forms is a popular choice, offering basic form creation and data storage in Google Sheets. It has a simple user interface that is easy to understand.

Jotform is another online form builder that allows the creation of up to 5 forms. It has a better UI design compared to Google Forms and offers various form templates. Data collected is stored in the Jotform database, and users can view and analyze the data using charts and insights. Jotform also allows users to connect to external databases like Google Sheets without requiring third-party software [6].

Typeform is well-known for its interactive user experience. It offers conversational forms, where questions are displayed card by card, creating a more engaging experience compared to traditional survey-style forms. Typeform is a great choice for those seeking a good and elegant form interface.

B. Facial Recognition

(a) OpenCV's Facial Recognition Library

The OpenCV Facial Recognition Library is widely used in Python programming. It provides various functionalities such as face detection, alignment, encoding, and recognition. The feature extraction process involves encoding the input image and extracting critical facial features or landmarks. The image is then encoded into an array of vectors using a classifier. These vectors are compared with facial images in a database, and the image with the lowest value indicates a match.

In traditional facial recognition algorithms, Eigenfaces are used to match similar images. Each image is represented as a grayscale and converted into a kxk bitmap of pixels [4]. The algorithm reshapes the image into a grayscale representation and calculates the Euclidean distance between the Eigenfaces of the captured image and those in the database using the k-NN classifier [4]. Similar images will have the lowest Euclidean distance value. However, this algorithm has low accuracy, especially when the person moves their head or changes facial expressions, as it affects the grayscale image and decreases accuracy.

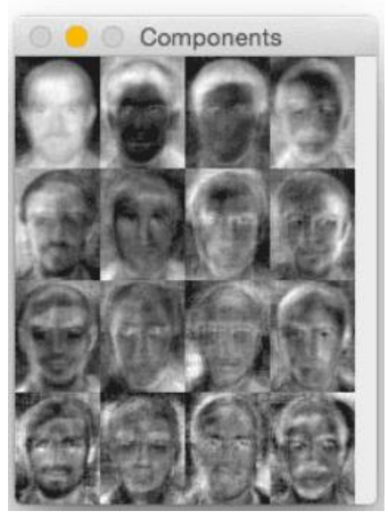

FIGURE 1. The grayscale image of eigenfaces.

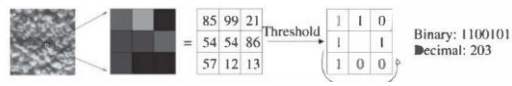

FIGURE 2. Operation of LBP [5].

Modern facial recognition systems utilize a combination of HOG (Histograms of Oriented Gradient), LBP (Local Binary Pattern), and the Haar Cascade classifier. HOG is a feature descriptor that identifies and highlights the object's edges in image processing. In facial recognition, HOG is used for feature extraction, where it extracts 25 facial landmarks that are crucial for the classification process [5]. The number of facial landmarks can be adjusted by the programmer.

The facial landmarks are obtained by dividing the images into small cells, and the gradient magnitude is calculated for each of these cells. This calculation helps the system recognize the facial edges accurately, contributing to the overall facial recognition process. It should be noted that the Haar Cascade classifier and the dlib algorithm, both included in the OpenCV Facial Recognition library, play significant roles in the system but will be discussed in detail later.

LBP (Local Binary Pattern) is another feature descriptor similar to HOG, and it is also used in the feature extraction process of facial recognition systems. LBP is known for its high efficiency and low computational cost [6].

In LBP operations, a 3x3 square neighborhood is defined around a pixel, as illustrated in Figure 2. The intensity of the center pixel is compared to the intensities of the surrounding 8 pixels. If a neighboring pixel's intensity is equal to or higher than the center pixel's intensity, it is assigned a value of 1; otherwise, it is assigned a value of 0. This process is performed for each pixel in the image, and the resulting binary values are then converted into decimal values [6].

(b) dlib Model

The dlib model is an open-source toolkit developed by Davis King and widely used in image processing and machine vision [7]. Although it is written in C++, it is compatible with various platforms and has been integrated into the Python facial recognition library.

One notable feature of the dlib model is its ability to implement different types of machine learning algorithms, including DNN (Deep Neural Networks), SVM (Support Vector Machines), and k-NN (k-Nearest Neighbors) [8]. In facial recognition, the dlib model utilizes a DNN algorithm to process images and extract a 128-dimensional vector as its output. This approach achieves high accuracy in facial recognition [3].

A study done by [9] demonstrated the effectiveness of combining the dlib model with HOG (Histograms of Oriented Gradient) for detecting driver drowsiness. The proposed system achieved an accuracy of 96.71%. [9] highlights the potential of using a combination of different models to improve the accuracy of a system.

(c) Haar Cascade

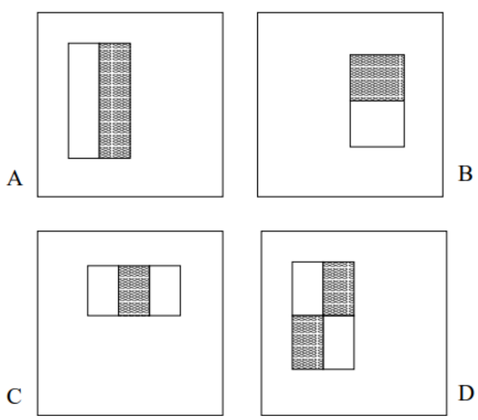

FIGURE 3. The rectangular features [10].

