

Implementation of Real-Time Face Recognition for Secure Weapon Storage Access Control

Anisa Anisa¹, Giva Andriana Mutiara², and Muhammad Rizqy Alfarisi³

Department of Applied Science, Telkom University, Bandung, Indonesia

Corresponding author: Giva Andriana Mutiara (e-mail: givamz@telkomuniversity.ac.id), **Author(s) Email:** Anisa (e-mail: nissaanisa171@gmail.com), Muhammad Rizqy Alfarisi (e-mail: mrizkyalfarisi@telkomuniversity.ac.id)

Abstract The security of weapon storage warehouses is a critical concern that requires an access control system with exceptionally high reliability, particularly in minimizing false acceptance, where unauthorized individuals are incorrectly granted access. In high-risk facilities, even a single false acceptance incident can lead to serious security consequences. Conventional systems based on physical keys or access cards present limitations, including risks of loss, duplication, and access forgery. Therefore, a biometric-based solution is necessary to enhance identification accuracy and strengthen overall security. This study aims to design and implement a reliable, high-security facial-recognition-based access control system for weapon storage facilities. The proposed system integrates a Multi-task Cascaded Convolutional Neural Network (MTCNN) for face detection, FaceNet for feature extraction, and a Support Vector Machine (SVM) for identity classification. The system is implemented as a standalone application on an edge computing device (mini PC) integrated with an electronic door lock. All detection and decision-making processes are performed locally without reliance on cloud services. System evaluation was conducted under various testing scenarios, including variations in lighting intensity, camera distance, facial attributes, and unregistered face testing. Experimental results show that the system achieved an accuracy of 96.25%. A precision of 100% indicates that no unauthorized access was granted. The recall reached 92.50%, reflecting a small proportion of rejected authorized users. The F1-score of 96.11% demonstrates balanced performance. The False Acceptance Rate was 0%, confirming complete prevention of illegal access. The False Rejection Rate was 7.50%, which remains acceptable in high-risk security environments. The system consistently rejected all unregistered faces and operated in real time with an average door unlocking response time of approximately 1.3 seconds. In conclusion, the proposed system provides reliable recognition performance with a strong emphasis on preventing false acceptance. These findings indicate its suitability for enhancing security in high-risk weapon storage facilities.

Keywords Face recognition, MTCNN, FaceNet, Support Vector Machine, access control security.

1. Introduction

The rapid advancement of information technology has had a significant impact on various aspects of life, including the field of security [1]. In the current digital era, threats to both physical and data security have become increasingly complex, necessitating a security system capable of providing effective, intelligent, and responsive protection. In a military context, security aspects are paramount as they pertain directly to personnel safety and the protection of a nation's strategic assets [2]. One of the critical facilities requiring a high level of security is the weapons storage warehouse, which stores various combat equipment and ammunition [3]. In practice, most weapon security systems in several Indonesian military institutions still rely on conventional methods, such as manual keys, padlocks, or access cards [4]. While these methods provide a degree of physical security, they possess

inherent limitations, such as the loss of physical keys, the risk of access duplication, and the lack of automated user activity logs [4]. These conditions increase the probability of unauthorized access and complicate both the auditing process and the early detection of security breaches [5]. With the continuous advancement of biometric technology, face recognition systems have emerged as a potential solution to enhance access security in high-risk facilities [6]. This system leverages individuals' unique facial characteristics, captured by cameras and processed with image processing algorithms and machine learning, to ensure access is granted exclusively to registered and verified personnel. In addition to its non-contact nature, face recognition operates autonomously and in real-time, thereby reducing reliance on human intervention and minimizing errors resulting from human oversight [6], [7].

In high-security environments such as weapon storage facilities, access control systems must operate under strict operational constraints, including time-critical access, minimal physical interaction, and high reliability. Unlike fingerprint-based systems, which require direct contact and may be affected by the use of gloves or environmental conditions, face recognition provides a non-contact and rapid authentication mechanism. Compared to iris recognition, which often requires precise positioning and specialized sensors, face recognition offers greater flexibility and ease of deployment using standard camera devices. Furthermore, unlike token-based methods such as access cards or RFID, facial biometrics are inherently non-transferable, significantly reducing the risk of credential sharing or unauthorized use. These characteristics make face recognition particularly suitable for military applications, where both security and operational efficiency are critical.

Various previous studies have developed face recognition systems using a range of detection, feature extraction, and classification methods [8]. Nevertheless, a number of existing approaches still face limitations, such as high computational overhead that complicates implementation on small-scale hardware, and a reliance on cloud infrastructure, which potentially introduces latency and additional security risks in the form of biometric data interception [9]. These conditions pose a distinct challenge to implementing face recognition systems in military operational environments, which demand high reliability, rapid response times, and stringent data security. Furthermore, most existing research remains focused solely on enhancing recognition accuracy without considering comprehensive system security aspects, particularly the risk of false acceptance in high-security environments such as military weapon storage warehouses [10]. Despite these advancements, a clear research gap remains in the development of face recognition-based access control systems specifically designed for high-security environments. Most existing studies primarily emphasize recognition accuracy without explicitly addressing the critical requirement of minimizing false acceptance in security-sensitive applications. In addition, several approaches still rely on cloud-based architectures or high-performance computing resources, which may introduce latency and increase the risk of biometric data exposure. Furthermore, limited attention has been given to evaluating system robustness under realistic operational conditions, such as variations in illumination, user distance, and facial attributes. Therefore, there is a need for a face recognition system that integrates computational efficiency, edge-based deployment, and a security-

oriented decision strategy to ensure reliable and safe access control in weapon storage environments.

This gap becomes particularly critical in weapon storage security, where the consequences of system failure are severe. In such environments, false acceptance errors carry far more severe consequences than false rejections, as they potentially grant unauthorized individuals access to firearms and ammunition. Consequently, this study is designed to function not only as a biometric identification system but also as a critical decision-making mechanism for high-security environments. In such settings, access reliability and the mitigation of false acceptance risks are prioritized over the mere attainment of high recognition accuracy. To address these challenges, this study proposes a weapon storage warehouse access control system based on face recognition, integrating the Multi-task Cascaded Convolutional Neural Network (MTCNN) for face detection, FaceNet for feature extraction, and Support Vector Machine (SVM) as the identity classifier. The selection of SVM as the final classifier aims to establish more stable and well-defined decision boundaries within the 128-dimensional embedding space generated by FaceNet, making it suitable for high-risk security environments. Unlike several previous systems that rely on cloud-based processing, the system proposed in this study is designed to operate as a standalone application on an edge computing device, thereby reducing latency and mitigating the risk of biometric data breaches.

While MTCNN, FaceNet, and SVM have been widely used in face recognition systems, most existing studies primarily focus on improving recognition accuracy without explicitly addressing the security requirements of high-risk environments. In contrast, this study emphasizes a security-oriented system design, where the integration of MTCNN, FaceNet, and SVM is specifically optimized to minimize the risk of false acceptance ($FAR = 0\%$), which is critical in weapon storage applications. Furthermore, the proposed system is implemented on an edge computing platform, enabling real-time processing without reliance on cloud infrastructure, thereby reducing latency and potential data exposure risks. Unlike conventional approaches that primarily emphasize recognition accuracy, this study adopts a security-oriented decision strategy. A high decision threshold is applied to enforce strict access control, even at the cost of increased false rejection. Therefore, the novelty of this study lies not only in the combination of algorithms but also in the system-level optimization and security-driven design tailored for high-risk operational environments.