Haar Cascade is a feature-based object detection method developed by Paul Viola and Michael Jones in 2001 [10]. It is used to detect and classify objects in an image. The Haar Cascade algorithm requires training with a large number of positive and negative images to improve its accuracy [10].

The algorithm utilizes four types of rectangular features that correspond to the detection window, capturing image corners, lines, and edges as shown in Figure 3. These rectangular features undergo a feature extraction process, where they are rapidly computed over the integral image of the image from the top left to the bottom right. This method is faster than a pixel-based system.

During the feature extraction process, the sum of the pixels within the white area of the rectangles is subtracted from the sum of the pixels within the gray area. The resulting difference values for each pixel are stored in rows and columns. The Adaboost algorithm is used to train the weak classifier.

[10] is recommended to use a pre-trained Haar Cascade rather than training the classifier from scratch, as the training process can be time-consuming.

(d) CNN

A Simple Neural Network and a Deep Learning Neural Network differ in the number of layers they have, with the latter having more layers. This study focuses on the Deep Learning Neural Network, which will be utilized in the face recognition system. Deep Learning Neural Networks are particularly effective in handling new data and can handle large-scale image processing and vast amounts of data [1]. As a result, they are suitable for application in various fields, such as audio processing and handwriting recognition.

Deep Learning Neural Networks are a subset of machine learning algorithms and are characterized by having multiple hidden layers for information processing. They are designed for feature learning and pattern categorization. Within the realm of Deep Learning Neural Networks, there are three main classes: RNN (Recurrent Neural Networks), CNN (Convolutional Neural Networks), and DNN (Deep Neural Networks) [1].

(e) CNN Architecture

CNN is a deep learning technique used for processing data arranged in a grid-like structure, such as digital images. CNN interprets images using a binary form and organizes them into a grid of pixels. Each grid contains pixel values that provide information about the color and brightness of the pixels [11].

The functioning of CNN is similar to the human brain, with individual neurons processing information and communicating with others. The layers of a CNN are designed to process simple patterns first and then progress to more complex patterns [11].

The convolution layer is the initial hidden layer in a CNN. This layer performs the majority of the computational tasks and plays a crucial role in feature extraction. It involves performing a dot product between two matrices: the input image matrix and a smaller matrix called the kernel [11]. Despite its smaller size, the kernel is more accurate in capturing relevant features from the image.

Below is some example of CNN architecture:

- FaceNet
- YOLOnet
- VGGnet
- Alexnet

C. Database

(a) Firebase

Firebase is a platform created by Google that offers various tools for developers to create and develop applications. It functions as a Backend-as-a-Service (BaaS) and allows developers to connect their applications to cloud-based services. Firebase simplifies the development of the application's backend, providing an easy-to-understand tool [12].

Firebase is a NoSQL server, which means it uses a document-based database where each document contains key-value pairs. The focus of NoSQL is on collecting documents, and it is also referred to as a non-relational database [13].

One notable tool offered by Firebase is the real-time database. It is user-friendly and compatible with different programming languages. One key feature of the Firebase real-time database is that it allows offline updates, meaning the database can be updated even when the application is not connected to the internet and can be synchronized when online again.

(b) MySQL

MySQL is an open-source relational database management system (RDBMS) based on Structured Query Language (SQL). It stores data in tables and supports different types of data. MySQL is programmed primarily in C++ and C and offers high-security features such as host-based verification and password encryption [14].

PHP is a popular scripting language used for web development. It is free and helps in reducing the length of code, especially for extensive HTML code. PHP is easy to learn and can run on various platforms as it is platform-independent [14].

phpMyAdmin is a free third-party software used for administering MySQL. It facilitates the connection between HTML websites and MySQL databases and provides services like database management and web hosting. phpMyAdmin can be run locally on any server or host [14].

D. Related Work

In the study conducted by Hiremani et al. (2022), an AI-powered contactless face recognition technique was developed for IoT access in smart mobility. They implemented Python's dlib library for facial recognition and achieved an accuracy of 92% with 94% sensitivity. The chances of error were approximately 8%, as determined using a confusion matrix with a sample size of 77 [15].

Bong and Lee (2021) created a contactless visitor access monitoring system using the MTCNN technique with a pre-trained ResNet model for facial recognition. They obtained an accuracy of 82% using five different sets of data, as measured through a confusion matrix [16].

Teoh et al. (2021) focused on face recognition and identification using a deep learning approach. They utilized the Haar Cascade Classifier with TensorFlow for facial recognition and achieved an accuracy of 91.7%. They also investigated the impact of face-to-camera distance and lighting intensity on the system's accuracy, finding that higher lighting intensity resulted in higher accuracy [17].

Overall, these studies demonstrate the development and evaluation of various facial recognition systems, showcasing their accuracies and exploring factors that can influence their performance.

III. Research Methodology

A. Flow Chart of The Registration Process

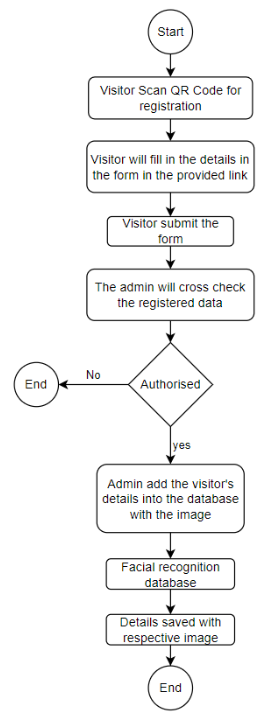

FIGURE 4. The flow chart of the registration procedure.