The objective of this study is to design and implement a secure, reliable, and real-time face recognition-based access control system for weapon

storage warehouses, taking into account recognition accuracy, system response time, and access security risks. The entire face recognition process is executed locally on a mini PC integrated with a relay module and an electric drop bolt lock. Consequently, the authentication process can be performed in real time without dependency on external network connections, while simultaneously enhancing system resilience against cybersecurity threats. The developed system is also designed to withstand real-world operational conditions, such as variations in illumination intensity, capture distance, and user face attributes. The primary contributions of this study are summarized as follows:

1. Design of a face recognition-based weapon storage warehouse access control system that operates autonomously within an edge computing environment, eliminating dependency on cloud infrastructure to enhance data security and reduce latency.
2. Integration of MTCNN, FaceNet, and SVM within a security-oriented framework, explicitly optimized to achieve a False Acceptance Rate (FAR) of 0%, prioritizing strict access control over conventional accuracy-focused approaches.
3. Comprehensive evaluation of system performance under diverse real-world operational conditions, including variations in illumination, camera distance, facial attributes, and unregistered user scenarios, with a specific focus on robustness and security performance (FAR and FRR).
4. Implementation of a digital access logging mechanism to support accountability and auditability in high-security weapon storage environments.

Based on the conducted design and evaluation, the proposed system demonstrates a practical and reliable solution for enhancing access control security in high-risk facilities. Furthermore, the edge computing-based approach and decision-making process oriented toward minimizing false acceptance are intended to serve as a reference for developing more secure and adaptive biometric security systems in the future.

The remainder of this paper is organized as follows: First, a literature review regarding face recognition methods and biometric-based security systems is presented. Second, the methodology and the proposed system architecture are described, including the data preparation process and the model processing workflow. Third, the experimental results and the evaluation of system performance under various operational conditions are presented, followed by a detailed discussion. Finally, the conclusion and suggestions for future research directions are provided.

II. Related Works

Research pertaining to face recognition systems has been extensively developed and applied across various domains, such as attendance systems, smart home security, and access control for strategic facilities [6], [11]. Various approaches have been employed, ranging from classical feature-based methods to deep learning-based approaches, with the objective of enhancing the accuracy and reliability of face recognition systems [12]. Early approaches to face recognition generally relied on traditional feature extraction methods, such as Local Binary Pattern Histogram (LBPH) [13], [14]. Although this method is computationally efficient and relatively easy to implement, various studies have shown that its performance decreases significantly under variations in illumination, pose, and facial expression [15]. These limitations render classical feature-based methods less reliable for deployment in security systems operating in high-risk environments that demand high precision and performance stability.

With the rapid advancement of deep learning technologies, Convolutional Neural Network (CNN)-based approaches have been widely adopted to improve face recognition accuracy [16]. Several studies have implemented end-to-end CNN architectures for face detection, feature extraction, and identity classification, demonstrating significant performance improvements compared to classical methods [17]. Nevertheless, pure CNN-based approaches generally require substantial computational resources and relatively high inference time, making them less suitable for deployment on resource-constrained devices such as embedded systems or mini PCs, particularly in real-time scenarios [18]. As an alternative for the face detection stage, the Multi-task Cascaded Convolutional Neural Network (MTCNN) was introduced as a more robust face detector capable of handling variations in illumination and viewing angles, thereby enhancing the stability of face recognition systems under real-world operational conditions [19], [20], [21].

To address these limitations, several studies have adopted a hybrid approach by separating the feature extraction and identity classification stages [29]. Models such as FaceNet are employed to generate discriminative 128-dimensional facial embedding vectors, which are subsequently classified using machine learning algorithms such as Support Vector Machine (SVM) [22]. This approach is capable of maintaining high recognition accuracy while reducing computational complexity compared to end-to-end CNN architectures, and it has demonstrated good stability on medium-scale datasets with balanced class distributions.

From a system architecture perspective, several studies still rely on cloud-based processing for face recognition and biometric data storage. Although this approach offers scalability, dependence on cloud infrastructure may introduce latency and pose security and privacy risks due to the transmission of biometric data over external networks [23], [24]. Consequently, more recent studies have shifted toward edge or fog computing approaches, in which face recognition is performed locally to enhance real-time responsiveness

and minimize exposure of sensitive data [25], [26]. This approach becomes particularly critical in high-security environments, such as military weapon storage warehouses, which require a high level of reliability, extremely low error tolerance, and strict access accountability. A comprehensive comparison of the position of this study relative to existing security systems is summarized in Table 1. This comparison highlights the differences in methodology, implementation platform, and application domain.

Table 1. Comparison of Related Face Recognition-Based Security Systems

Ref	Face Detection	Feature Extraction	Classification	Platform	Application Domain	Main Limitation
[4]	N/A	RFID Tag	RFID Reader	Server	Weapon rack security	No biometric authentication
[5]	Haar Cascade	CNN	Softmax	PC	Attendance system	Not designed for critical security
[6]	CNN	CNN	Softmax	IoT Device	Smart door	Network latency and privacy risk
[9]	N/A	Biometric features	Cloud-based classifier	Server	Biometric identification	Risk of data interception
[55]	Haar Cascade	Handcrafted features	KNN	PC	Attendance system	Limited real-time performance
[56]	CNN	CNN	Softmax	PC	Face recognition	High computational cost
This Study	MTCNN	FaceNet	SVM	Embedded (Mini PC)	Weapon warehouse security	Low latency, high security, no cloud dependency

Based on the comparison presented in Table 1, previous studies can be classified into two main approaches: non-biometric and biometric systems. Non-biometric security systems, such as those based on Radio Frequency Identification (RFID), are relatively easy to implement and involve low system complexity [27]. However, this approach has fundamental limitations in ensuring the validity of user identity, as it relies on physical media that are vulnerable to loss, unauthorized lending, or misuse [27], [28]. These limitations make non-biometric systems less suitable for deployment in high-security environments [29]. In contrast, biometric-based security systems, particularly those employing face recognition, offer a stronger level of authentication by leveraging the unique physiological characteristics of users [30]. Nevertheless, the literature indicates that face recognition systems still face several challenges, particularly in terms of computational efficiency, performance stability under real-world operational conditions, and biometric data security risks [31], [32]. These challenges become increasingly significant when such systems are deployed in critical security

environments that demand extremely low error rates and fast, consistent system responses.

In addition to authentication methods and system architecture, the selection of the final classification algorithm also has a significant impact on the efficiency and stability of face recognition systems [12], [33], particularly when deployed on edge devices with limited computational resources [9]. The k-Nearest Neighbor (KNN) method, although simple and easy to implement, is known to be sensitive to the curse of dimensionality when applied to high-dimensional embedding vectors [34]. This condition leads to increased inference latency and reduced system efficiency as the number of registered identities grows. Meanwhile, Random Forest offers strong non-linear classification capabilities; however, it is prone to overfitting when applied to datasets with a limited number of subjects, which may reduce the model's generalization performance on unseen data [35], [36]. This limitation makes Random Forest less optimal for face recognition scenarios involving a relatively small number of identities but requiring high levels of accuracy and consistency [37]. In contrast, SVM offers the advantage of maximizing the separation margin

between classes in high-dimensional embedding spaces [38], [39]. This characteristic enables SVM to deliver stable and consistent performance on medium-scale datasets with relatively balanced class distributions, while maintaining low inference latency. Therefore, SVM is considered more suitable for embedding-based face recognition scenarios,

particularly when combined with feature extraction models that generate discriminative facial representations, such as FaceNet [22]. A comparison of the characteristics of several commonly used classification algorithms for facial embedding mapping is presented in Table 2. This comparison highlights their performance differences.

Table 2. Comparison of Classification Algorithms for Face Embedding Mapping

Algorithm	High-Dimensional Stability	Medium Dataset Performance	Inference Latency	Primary Limitation	Ref
KNN	Poor	Moderate	High	Sensitive to the curse of dimensionality	[57], [58], [59]
Random Forest	Moderate	Moderate	Moderate	Prone to overfitting on a limited number of subjects	[60], [61]
Softmax (DNN)	High	Low	Low	Requires a large-scale identity dataset	[62], [63]
SVM	Excellence	High	Very Low	Requires appropriate kernel and parameter selection	[38], [39], [48]

Based on the literature review and comparative analysis, a research gap remains in the development of face recognition-based access control systems that not only focus on improving algorithmic accuracy but also consider computational efficiency, performance stability on edge devices, and system suitability for critical security environments. Therefore, this study proposes the integration of MTCNN as a face detector robust to environmental variations, FaceNet as a low-dimensional embedding-based feature extractor, and SVM as the identity classifier within an edge computing system integrated with a door-locking mechanism [19], [22], [40], [41]. This approach is designed to achieve an optimal balance among recognition accuracy, response speed, and high security in weapon storage warehouse protection systems.

Although numerous previous studies have demonstrated promising performance in the implementation of face recognition systems for access control and security applications, most of these approaches primarily focus on improving overall recognition accuracy. Specific security aspects, particularly the risk of false acceptance in high-security environments such as weapon storage warehouses, have not been extensively addressed. Moreover, several studies still rely on cloud-based infrastructures or high-performance computing devices, which may introduce latency and increase the risk of biometric data exposure. Therefore, a research gap remains in the development of face recognition systems capable of operating independently (standalone) on edge computing devices, with a primary focus on minimizing

the risk of false acceptance to support the protection of critical military facilities.

III. Methodology

A. Proposed System

The system proposed in this study is a face recognition-based access control system designed to operate in a standalone (offline) manner and implemented in a critical security environment, specifically a weapon storage warehouse. All face recognition processes, decision-making mechanisms, and access logging are performed locally on an edge device without dependence on external network connections or cloud services. This approach aims to minimize latency, enhance system reliability, and reduce the risk of biometric data leakage.

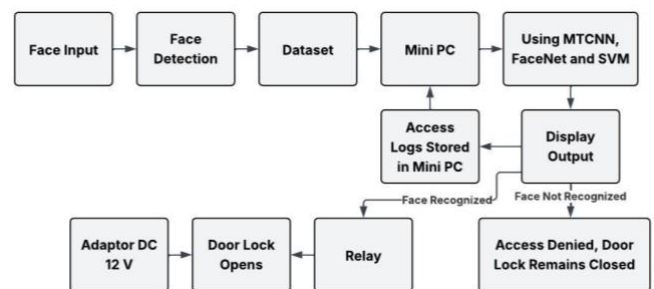


Fig. 1. System architecture and hardware

The overall system architecture is illustrated in Fig. 1. The system consists of a USB camera as the facial image acquisition device, a mini-PC as the main

processing unit, and a relay module connected to an electric drop bolt lock serving as the door-locking actuator. When a user stands in front of the camera, the captured facial image is processed in real time by the mini PC using a combination of MTCNN, FaceNet, and Support Vector Machine (SVM) algorithms. All access attempts, whether granted or denied, are recorded and stored locally as access logs for security auditing purposes. The workflow of the face recognition and access control system is illustrated in Fig. 2. The process begins with facial image acquisition through the camera, followed by face detection using MTCNN. The detected

time face recognition on CPU-based devices. For image acquisition, a Logitech C922x HD Pro webcam with a resolution of 1080p was used to capture facial images during the authentication process.

Table 3. Distribution of Dataset

Category	Total Personnel	Images per Subject	Total Images
Training Set (80%)	25	192	4,800
Testing Set (20%)	25	48	1,200
Total	25	240	6,000

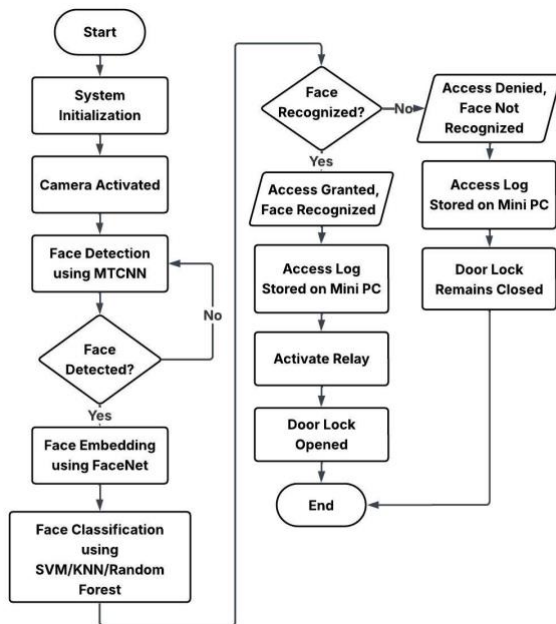


Fig. 2. Operational flowchart access control

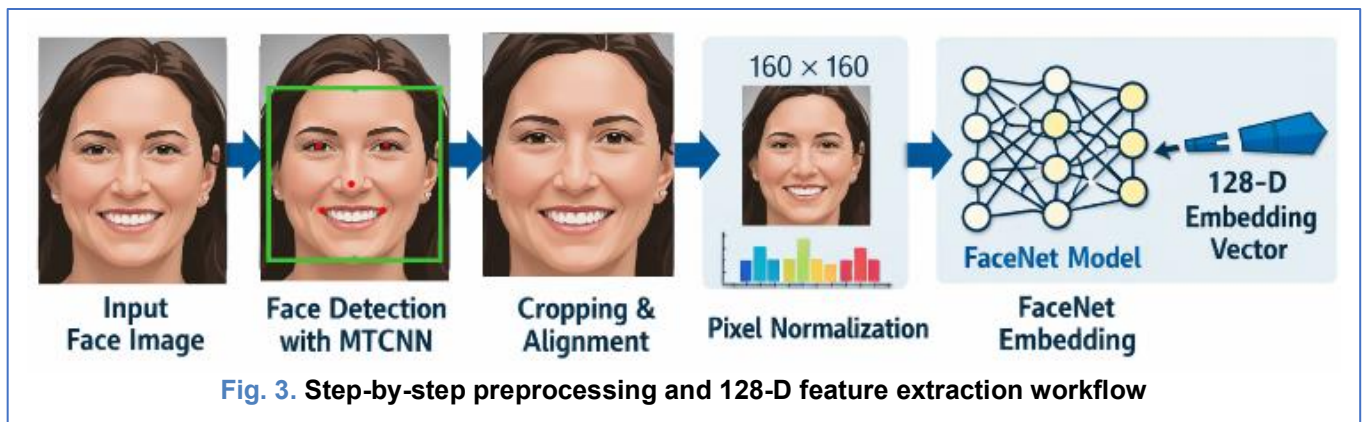
face is then transformed into a 128-dimensional embedding vector using FaceNet and classified using SVM to determine the user's identity. If the face is recognized as a registered personnel, the system activates the relay to unlock the door; otherwise, unrecognized faces are denied access and the door. The integration of face detection, embedding-based feature extraction, and classification on an edge device enables the system to provide real-time responses with a level of security appropriate for protecting a weapons storage warehouse. The hardware configuration used to implement the proposed system is described as follows. The proposed system was implemented on a mini PC (GMKtec M6 Ultra) equipped with an AMD Ryzen 7 7640HS processor, 32 GB DDR5 RAM, and a 512 GB SSD. The system operates without a dedicated GPU, ensuring compatibility with edge computing environments and demonstrating the feasibility of real-

B. Dataset Collection

This section describes the methodology for primary dataset acquisition and processing, which forms the foundation for the development of the weapon storage warehouse access control system. The dataset used in this study consists of primary data collected directly to support the development and evaluation of the face recognition-based access control system. The use of primary data is considered crucial to ensure that the model achieves a high level of reliability when deployed in real military operational scenarios. Military environments possess unique characteristics such as controlled access distance (0.5–1.5 m), variable illumination (10–150 lux), and operational facial attributes, which are often not represented in publicly available datasets. Therefore, on-site data collection provides improved validity and applicability for real-world deployment.

This primary dataset was independently collected from 25 registered personnel with authorized access. To minimize intra-class variation and prevent identity bias, data acquisition was conducted across three recording sessions performed on different days to introduce natural appearance variations. For each individual, 240 facial images were collected, resulting in a total dataset of 6,000 images. This medium-scale dataset, with a balanced class distribution across identities, is considered well-suited for evaluating the robustness of embedding-based classification algorithms in maintaining optimal inter-class separability. Detailed information regarding the dataset distribution used in this study is summarized in Table 3.