The registration procedure of the monitoring system begins with visitors registering through an online form. The submitted data is then cross-checked by the admin to ensure the security of the premises. If the submission is approved, the admin inserts the details into the program, and the visitor becomes registered personnel. If the submission is not approved, the admin may ask the visitor to resubmit the form or deny authorization. After inserting the details and the visitor's image, the backend process encodes the image with the visitor's information. The encoded details, along with the image, are stored in both the real-time database in Firebase and the local database, completing the process.

B. Flow Chart of The Monitoring System

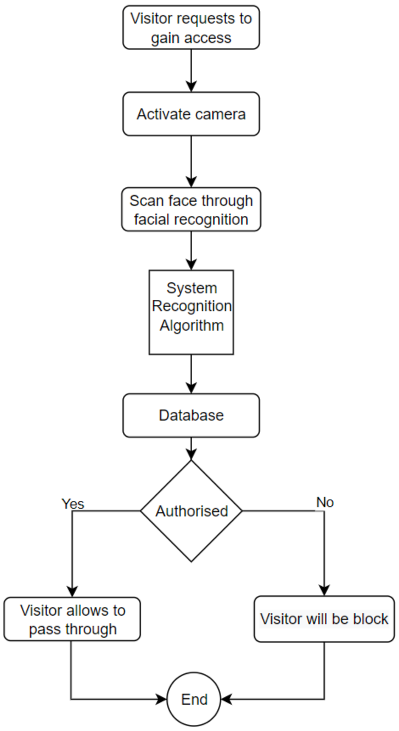

FIGURE 5. The flow chart of the monitoring system.

The monitoring system operates separately from the registration process and is deployed at authorized areas to monitor visitors. When a visitor approaches the authorized area, the system activates the camera to capture their facial image. The captured image is then processed through the system's recognition algorithm, where it undergoes encoding in the backend programming of the facial recognition system.

Once the facial image is encoded, it is compared with the database. If a match is found between the facial image and an image in the database, the system grants authorization for the registered personnel to enter the authorized sector. However, if the facial image is not recognized, the system blocks the visitor from accessing the area.

C. UML (Unified Modeling Language)

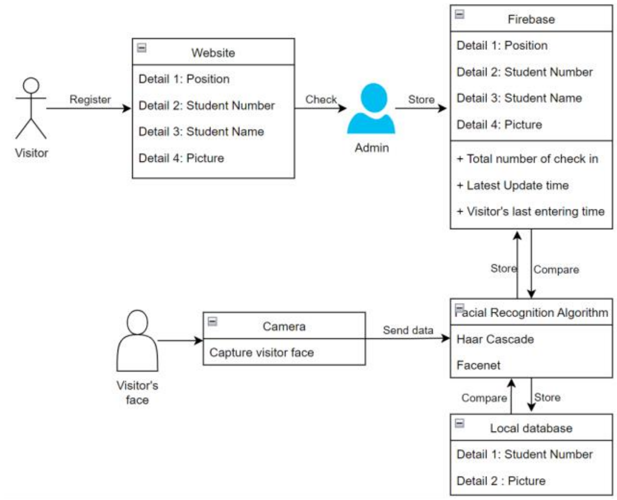

FIGURE 6. The UML of the facial recognition monitoring system.

Figure 6 presents the UML (Unified Modeling Language) diagram of the system, which helps visualize the system's design in a simplified manner. UML focuses on illustrating the interaction between different components of the system, while a flow chart provides a summary of the system's flow.

Analyzing the UML diagram, in the registration process, visitors fill in their details and submit an image through a website form. The data is then sent and stored in a Google Sheet for admin verification. Upon approval, the admin inserts the data into Firebase using program software. The facial image undergoes encoding through the facial recognition algorithm, and both the encoded image and details are stored in a local and cloud database.

In the monitoring process, the camera captures the visitor's facial image. The program sends the image data to the facial recognition algorithm for encoding. The encoded images are compared with those in the local database. If a match is found, the image details are compared and updated in Firebase. However, if no match is found, the program informs the visitor that they are not recognized and returns to standby mode. The process concludes at this point.

IV. DESIGN

A. Online Form

Jotform was selected to be the website for the registration process in this research. Jotform is a software company which offers an online platform for form creators. This online form creator platform offers a free package of up to 5 forms. Jotfrom also provides compatibility flexibility for the user. The online form can be linked and stored in the database selected by the user. For instance, Google Sheet is selected as the database for the online form. All registered details will be stored in the Google sheet and only can be accessed by the authorized admin.

The original plan of this research was to create HTML website hosted by Firebase that automatically updates and sync with cloud database upon registration. Microsoft Visual Studio was chosen as the programming platform to create the HTML website. However, there were some issues in updating and syncing data with the Firebase real-time database, Therefore, a free online form creator platform is chosen to create the registration website.

B. Facial Recognition System

(a) Face Recognition Library

The Face Recognition Library in OpenCV utilizes a combination of traditional computer vision techniques, such as Haar Cascades, LBP, HOG, and machine learning algorithms, to process images. It is commonly employed for detecting, aligning, and recognizing faces in images or video frames, offering high accuracy. However, it does not utilize Convolutional Neural Networks (CNN) commonly associated with deep learning.