The image acquisition process was conducted using a USB camera directly integrated with the edge computing unit in the form of a mini PC serving as the main processing device. Image capture was performed in an indoor environment designed to simulate the operational conditions of a weapon storage warehouse, without the use of artificial backgrounds or additional



visual manipulation. The testing area had an approximate dimension of 4×5 meters with a consistent background and lighting conditions categorized into bright (100–150 lux) and dim (10–50 lux) environments.

The camera was positioned at a fixed height of 1.5 meters with a horizontal orientation and oriented horizontally to capture frontal facial images during the authentication process. The distance between the user and the camera was maintained between 0.5 and 1.5 meters to ensure consistent image quality. To ensure reproducibility, several variables were controlled during the testing process. The background was kept consistent, with no significant visual distractions, and user movement was limited to natural standing positions in front of the camera. Variations were intentionally introduced only in specific test scenarios, such as changes in illumination, facial attributes, and user distance, to evaluate system robustness under realistic operational conditions.

During the training dataset collection phase, facial images were captured without additional attributes to ensure clean and consistent data quality. Variations introduced at this stage included differences in lighting conditions (bright and dim) as well as pose variations, including frontal position (0°) and slight head rotation ($\pm 15^\circ$), to enhance the model's generalization capability against changes in viewing angles. Meanwhile, system robustness testing was conducted separately by incorporating various facial attributes, such as the use of military helmets, caps, glasses, and more extreme facial pose variations. This scenario was designed to evaluate the system's resilience under realistic operational conditions that may be encountered during field deployment. In addition, the evaluation also considered variations in the distance between the user and the camera to simulate a realistic authentication process. Prior to the training and evaluation stages, all facial images underwent a series of systematic preprocessing steps to ensure consistent feature representation. Face detection was performed using MTCNN to accurately obtain facial bounding box coordinates and key

landmark points. Based on the detection results, the facial images were subsequently subjected to cropping and alignment processes to standardize their spatial configuration. The preprocessed images were then resized to 160×160 pixels to match the input requirements of the FaceNet model, and pixel values were normalized to maintain stability in the data distribution. Each facial image was subsequently transformed by the FaceNet model into a 128-dimensional embedding vector that is robust to variations in illumination and facial pose. The overall preprocessing and feature extraction stages are summarized in Fig. 3.

These embedding representations were used as input for the SVM algorithm. The entire dataset was divided into 80% training data and 20% testing data, with the split performed at the identity level to prevent data leakage and ensure an objective performance evaluation. This strategy was designed to represent a realistic authentication scenario, in which the system is expected to reliably recognize registered identities under varying operational conditions. The collected data include variations in facial appearance, lighting conditions, and pose to represent realistic operational scenarios. However, detailed demographic attributes such as age and gender distribution were not explicitly categorized, as the primary objective of this study is to evaluate system performance under operational conditions rather than to conduct demographic analysis. To evaluate the system's ability to reject unauthorized access, additional test scenarios were conducted using unregistered individuals. These unregistered samples were used exclusively during the testing phase to simulate real-world conditions where unknown users attempt to gain access. This approach ensures that the evaluation includes both positive (registered) and negative (unregistered) classes, thereby providing a more objective assessment of system performance. Furthermore, the dataset was constructed with a balanced number of samples per registered individual, and the training-testing split was performed at the identity level to prevent data leakage

and ensure unbiased evaluation. This strategy allows the system to be evaluated under conditions that closely resemble real-world deployment scenarios.

C. System Algorithm Stages

All algorithmic stages were designed to ensure that the system operates accurately, in real time, and securely on edge computing devices with limited computational resources. To achieve this objective, the system algorithm is divided into two main phases: the training phase (offline training) and the recognition phase (online inference). This separation aims to reduce computational load during field operation while enhancing system stability and response speed.

1. Training Phase (Offline Training)

The training phase represents the initial stage conducted prior to system deployment in the operational environment. All training processes were executed locally without involving any external network connections. The facial image dataset used in this phase was derived from primary data of authorized personnel and had undergone a curation process to ensure image quality and balanced class distribution. In the initial stage, each facial image was processed using the Multi-task Cascaded Convolutional Neural Network (MTCNN) algorithm to detect and localize the facial region. MTCNN was selected as the face detection method due to its consistent performance under varying illumination conditions and changes in facial orientation [19], [20]. In addition to generating facial bounding boxes, MTCNN also detects five key facial landmarks, namely the two eyes, the nose, and the corners of the mouth, which are utilized for the facial alignment process.

MTCNN consists of three cascaded convolutional neural networks: the Proposal Network (P-Net), the Refine Network (R-Net), and the Output Network (O-Net) [19]. The P-Net operates on an input size of 12×12 pixels with a stride of 2, enabling rapid generation of face candidate regions across multiple image scales. These candidate regions are subsequently filtered using non-maximum suppression (NMS) to reduce overlap among bounding boxes. The resulting face candidates are then processed by R-Net, which uses a 24×24 -pixel input to eliminate false positives and refine bounding-box coordinates. In the final stage, O-Net processes facial regions using an input size of 48×48 pixels to produce high-precision facial bounding boxes while simultaneously detecting five key facial landmarks, namely the two eyes, nose, and mouth corners [20]. These landmark points are utilized in the

where B_1 and B_2 represent the bounding boxes generated during face detection, while the numerator indicates the intersection area between the two bounding boxes, and the denominator represents the

alignment process, which plays a crucial role in improving the spatial consistency of facial images and the quality of embeddings generated by FaceNet.

The detailed architecture of each MTCNN stage is illustrated in Fig. 4, which presents the cascaded convolutional neural network architecture of P-Net, R-Net, and O-Net, including convolution layers, pooling operations, and fully connected layers. The P-Net receives an input image of $12 \times 12 \times 3$ pixels. The first convolution layer applies 10 filters with a kernel size of 3×3 , followed by a 2×2 max pooling layer with a stride of 2. The second convolution layer uses 16 filters, followed by a third convolution layer with 32 filters. The output is then processed by classification and bounding box regression layers. The R-Net processes $24 \times 24 \times 3$ input images and consists of three convolution layers followed by a fully connected layer to refine face candidates and reduce false positives. The O-Net receives $48 \times 48 \times 3$ input images and consists of four convolution layers followed by a fully connected layer that produces final bounding boxes and five facial landmark outputs. The convolution operation used in the MTCNN architecture can be mathematically expressed as shown in Eq. (1) [41], where the input image is convolved with kernel weights to extract spatial features:

$$F(x, y) = \sum_i \sum_j I(x+i, y+j)K(i, j) \quad (1)$$

where $I(x, y)$ represents the input image and $K(i, j)$ denotes the convolution kernel. After convolution, a non-linear activation function is applied using the Rectified Linear Unit (ReLU) to introduce non-linearity into the network, as defined in Eq. (2) [42]:

$$f(x) = \max(0, x) \quad (2)$$

To reduce spatial dimensionality and improve computational efficiency, max pooling is applied as shown in Eq. (3) [42]:

$$P(x, y) = \max_{i, j} F(x+i, y+j) \quad (3)$$

Bounding box regression is performed to refine detected face regions, as formulated in Eq. (4) [42].

$$B' = B + \Delta B \quad (4)$$

To eliminate overlapping detections, Non-Maximum Suppression (NMS) is applied based on Intersection over Union (IoU), as defined in Eq. (5) [42]:

$$IoU = \frac{Area(B_1 \cap B_2)}{Area(B_1 \cup B_2)} \quad (5)$$

union area of both bounding boxes. The Intersection over Union (IoU) value ranges from 0 to 1, where higher values indicate greater overlap between bounding boxes. In the MTCNN framework, IoU is used to

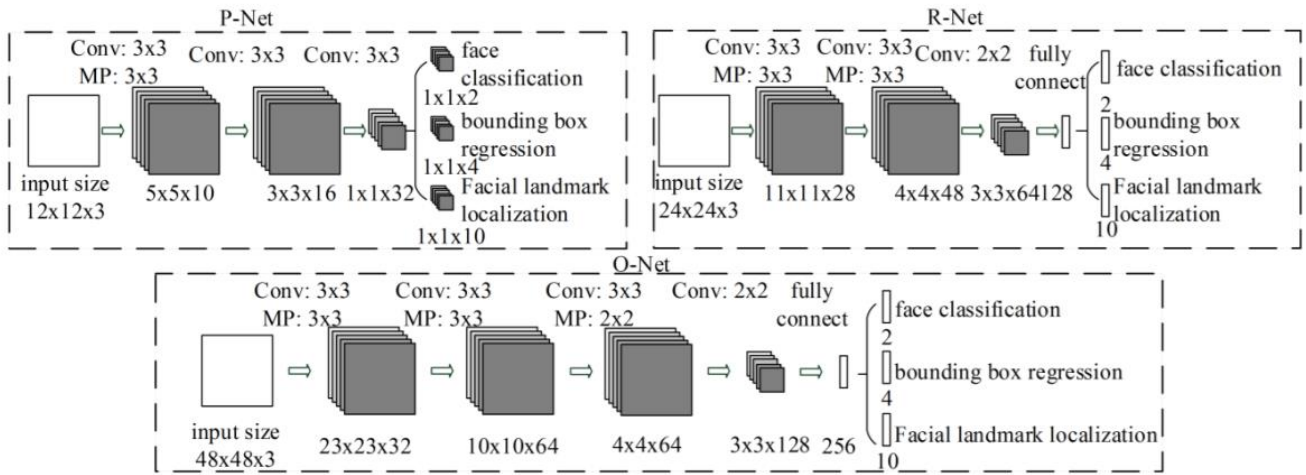


Fig. 4. The cascaded convolutional neural network architecture of MTCNN, consisting of P-Net, R-Net, and O-Net

eliminate redundant detections by retaining bounding boxes with higher confidence scores and suppressing overlapping regions during the Non-Maximum Suppression (NMS) process. The detected facial images subsequently underwent a processing stage that included cropping, facial alignment, resizing to 160×160 pixels, and pixel value normalization. These steps were performed to ensure input uniformity prior to feature extraction [43]. Feature extraction was performed using the FaceNet model to generate facial representations in the form of 128-dimensional embedding vectors. The selection of the FaceNet model is based on its ability to generate highly discriminative and compact 128-dimensional embedding vectors, which are effective for distinguishing identities while maintaining computational efficiency [44]. This makes FaceNet particularly suitable for real-time face recognition systems deployed on edge devices. The FaceNet model maps facial images into a Euclidean embedding space, where the distance between embedding vectors represents the degree of identity similarity [45]. The distance between two embedding vectors is computed using the Euclidean distance, as expressed in Eq. (6) [46] below:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (6)$$

where x and y denote the 128-dimensional facial embedding vectors generated by FaceNet. To improve feature discrimination, FaceNet employs a triplet loss function that minimizes intra-class distance and maximizes inter-class separation, as formulated in Eq. (7) [46]:

$$L = \max(d(a, p) - d(a, n) + \alpha, 0) \quad (7)$$

where a represents the anchor embedding, p denotes the positive embedding from the same identity, n represents the negative embedding from a different identity, and α is the margin parameter. To maintain feature scale consistency, L2 normalization is applied to the embedding vector as shown in Eq. (8) [46]:

$$x' = \frac{x}{\|x\|_2} \quad (8)$$

where x represents the original embedding vector and x' denotes the normalized embedding. Similarity between facial embeddings can also be evaluated using cosine similarity, as defined in Eq. (9) [46]:

$$\cos(\theta) = \frac{x \cdot y}{\|x\| \|y\|} \quad (9)$$

where x and y denote embedding vectors and θ represents the angle between them. The FaceNet model maps facial images into a 128-dimensional embedding space, represented in Eq. (10) [46]:

$$f(x) \in \mathbb{R}^{128} \quad (10)$$

where $f(x)$ denotes the embedding function. The FaceNet model employed in this study is a pre-trained model based on the Inception-ResNet architecture, enabling it to produce highly discriminative feature representations without requiring complex retraining of the convolutional network. This characteristic makes it computationally efficient and suitable for deployment on edge devices [47]. The resulting embedding vectors were subsequently normalized using L2 normalization to maintain feature scale consistency. These normalized embeddings were then used as training data for the SVM model, selected due to its capability to maximize the separation margin between classes in high-dimensional feature spaces, thereby providing stable classification performance on medium-scale datasets. The trained SVM model, along with the

associated identity label information, was stored locally for use during the face recognition phase. SVM was selected as the classification model due to its capability to construct optimal decision boundaries in high-dimensional feature spaces and its relatively low computational complexity compared to deep learning-based classifiers [48]. This makes it well-suited for real-time implementation on edge computing devices without requiring GPU acceleration.

In this study, an SVM algorithm with a linear kernel was employed as the face identity classification model. The linear kernel was selected due to its high computational efficiency and strong generalization capability in high-dimensional feature spaces, such as the 128-dimensional facial embedding vectors generated by FaceNet. The choice of this kernel also considered the system's requirement to operate in real time on edge computing devices with limited computational resources [49]. The regularization parameter C was set to 1.0, providing a balance between maximizing the separation margin between classes and minimizing classification errors. The parameter value was determined through preliminary experiments on the training data to ensure model performance stability and to prevent overfitting [50]. In addition, probability estimation was enabled in the SVM model to generate confidence scores. In this study, the confidence score represents the likelihood that the input facial embedding belongs to a particular identity class, based on the classification output of the SVM. The score is normalized to the range 0 to 1, where higher values indicate greater similarity and higher confidence in the predicted identity. These confidence scores are subsequently used as the basis for access decision-making during the recognition phase. The training protocol of the SVM model was designed to ensure stable classification performance while maintaining computational efficiency. The model was trained using embedding vectors generated from the FaceNet model, with a linear kernel and a regularization parameter (C) set to 1.0. The parameter value was determined through preliminary experiments on the training dataset to achieve a balance between classification accuracy and generalization capability. To construct an optimal decision boundary, the Support Vector Machine aims to determine a hyperplane that maximizes the separation margin between classes, as represented in Eq. (11) [46]:

$$w \cdot x + b = 0 \quad (11)$$

where w represents the weight vector, x denotes the input feature vector, and b is the bias term. The optimization objective of SVM is to maximize the margin between classes, which can be formulated as shown in Eq. (12) [38]:

$$\min_w \frac{1}{2} |w|^2 \quad (12)$$

This formulation ensures optimal class separation and improved classification performance. To determine the final access decision, a threshold-based rule is applied as formulated in Eq. (13) [38]:

$$\begin{aligned} y &= 1 & \text{if } score \geq 0.90 \\ y &= 0 & \text{if } score < 0.90 \end{aligned} \quad (13)$$

Cross-validation was not applied in this study, as the evaluation strategy was based on an identity-level data split. Specifically, the dataset was divided into 80% training data and 20% testing data, where all images belonging to a particular individual were exclusively assigned to either the training or testing set. This approach prevents data leakage and ensures that the model is evaluated on unseen identities, providing a more realistic assessment of system performance. This training strategy reflects real-world deployment scenarios, where the system must generalize to variations in facial appearance without prior exposure to testing samples during training.