This research combines the traditional face recognition library in OpenCV with the dlib model, which is a powerful tool for image processing. The dlib model incorporates deep learning algorithms, including CNN, to process images. In this research, the dlib model is used to detect and align faces captured in images. It preprocesses the captured images before passing them to the facial recognition model.

This research utilizes a pre-trained DNN (Deep Neural Network) from the dlib model, which provides a 128-dimensional feature vector for detecting and extracting features from the images. The dlib model also employs a shape predictor model to estimate the location of key facial landmarks. Each pictures of the personnel used in this research will be trained an encoder. Figure 7 illustrates the flow chart of the image encoding process to the end product in this research.

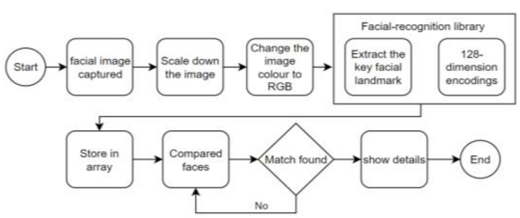

FIGURE 7. The flow chart of the facial recognition and the system.

(b) FaceNet

This research employs the FaceNet algorithm, which incorporates the Haar Cascade classifier, for deep learning-based facial recognition. The Haar Cascade classifier is utilized for face detection, while FaceNet handles the face recognition part of the system. The images are pre-processed and passed through the Haar Cascade classifier and FaceNet. FaceNet uses a CNN architecture and Euclidean distance to determine image similarity. The image encoding process involves resizing, color conversion, and comparison with images in the database. The system captures images from a webcam and matches them with database images, without storing them in the database itself as shown in Figure 8.

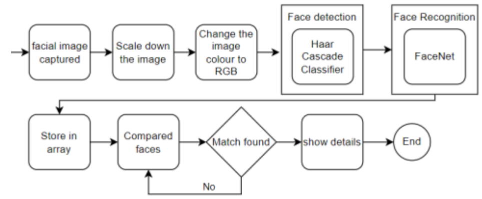

FIGURE 8. The flow chart of the facial recognition and the system.

C. Database

In this research, the chosen real-time database is Firebase, an online platform offering various services and integration for applications. Firebase is preferred due to its compatibility with Python, ease of integration, and availability of a free package for storing large amounts of data. Initially, MySQL or SQLite databases were considered, but encountered challenges in updating and syncing data with the MySQL server. While MySQL offers enhanced security through local host connections and username/password authentication. Firebase was opted as the main database due to its ease of use and compatibility with the research requirements.

D. Software

There are only 1 primary software used in this research. Python is the selected programming language to be used for facial recognition system in this research. Therefore, PyCharm was selected as the IDE for this research. PyCharm is a very powerful IDE for Python development environment and it is suitable for image processing.

E. Hardware

There are only 1 hardware used in this research. Laptop webcam is the selected camera to be used to capture images. The highest resolution of the webcam is 720p with size of 16:9 and 30hz refresh rate.

F. Design Constraints

(a) Lack of Images Training

A constraint has been identified in the facial recognition system related to insufficient training on images. However, this research aims to develop a system for a specific scenario where images are provided shortly before applying the facial recognition model to visitors. This scenario is designed to simulate a situation with a large influx of visitors, where there is limited time for training the system. This research intends to investigate and analyze the accuracy of the facial recognition system when it lacks training on the specific images.

(b) Performance of GPU and CPU

Another constraint was identified which related to the GPU power of the laptop used. The facial recognition algorithm employed in this research requires a high-performance GPU to effectively analyze and identify facial features in images with high accuracy. A powerful GPU enables faster processing and analysis of each image, which is particularly important for facial recognition algorithms utilizing CNN architecture. CNN algorithms rely on extensive data processing and matrix operations to achieve accurate face recognition. GPUs outperform CPUs in terms of accuracy and detection speed, making them essential for efficient facial recognition tasks [18].

V. Risk Assessment

A. Data Privacy and Security

Data privacy and security are significant concerns in today's world, with recent events raising questions about the protection of user information. In this research, the researcher collected facial images and details of 20 students who provided consent. The data is stored in both a local database on the researcher's laptop and a real-time cloud database using Firebase and Google Sheet. However, there are inherent risks of data breaches and unauthorized access by malicious parties. Even though the data is stored locally, there is a possibility of hackers compromising the researcher's laptop and gaining access to the data.

Using a cloud-based database like Firebase increases the risk further, access rules could be modified, potentially allowing unauthorized reading or writing of the data. While Google Sheet provides a higher level of security and privacy, there have been instances of data breaches and security incidents involving Google services, such as the exposure of 52.5 million user data from Google+ in 2018. Any unauthorized exposure of user data poses a significant threat to their safety and privacy.

B. False Positives and Negatives

The inconsistency of a facial recognition system can lead to false positives and false negatives, posing a safety risk. Factors contributing to inconsistency include illumination conditions, variations in facial features among registered individuals, distance between the camera and the person, and the presence of twins among registered users. Inaccuracies in the system can potentially allow unauthorized individuals to gain access to restricted areas, jeopardizing the safety of users and valuable assets. To mitigate this risk, it is important to achieve high accuracy in facial recognition. One approach is to establish a threshold value for the system's accuracy. If the recognition accuracy falls below the threshold, the user will not be authorized, and access will only be granted when the accuracy meets or exceeds the threshold value.