2. Recognition Phase (Online Inference)

The recognition phase represents the operational stage of the system, executed in real time on an edge device in the form of a mini PC and directly integrated with the door-locking mechanism when a user stands in front of the access point. Each captured frame is processed using the MTCNN algorithm to detect the presence of a face. If no face is detected, the system returns to the image acquisition stage. Once a face is detected, the facial image undergoes the same preprocessing steps applied during the training phase and is subsequently processed by FaceNet to generate a 128-dimensional embedding vector. The facial embedding is subsequently classified using the trained SVM model to determine the user's identity. The classification process using the SVM model is mathematically represented by the following decision function, which determines the identity class of a given feature vector x , as shown in Eq. (14) [38]:

$$f(x) = \text{sign} \left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right) \quad (14)$$

where $K(\cdot)$ denotes the kernel function, α_i represents the learned coefficients obtained during training, y_i corresponds to the class labels, and b is the bias term. The sign of the decision function indicates the predicted class label, which is subsequently mapped to the corresponding user identity. The classification output, consisting of the predicted user identity along with the associated confidence score, is used as the basis for access decision-making. Based on the generated confidence scores, a decision threshold of 0.90 is applied to the classification output to determine

whether the input face is classified as an authorized or unauthorized user. This threshold is intentionally set to prioritize security by minimizing the risk of false acceptance. The threshold value of 0.90 was determined through empirical evaluation on the testing dataset by analyzing the trade-off between the false acceptance rate (FAR) and the false rejection rate (FRR), with primary emphasis on minimizing false acceptance. If the resulting confidence score is greater than or equal to the defined threshold, the user identity is considered valid, and the system activates the relay module to unlock the electric drop lock. Conversely, if the confidence score falls below the threshold or the identity is not registered in the system database, access is denied, and the door remains securely locked. The relatively high threshold value is intentionally selected to prioritize the security aspect of the system. In the context of weapon storage security, the risk of unauthorized access carries far more serious consequences than the inconvenience of rejecting legitimate users. Therefore, the system is deliberately configured to be more tolerant of false rejection than false acceptance, adopting a security-oriented decision strategy. All access activities, whether granted or denied, are automatically recorded in digital logs as part of the system's security audit mechanism.

D. Performance Evaluation Metrics

The performance evaluation of the face recognition system was conducted using standard classification metrics derived from the confusion matrix. The evaluated metrics include accuracy, precision, recall, F1-score, false acceptance rate (FAR), and false rejection rate (FRR). All metrics were calculated based on the system's classification results by applying a decision threshold of 0.90. The confusion matrix consists of four fundamental components: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Based on these components, several evaluation metrics are calculated. Accuracy is computed using Eq. (15) [51], precision is calculated using Eq. (16) [51], recall is determined using Eq. (17) [51], and F1-score is formulated using Eq. (18) [51]. Furthermore, the False Acceptance Rate (FAR) and False Rejection Rate (FRR) are calculated using Eq. (19) [51] and Eq. (20) [51], respectively.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (15)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (16)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (17)$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (18)$$

$$\text{False Acceptance Rate (FAR)} = \frac{FP}{FP + TN} \quad (19)$$

$$\text{False Rejection Rate (FRR)} = \frac{FN}{FN + TP} \quad (20)$$

In the context of the proposed face recognition system, a true positive (TP) represents the number of cases in which a registered user's face is correctly recognized and granted access by the system. False negative (FN) refers to instances where a registered user's face fails to be recognized, resulting in access denial. Furthermore, true negative (TN) denotes the number of cases in which an unregistered face is correctly rejected by the system, while false positive (FP) represents instances where an unregistered face is incorrectly recognized as an authorized user and granted access. These definitions ensure that the system performance evaluation directly reflects its capability to accurately recognize legitimate users while simultaneously preventing unauthorized access.

IV. Result

This section presents the experimental results and performance evaluation of the proposed face recognition-based access control system under various operational conditions. The evaluation was conducted using multiple testing scenarios, including variations in illumination intensity, camera distance, and facial attributes, as well as testing using unregistered individuals. These experiments were designed to assess the robustness, reliability, and security performance of the proposed system in realistic operational environments.

A. Experimental Setup and Testing Scenarios

This subsection describes the testing scenarios conducted to evaluate the performance of the proposed face recognition-based access control system for weapon storage security. The experiments were carried out in an indoor environment representing real operational conditions, incorporating variations in lighting intensity, user-to-camera distance, and the use of specific facial attributes. All tests were performed on an edge device in the form of a mini PC without external network connectivity, ensuring that the measured performance reflects actual operational deployment conditions. The primary objective of this evaluation is to assess the system's robustness, accuracy, and reliability when subjected to variations in environmental conditions and user characteristics.

B. Robustness Evaluation under Operational Conditions

This subsection presents a comprehensive robustness evaluation of the face recognition system under various operational conditions commonly encountered in weapon storage environments. To ensure a clear and

systematic presentation, the evaluation is divided into three main testing scenarios: (1) lighting intensity variation, (2) camera distance variation, and (3) user facial attribute variation. Each scenario is analyzed independently to provide a structured understanding of how different environmental and user-related factors affect system performance.

1. Lighting Intensity Testing

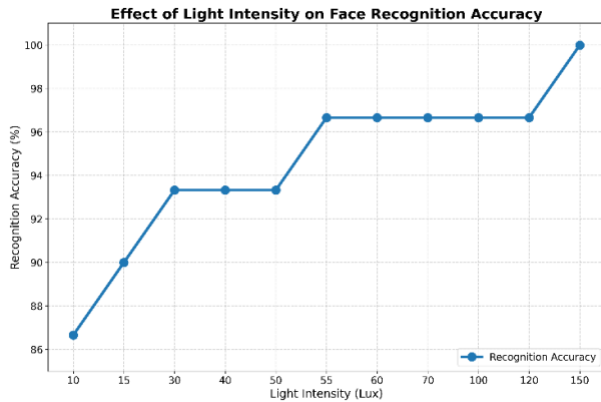


Fig. 5. Accuracy of the Proposed Face Recognition System Under Varying Illumination Intensities

The lighting intensity test was conducted to evaluate the effect of illumination variations on the performance of the face recognition system in a weapon storage environment. Illumination levels were measured using a lux meter and varied within a range of 10–150 lux, representing indoor lighting conditions from low to high intensity. The evaluation was performed using the registered dataset described in Section III.B, which consists of 25 authorized personnel with a total of 6,000 facial images. For each illumination level, 30 testing trials were conducted using samples from the testing subset (20% of the dataset), ensuring that the evaluation reflects unseen data during the training phase. Each trial corresponds to a real-time recognition attempt of a registered user under the specified lighting condition. System performance was evaluated based on the number of correct and incorrect recognitions, overall accuracy, confidence scores (ranging from 0 to 1), as well as additional evaluation metrics including precision, recall, False Acceptance Rate (FAR), and False Rejection Rate (FRR). These metrics were computed based on the confusion matrix formulation defined in Section IV.D, with a decision threshold of 0.90 applied during classification.

The experimental results indicate that under very low lighting conditions (10–15 lux), the system is still capable of recognizing faces with an accuracy of 86.66%–90.00%. Under these conditions, precision remains at 100%, and FAR remains at 0% across all tests, indicating that no unauthorized access is granted. However, recall ranges from 86.66% to

90.00%, resulting in an FRR of approximately 10.00%–13.34%, indicating a higher rate of false rejection of authorized users due to degraded image quality. As illumination intensity increases, the system performance improves consistently. Under moderate lighting conditions (30–50 lux), the system achieves an accuracy of 93.33%, with precision at 100%, FAR at 0%, recall at 93.33%, and FRR at approximately 6.67% across all test points. Under higher illumination levels (55–120 lux), the system performance becomes more stable, achieving accuracy and recall of 96.66%, with precision remaining at 100%, FAR at 0%, and FRR further reduced to approximately 3.34%. This indicates improved system reliability under more optimal lighting conditions. The highest accuracy is achieved at an illumination level of 150 lux, where the system demonstrates optimal performance across all evaluation metrics, with accuracy, precision, and recall each reaching 100%, and both FAR and FRR equal to 0%.