VI. Experiment Plan

This research aims to determine the accuracy of two facial recognition systems through experiments. The accuracy of using FaceNet algorithm is evaluated using a confusion matrix, which allows calculation of False Acceptance Rate (FAR) as in (1) and False Rejection Rate (FFR) as in (2).

$$FAR = \frac{n(FP)}{n(FP+TN)} \tag{1}$$

$$FRR = \frac{n(FN)}{n(FN+TP)} \tag{2}$$

Where,

FP = False Positive

FN = False Negative

FP = True Positive

TN = True Negative

n = Total Number

The values obtained from the confusion matrix are used to plot a graph and determine the Equal Error Rate (EER) as shown in Figure 9. The threshold value for the systems varied from 0.1 to 0.9. The threshold value indicates acceptable range of percentage accuracy. For instance, if threshold value is equal to 1 which is accuracy of 100%, this indicates that all percentage accuracy will be accepted by the system. If the threshold value is 0.2 which is 20%, this indicates that only percentage accuracy that is higher than 80% will be accepted by the system. The threshold value was set manually by the researcher of this research.

The experiment involves collecting twenty facial images, with ten categorized as registered personnel and the rest as unregistered personnel. Each student's image is tested five times, resulting in a sample size of 100.

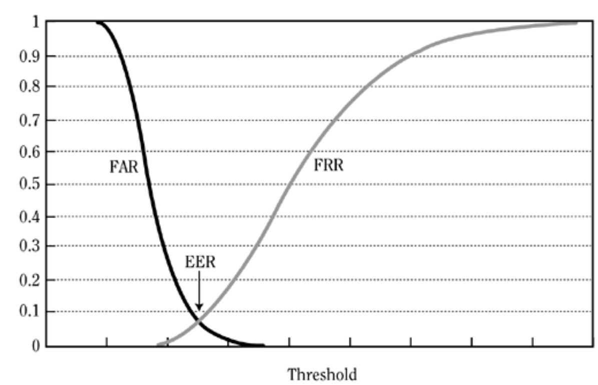

FIGURE 9. The flow chart of the facial recognition and the system.

The experiment also includes testing the Facial Recognition Library's and FaceNet algorithm's accuracy in correctly identifying registered personnel and their details. Both algorithims are given the same number of registered personnel images (10 people), with each

E-ISSN: 2682-860X

image tested five times, resulting in a sample size of fifty. The accuracy of both facial recognition systems are determined by the total number of correctly recognised personnel divided by the total number of sample as in (3).

$$\text{Accuracy} = \frac{\text{Total Number of Correctly Recognised Personnel}}{\text{Total Number of Sample}} \times 100\% \tag{3}$$

The accuracy of various facial features was determined through the previous experiment. The data collected from these experiments, including the facial feature and its corresponding accuracy, will be presented using a graph. The accuracy of each face feature is calculated by dividing the total sample size by the number of correct results.

As for the experiment regarding the impact of the number of registrations on the speed and accuracy of the authentication system, it will not be conducted. This decision is based on the findings which indicate that the number of registrations has an insignificant effect on the system.

VII. Results and Discussion

A. User Interface (UI)

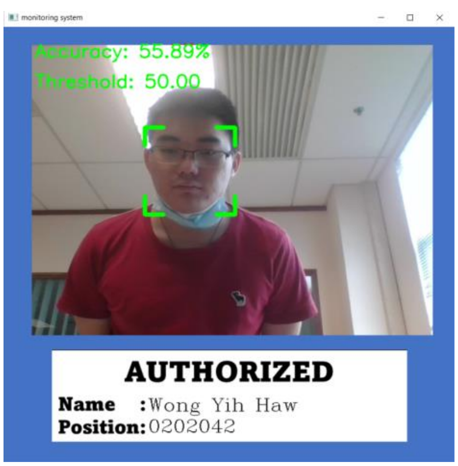

FIGURE 10. The UI shown when the registered personnel is authorized.

FIGURE 11. The UI shown when the visitor is not registered.

When the registered personnel is authorized, the UI will display the name of the registered personnel with the position. If the visitor is not registered, the UI will display "Unregistered" to notify the visitor that his or her details are not found in the system. The accuracy percentage and threshold value will be displayed. If the accuracy percentage fell under the threshold value, the system will recognise it as "unregistered" due to low accuracy.

B. The Graph of Percentage against Threshold Value of FaceNet Algorithm

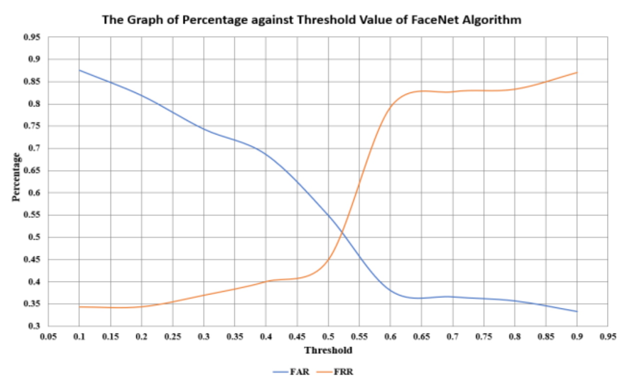

FIGURE 12. The graph of percentage against Threshold Value of FaceNet Algorithm.

It is found that the ERR of the FaceNet algorithm is 51% with allow threshold of 0.52.

The graph presented in Figure 12 illustrates the relationship between the percentage and threshold value for the FaceNet algorithm. It closely resembles the ideal graph depicted in Figure 9, where the number of False Rejection Rates (FRR) is inversely proportional to the number of False Acceptance Rates (FAR). The point where these lines intersect is called the Equal Error Rate (EER), indicating an equal percentage of false acceptances and false rejections.

The threshold value is set within the range of 0.1 to 0.9, with higher values indicating that correctly identified faces with a lower percentage will be considered successful. Each threshold value represents a specific accuracy percentage. For example, if the threshold is set to 1, any accuracy percentage will be considered successful. On the other hand, if the threshold is 0.4 (40%), only an accuracy greater than 60% will be considered a success.

The FRR is determined by the number of False Negatives (FN), while the FAR is based on the number of False Positives (FP) within the sample. The study involves ten registered individuals tested five times each, resulting in a sample size of 50 to assess True Positives (TP) and True Negatives (TN). Additionally, ten random unregistered individuals were chosen as a comparison group, also tested five times each, resulting in another sample size of 50. An example of FP secenario where an unregistered individual is labeled as "unknown" since their face image is not present in the database.

The study determined that the EER occurs at a threshold of 0.52 with an associated accuracy of 51%. This finding suggests that the system can be considered stable and accurate. However, providing more images to the system could lower the EER and improve overall accuracy. Notably, some of the obtained results do not exhibit a steep gradient as seen in the ideal case graph (Figure 10). Particularly between threshold values of 0.6 and 0.7, the FRR and FAR lines appear nearly flat, indicating similarity in their values. This inconsistency may be attributed to factors such as insufficient training images, environmental conditions, or image quality.

During the experiment, it was observed that the accuracy of identified faces using the data pool was generally low, ranging between 30% and 50%. The highest accuracy achieved was only 59%.

Due to the relatively low accuracy percentages (between 50% and 60%), this experiment was not conducted in the Face Recognition Library. If a threshold is set above 60%, all registered individuals would be categorized as False (TF), making it impossible to plot the graph and determine the EER. Figure 13 provides an example of registered individuals with accuracy percentages below the 0.6 threshold, resulting in their classification as TF.

FIGURE 13. Registerd personnel with low accuracy percentage that does not fall under the threshold value of 0.6.

C. Accuracy of the Facial Recognition System

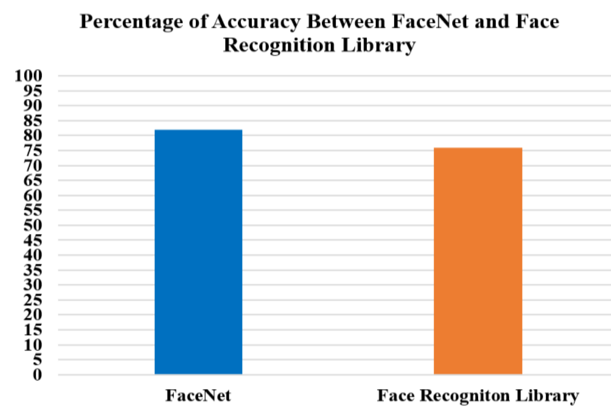

FIGURE 14. The difference between both system in term of accuracy.

It is found that FaceNet algorithm recorded a higher accuracy compared to the Face Recognition Library. The percentage of accuracy for FaceNet algorithm is 82% while the Face Recognition Library recorded with 76% which is 6% lower than FaceNet algorithm.

The accuracy of two facial recognition systems was evaluated using a different approach. Instead of using confusion matrix, this experiment measured the time it took for the systems to correctly identify registered individuals with the correct details. The accuracy of both facial recognition systems are determined by the total number of correcly recognised personnel divided by the total number of sample.

To facilitate comparison, both systems were provided with the same set of registered personnel images, consisting of 10 individuals. Each individual was tested five times, resulting in a sample size of 50.

The bar chart in Figure 14 indicates that the FaceNet algorithm achieved a higher accuracy of 82% compared to the Facial Recognition Library, which achieved an accuracy of 76%. The difference between the two algorithms was found to be 6%.

FaceNet utilizes a CNN architecture for image processing, while the Facial Recognition Library compares images to determine the one with the highest similarity. FaceNet employs multiple layers to identify the most similar image. In this study, FaceNet was combined with the Haar Cascade algorithm to enhance the accuracy percentage.

FaceNet demonstrated a higher accuracy percentage compared to the Facial Recognition Library as anticipated.

D. Accuracy of the Facial Recognition System on Different Facial Features

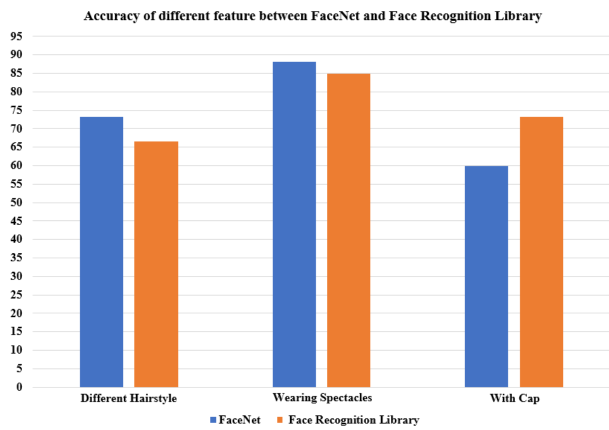

FIGURE 15. The accuracy of different feature between FaceNet and face recognition library.