Overall, the system consistently maintains a FAR of 0% across all illumination variations, demonstrating its effectiveness in preventing unauthorized access. However, illumination variations primarily affect the FRR, which decreases significantly as lighting intensity increases. These results indicate that lighting quality plays a crucial role in improving the reliability of the face recognition system, particularly in reducing false rejection of authorized users. The overall trend of system accuracy under varying illumination levels is illustrated in Fig. 5. The trend shown in Fig. 5 indicates a positive correlation between illumination intensity and system accuracy. As lighting intensity increases from 10 lux to 150 lux, system accuracy improves from 86.66% to 100%, demonstrating that higher illumination levels contribute to improved facial feature extraction and more reliable recognition performance. This improvement occurs because better lighting conditions enhance facial texture visibility, enabling MTCNN to detect facial landmarks more accurately and allowing FaceNet to generate more discriminative embeddings.

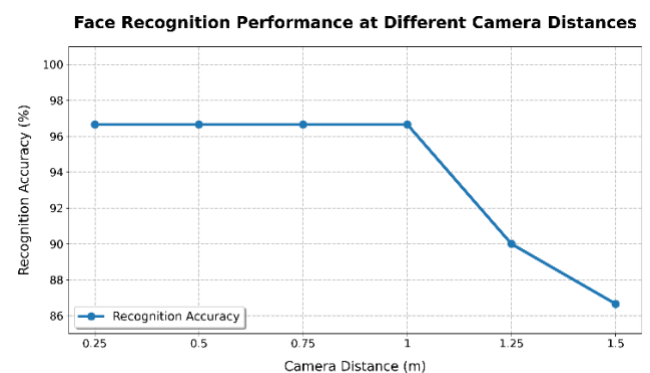


Fig. 6. Accuracy of the Proposed Face Recognition System Under Varying Camera Distances

Furthermore, although performance degradation is observed under low-light conditions, the system consistently maintains a False Acceptance Rate (FAR) of 0% across all illumination levels. This indicates that the system prioritizes security even when recognition performance decreases. The reduction in performance under low illumination is primarily reflected in increased False Rejection Rate (FRR), rather than incorrect acceptance of unauthorized users. This behavior confirms that the proposed system adopts a security-oriented decision strategy, making it suitable for deployment in high-security environments such as weapon storage warehouses.

2. Camera Distance Testing

This scenario evaluates the robustness of the system against variations in user positioning relative to the camera, which is a critical factor in real-world authentication conditions within weapon storage environments. The camera distance test was conducted to evaluate the effect of variations in the distance between the user's face and the camera on the performance of the face recognition system in a weapon storage environment. This variation represents real operational conditions, where the user's position relative to the camera may change depending on the room layout and user behavior during authentication.

The distance between the camera and the user's face was varied within a range of 0.25 m to 1.5 m. At each distance, the test was conducted 30 times using registered personnel faces, with identical system parameters applied in every trial. System performance was evaluated based on the number of correct and incorrect recognitions, overall accuracy, as well as additional evaluation metrics including precision, recall, FAR, and FRR. The experimental results indicate that within a distance range of 0.25 m to 1.0 m, the system maintains stable recognition performance with an accuracy of 96.66%. In this range, precision remains at 100%, and FAR remains at 0% across all tests, indicating that no unauthorized access is granted. The recall is consistently 96.66%, resulting in a low FRR of approximately 3.34%, reflecting reliable recognition performance at typical operational distances.

As the distance increases beyond 1.0 m, a noticeable decline in system performance is observed. At a distance of 1.25 m, the accuracy decreases to 90.00%, with recall reduced to 90.00% and FRR increasing to approximately 10.00%. This degradation becomes more significant at a distance of 1.5 m, where the system achieves an accuracy of 86.66%, a recall of 86.66%, and an FRR of approximately 13.34%. Despite this decline, precision remains at 100%, and FAR remains at 0% across all tested distances, indicating that the system consistently prevents false acceptance of unauthorized users. The observed performance degradation is primarily attributed to the reduction in

effective facial resolution at greater distances, which negatively impacts face detection by MTCNN and the quality of feature embeddings generated by FaceNet. Overall, the results demonstrate that the system exhibits strong robustness to distance variations up to 1.0 m, while performance degradation at longer distances is mainly reflected in increased false rejection. This trend shows that accuracy decreases by approximately 6.66% at 1.25 m and 10% at 1.5 m compared to the optimal distance range. Similarly, the FRR increases from 3.34% to 10.00% and 13.34%, respectively, indicating that increased distance primarily affects recognition reliability through higher false rejection rates while maintaining consistent access security. The trend of system accuracy under varying camera distances is illustrated in Fig. 6. The trend illustrated in Fig. 6 shows that system accuracy remains stable within the distance range of 0.25 m to 1.0 m, maintaining a consistent accuracy of 96.66%. However, when the distance increases to 1.25 m and 1.5 m, system accuracy decreases to 90.00% and 86.66%, respectively. This gradual decline indicates that increased distance reduces the effective facial resolution captured by the camera, which affects the performance of MTCNN in detecting facial landmarks and reduces the discriminative capability of embeddings generated by FaceNet. Despite the decrease in accuracy at longer distances, the system consistently maintains a precision of 100% and a FAR of 0% across all tested distances, indicating strong access security. The performance decline is primarily reflected in the increase of FRR from 3.34% to 13.34% at 1.5 m. These results suggest that the system performs optimally within the 0.25 m to 1.0 m operational range.

3. Facial Attribute Testing

This scenario evaluates the robustness of the system against variations in user appearance caused by operational equipment and real-world usage conditions. The facial attribute test was conducted to evaluate the robustness of the face recognition system against variations in users' physical appearance that may affect the facial feature extraction process. The evaluated attributes include faces without accessories, the use of hats, eyeglasses, military helmets, tilted head orientation, and faces with markings or camouflage. These variations represent real operational conditions in weapon storage environments, where personnel may wear additional equipment or experience changes in facial appearance. The tests were performed at camera distances ranging from 0.25 m to 1.50 m using identical system parameters and a decision threshold of 0.90. System performance evaluation was based on the range of confidence scores generated by the classification model. A summary of the facial attribute

Table 4. Facial Attribute Testing at Various Camera Distances Based on Confidence Scores (%)

Distance (m)	Plain Face	Hat	Glasses	Military Helmet	Tilted Face	Camouflage Paint
0.25	97–98%	97–98%	91–92%	97–98%	90–92%	90–92%
0.50	98–99%	98–99%	90–91%	96–97%	90–91%	90–92%
0.75	98–99%	97–98%	90–91%	95–97%	91–93%	90–91%
1.00	98–99%	93–97%	90–91%	95–96%	90–92%	50–68%*
1.25	98–99%	93–96%	90–91%	94–95%	90–92%	50–55%*
1.50	95–96%	90–92%	90–91%	90–94%	89–90%	40–48%*

testing results at various camera distances is presented in Table 4. Based on the results, the face recognition system demonstrated highly stable performance in the scenario without facial accessories, achieving high confidence scores ranging from 95% to 99% across all distance variations from 0.25 m to 1.50 m. The highest confidence values were observed at 0.50 m and 0.75 m, reaching 98%–99%, while slightly lower values of 95%–96% were observed at 1.50 m. This indicates that facial features can be optimally extracted when no visual obstruction is present in the primary facial regions. In the scenarios involving the use of hats, eyeglasses, and military helmets, the system was also able to recognize faces reliably at all tested distances. For the hat condition, confidence scores ranged from 90% to 99%, with the highest performance observed at 0.50 m (98%–99%) and the lowest at 1.50 m (90%–92%). In the eyeglasses scenario, confidence scores ranged from 90% to 92% across all distances, indicating stable performance. Similarly, in the military helmet condition, confidence scores ranged from 90% to 98%, with a gradual decrease observed at longer distances. These results suggest that MTCNN consistently detects facial landmarks, while FaceNet continues to generate representative embeddings even when certain facial areas are partially occluded by accessories. In the tilted-face scenario, the system was still able to perform recognition up to a distance of 1.50 m, with confidence scores ranging from 89% to 93%. The highest value of

91%–93% was observed at 0.75 m, while the lowest value of 89%–90% occurred at 1.50 m. Although a decrease in confidence was observed compared to the frontal-face condition, most values remained around or above the defined threshold. This result indicates that landmark-based face alignment plays a crucial role in maintaining spatial consistency of facial images prior to feature extraction.