Both the FaceNet algorithm and the Facial Recognition Library were evaluated using the same number of samples for each facial feature. According to Figure 15, it is observed that the FaceNet algorithm generally achieves a slightly higher accuracy percentage compared to the Facial Recognition Library. The FaceNet algorithm performs better in terms of accuracy for different hairstyles and people wearing spectacles. However, the Facial Recognition Library exhibits higher accuracy when it comes to individuals wearing caps.

Regarding individuals with different hairstyles than their submitted images, the Facial Recognition Library achieves an accuracy of 66.67%, while the FaceNet algorithm records an accuracy of 73.3%. People wearing spectacles have the highest accuracy percentage, with the FaceNet algorithm achieving 85% accuracy. On the other hand, individuals wearing caps have the lowest accuracy percentage, with the Facial Recognition Library recording 73.3% accuracy, which is higher than the FaceNet algorithm.

This study concludes that different facial features have an impact on the accuracy of the facial recognition system. Three types of facial features were categorized: different hairstyles, individuals wearing spectacles, and individuals wearing caps. These findings were extracted from previous experiments.

Within the database, there were three individuals with different hairstyles compared to their submitted images. Figure 15 shows that the FaceNet algorithm achieved an accuracy percentage of 73.3%, while the Facial Recognition Library recorded a lower accuracy percentage of 66.66%. In one case, there was a registered individual with longer hair than the submitted image, and both systems failed to recognize him in the first two tests. The researcher of this research suspects that the hair covered important facial features processed by the system, making it unable to find the most similar image in the database initially.

Furthermore, out of the total 20 people in the database, 17 of them wore glasses. Both algorithms demonstrated high accuracy in recognizing individuals wearing spectacles. The FaceNet algorithm achieved the highest accuracy percentage at 88.23%, while the Facial Recognition Library reached 85% accuracy. This indicates that wearing glasses does not significantly affect the accuracy of the facial recognition algorithms.

In addition, there were a few registered individuals who wore caps during the experiment, totaling three individuals. Surprisingly, the FaceNet algorithm recorded a lower accuracy percentage (60%) compared to the Facial Recognition Library (73.3%). However, both accuracy percentages are still considered low. This experiment highlights that different face features, such as wearing a cap, can impact the accuracy of the algorithms. Nonetheless, FaceNet achieved the highest accuracy percentage in terms of image similarities when one of the individuals wore a cap. These inconsistencies suggest that the system may have some limitations or inconsistencies in certain scenarios.

E. Effect of The Distance Between The Face and Camera on Accuracy of The Facial Recognition System.

In this research, the researcher found that the distance between the face and camera would affect the accuracy of the facial recognition system. It is found that when the registered personnel move his body backward, the facial recognition system could not recognize him. The system could not identify any faces located in the frame. Therefore, the facial recognition system displays “standby” on UI to indicate that the system is ready to identify faces. Therefore, the researcher concludes that if anyone stand further more than 60 cm away from the camera, the facial recognition will not identify any faces.

F. Limitations

Several limitations contribute to the inconsistency of the facial recognition system. The first limitation is the surrounding illumination conditions. In environments with dark lighting, the accuracy of the system is significantly lower compared to well-lit environments.

Insufficient lighting results in poorly illuminated faces, making it difficult for the algorithm to detect and process facial features. Additionally, when individuals face bright lights from behind, the camera captures darker faces, leading to detection difficulties.

The quality of the camera used also impacts the accuracy of the algorithm. Low-resolution cameras produce unclear and noisy images, which can lead to inaccuracies during image processing.

Furthermore, obstacles between the camera and the individual pose another limitation. In some cases, individuals deliberately covered their foreheads or mouths to test the system's accuracy. However, both algorithms failed to recognize them because they could not detect a face or find a match in the database. It's important to note that the algorithms used in this study did not include mask detection or detection through mask algorithms. Therefore, registered personnel were required to remove their masks for facial recognition purposes.

VIII. Conclusion

In conclusion, this research successfully developed an AI-Enabled Contactless Visitor Access Monitoring System with a facial recognition system and real-time database. A deep-learning facial recognition algorithm were implemented using the Face Recognition Library and compared it to other deep learning approaches such as FaceNet, VGGnet, and YOLOnet. The second facial recognition system, which integrated the FaceNet model with the Haar Cascade Classifier, achieved an EER value of 51% with a threshold of 0.52. However, the EER value of the Facial Recognition Library cannot be obtained due to accuracy percentages falling between 50% to 60%.

The accuracy comparison between the two facial recognition systems showed that the FaceNet algorithm achieved higher accuracy (82%) compared to the Face Recognition Library (76%). Experiments on different facial features, including caps, hairstyles, and spectacles were conducted in this research. It is found that the FaceNet model performed better in identifying individuals with different hairstyles and wearing spectacles, while the Facial Recognition Library was more accurate in recognizing individuals wearing caps.

Furthermore, the researcher identified the distance between the camera and the face as a manipulated variable. It was found that when registered personnel stood more than 60 cm away from the camera, the facial recognition system struggled to identify them accurately. Therefore, it is recommended for registered personnel to stand closer to the camera to improve accuracy.

Overall, this research successfully achieved its aims and objectives, developing a functional monitoring system and providing insights into the accuracy and variables affecting facial recognition systems.

Acknowledgment

The authors would like to acknowledge the support and encouragement of members from UOW Malaysia KDU Penang University college for the countless hours spent for discussing new technologies, debating various topics, and helping each other to overcome academic challenges.

Funding Statement

No funding was funded for this paper.

Author Contributions

Gin Chong Lee: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation, Project Administration, Writing – Review & Editing;

Gin Chong Lee: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation, Project Administration, Writing – Review & Editing;

Hock Kheng Sim: Conceptualization, Data Curation, Methodology, Validation, Writing – Original Draft Preparation, Project Administration, Writing – Review & Editing.

Conflict of Interests

No conflict of interests were disclosed.

Ethics Statements

Our publication ethics follow The Committee of Publication Ethics (COPE) guideline. https://publicationethics.org/

References

[1] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, “Past, present, and future of face recognition: A review,” Electronics, vol. 9, no. 8, p. 1188, 2020. DOI: https://doi.org/10.3390/electronics9081188

[2] M. Çarıkçı and F. Özen, “A face recognition system based on eigenfaces method,” Procedia Technology, vol. 1, pp. 118–123, 2012. DOI: https://doi.org/10.1016/j.protcy.2012.02.023

[3] B. R. Singh, S. K. Singh, and R. K. Singh, “Bacteria on fingerprint scanners of biometric attendance machines,” Microbiology Research International, vol. 7, no. 4, pp. 31–39, 2019. URL: https://www.netjournals.org/pdf/MRI/2019/4/19-027.pdf

[4] A. M. Jagtap, A. S. Kapse, and S. S. Giri, “A study of LBPH, eigenface, fisherface and haar-like features for face recognition using OpenCV,” in Proceedings of the 2019 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, India, 2019, pp. 719–724. DOI: https://doi.org/10.1109/ISS1.2019.8907965

[5] M. Ghorbani, A. T. Targhi, and M. M. Dehshibi, “HOG and LBP: Towards a robust face recognition system,” in Proceedings of the 2015 Tenth International Conference on Digital Information Management (ICDIM), Jeju, South Korea, 2015, pp. 148–152. DOI: https://doi.org/10.1109/ICDIM.2015.7381860

[6] H. Ren, “A comprehensive study on robustness of HOG and LBP towards image distortions,” Journal of Physics: Conference

Series, vol. 1325, no. 1, p. 012012, 2019. DOI: https://doi.org/10.1088/1742-6596/1325/1/012012
[7] D. E. King, "Dlib-ml: A machine learning toolkit," Journal of Machine Learning Research, vol. 10, pp. 1755–1758, 2009. URL: https://www.jmlr.org/papers/volume10/king09a/king09a.pdf (accessed: 28 July, 2023)
[8] L. Chen, Y. Zhang, and J. Zhao, "Driver fatigue detection based on facial key points and LSTM," Security and Communication Networks, vol. 2021, pp. 1–9, 2021. DOI: https://doi.org/10.1155/2021/5383573
[9] S. Mohanty, S. R. Mishra, and B. K. Patnaik, "Design of real-time drowsiness detection system using Dlib," in Proceedings of the 2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), Bangalore, India, 2019, pp. 1–4. DOI: https://doi.org/10.1109/WIECON-ECE48653.2019.9019910
[10] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 2001, pp. 511–518. DOI: https://doi.org/10.1109/CVPR.2001.990517
[11] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," Neural Computation, vol. 29, no. 9, pp. 2352–2449, 2017. DOI: https://doi.org/10.1162/neco_a_00990
[12] D. Vaz, M. Matos, J. Silva, and L. Veiga, "MIRES: Intrusion recovery for applications based on backend-as-a-service," IEEE Transactions on Cloud Computing, vol. 11, no. 2, pp. 2011–2027, 2023. DOI: https://doi.org/10.1109/TCC.2022.3178982
[13] D. Guo and E. Onstein, "State-of-the-art geospatial information processing in NoSQL databases," ISPRS International Journal of Geo-Information, vol. 9, no. 5, p. 331, 2020. DOI: https://doi.org/10.3390/ijgi9050331
[14] E. Abdullah, M. A. Hannan, A. Mohamed, and A. Hussain, "Development of real-time energy monitoring system and data log using NodeMCU ESP8266 and MySQL database," International Journal of Power Electronics and Drive Systems, vol. 10, no. 1, pp. 245–252, 2019. URL: https://www.researchgate.net/publication/356509140_Development_of_Real-Time_Energy_Monitoring_System_and_Data_Log_Using_NodeMCU_ESP_8266_and_MYSQL_Database (Accessed:27 july 2023)
[15] N. Hiremani, S. S. Manvi, and P. B. Hiremani, "Artificial intelligence-powered contactless face recognition technique for internet of things access for smart mobility," Wireless Communications and Mobile Computing, vol. 2022, pp. 1–11, 2022. DOI: https://doi.org/10.1155/2022/8750840
[16] Y.S. Bong and G.C. Lee, "A contactless visitor access monitoring system," International Journal on Robotics, Automation and Sciences, vol. 3, pp. 33–41, 2021. DOI: https://doi.org/10.33093/ijoras.2021.3.6
[17] K. H. Teoh, S. S. Yuhaniz, and M. H. Husni, "Face recognition and identification using deep learning approach," Journal of Physics: Conference Series, vol. 1755, no. 1, p. 012006, 2021. DOI: https://doi.org/10.1088/1742-6596/1755/1/012006
[18] V. Jain and D. Patel, "A GPU based implementation of robust face detection system," Procedia Computer Science, vol. 87, pp. 156–163, 2016. DOI: https://doi.org/10.1016/j.procs.2016.05.142