However, in the camouflage scenario, a significant performance decline was observed at distances of 1.00 m to 1.50 m, with confidence scores decreasing from 50%–68% at 1.00 m, 50%–55% at 1.25 m, and further to 40%–48% at 1.50 m, resulting in recognition failures. This degradation is caused by the obstruction of key facial features, leading to less representative embeddings generated by FaceNet. Overall, the results indicate that the system maintains stable recognition performance under most operational facial attribute conditions, particularly when accessories such as hats, glasses, and helmets are used. The confidence scores consistently remained above the 0.90 threshold, demonstrating that the proposed MTCNN–FaceNet–SVM pipeline is robust against moderate facial occlusions and pose variations. Furthermore, performance degradation was primarily observed only under extreme facial modifications, such as camouflage paint, which significantly alters facial texture and contour information. This finding suggests that while the system is highly reliable for operational deployment, additional feature enhancement or multi-

Table 5. Performance Metrics Across Facial Attribute Variations

Facial Attribute	Accuracy	Precision	Recall	FAR	FRR
Plain Face	97.78%	100%	97.78%	0.00%	2.22%
Hat	95.56%	100%	95.56%	0.00%	4.44%
Glasses	90.56%	100%	90.56%	0.00%	9.44%
Military Helmet	95.56%	100%	95.56%	0.00%	4.44%
Tilted Face	90.56%	100%	90.56%	0.00%	9.44%
Camouflage Paint	61.11%	100%	61.11%	0.00%	38.89%

Table 6. Testing Unregistered Face Dataset with Confidence Score (%)

Distance (m)	Plain Face	Hat	Glasses	Military Helmet	Tilted Face	Camouflage Paint
0.25	30–50%	20–50%	20–50%	20–50%	20–50%	20–50%
0.50	30–50%	20–50%	20–50%	20–50%	20–50%	20–50%
0.75	30–50%	20–50%	20–50%	20–50%	20–50%	20–50%
1.00	20–50%	20–50%	20–50%	20–50%	20–50%	20–50%
1.25	20–50%	20–50%	20–50%	20–50%	20–50%	20–50%
1.50	20–50%	20–50%	20–50%	20–50%	20–50%	20–50%

modal biometric integration may further improve robustness under extreme appearance variations. To provide a more comprehensive evaluation, system performance under various facial attribute conditions is further quantified using accuracy, precision, recall, False Acceptance Rate (FAR), and False Rejection performance and potential misclassification under different facial attribute variations and operational conditions. As shown in Table 5, the system achieves the best performance in the plain face scenario, with an accuracy and recall of 97.78% and a low FRR of 2.22%, indicating optimal conditions without visual obstruction. For moderate variations such as hats and military helmets, the system maintains stable performance with an accuracy and recall of 95.56% and an FRR of 4.44%, suggesting that partial occlusion does not significantly affect facial feature extraction. In contrast, the use of eyeglasses and tilted face orientation leads to a performance decline, with accuracy and recall decreasing to 90.56% and FRR increasing to 9.44%. This indicates that occlusion in critical facial regions and variations in face orientation affect the feature extraction and alignment processes. The most significant degradation is observed in the camouflage scenario, where the system achieves only 61.11% accuracy and recall, with a substantially higher FRR of 38.89%. This result indicates that major alterations in facial texture significantly disrupt the embedding generation process.

Overall, the system consistently maintains a precision of 100% and FAR of 0% across all scenarios, demonstrating strong resistance to unauthorized access. However, variations in facial attributes primarily affect the FRR, particularly under extreme conditions. Overall, the robustness evaluation results demonstrate that the proposed face recognition system maintains stable performance under moderate variations in lighting intensity, camera distance, and facial attributes. However, performance degradation is observed under more extreme conditions, such as very low illumination, increased distance, and significant

Rate (FRR). These evaluation metrics provide a quantitative assessment of the system's reliability, allowing analysis of both successful recognition

facial alterations. These findings highlight the importance of environmental and user-related factors in influencing system reliability in real-world weapon storage applications.

C. Unregistered Dataset Evaluation and FAR–FRR Analysis

The evaluation using an unregistered face dataset was conducted to assess the system's ability to reject access attempts from individuals who do not have reference data stored in the system database. This testing is essential for measuring the system's security level, particularly in preventing false acceptance of unauthorized identities. The tests were performed using unregistered faces under various camera distances ranging from 0.25 m to 1.50 m, and multiple facial attribute conditions, including plain face, hats, eyeglasses, military helmets, tilted faces, and camouflage markings. All experiments were conducted using identical system parameters and the same decision threshold of 0.90 applied in the registered

Table 7. FAR and FRR Evaluation

Dataset	FAR (%)	FRR (%)	Remarks
Registered	-	7.50	Errors on extreme conditions
Unregistered	0	-	All faces correctly rejected

dataset evaluation. As shown in Table 6, the confidence scores for unregistered faces remained consistently low across all scenarios. In the plain face

condition, confidence scores ranged from 30%–50% at distances of 0.25 m, 0.50 m, and 0.75 m, and decreased to 20%–50% at 1.00 m to 1.50 m. For the hat, glasses, military helmet, tilted face, and camouflage conditions, confidence scores consistently ranged between 20%–50% across all tested distances from 0.25 m to 1.50 m. These values are significantly below the defined decision threshold, indicating that all unregistered individuals were correctly rejected by the system. No instances of false acceptance were observed across all tested distances and facial attribute variations. A summary of the unregistered dataset evaluation results is presented in Table 6 for clarity easier quantitative interpretation.

Based on Table 6, the face recognition system consistently rejected all unregistered individuals across all tested distances and facial attribute variations. The confidence scores remained within the 20%–50% range, which is significantly below the defined decision threshold of 0.90, indicating that no unauthorized users were incorrectly accepted by the system. These results demonstrate the system's strong capability in distinguishing unknown identities and preventing unauthorized access under various operational conditions. Furthermore, the consistently low confidence scores across all testing scenarios confirm that the proposed system effectively minimizes the risk of false acceptance and maintains a high level of access security.

To provide a more comprehensive overview of the system's performance, an evaluation was conducted using the False Acceptance Rate (FAR) and False Rejection Rate (FRR) metrics. A summary of this evaluation is presented in Table 7, comparing the system's performance on registered and unregistered face datasets. Based on Table 7, the system achieved a FAR of 0% on the unregistered face dataset, indicating that no unauthorized access attempts were

successfully accepted. Meanwhile, the FRR for the registered face dataset was 7.50%, caused by a small number of recognition failures under extreme operational conditions, such as variations in facial attributes and increased distance between the user and the camera. The zero FAR demonstrates that the system effectively prioritizes security by preventing unauthorized access. On the other hand, the relatively low FRR remains within an acceptable range for a high-security system. The combination of these FAR and FRR values indicates that the proposed face recognition system maintains a good balance between security and reliability, making it suitable for deployment in high-risk environments such as weapon storage warehouses.

D. Confusion Matrix and Performance Metrics

The performance of the face recognition system was evaluated using a confusion matrix to assess its ability to correctly recognize registered users and reject unregistered users. This evaluation was conducted based on the testing dataset, which represents 20% of the total dataset as shown in Table 3 from the total of 6,000 collected facial images, 1,200 images were used for testing, while the remaining images were used for system training. The testing dataset included both registered and unregistered faces to objectively evaluate the classification capability of the system, independent of the training process. This data split strategy was implemented to prevent evaluation bias and ensure that the system's performance reflects realistic testing conditions. Additionally, the inclusion of unregistered samples enables a more comprehensive assessment of the system's ability to distinguish between authorized and unauthorized users. Based on the evaluation results using the testing dataset, the system achieved 555 True Positives (TP), 45 False Negatives (FN), 600 True Negatives (TN), and 0 False Positives (FP). The resulting confusion matrix illustrates that the system successfully recognized the

Confusion Matrix of Face Recognition System (N=1200)

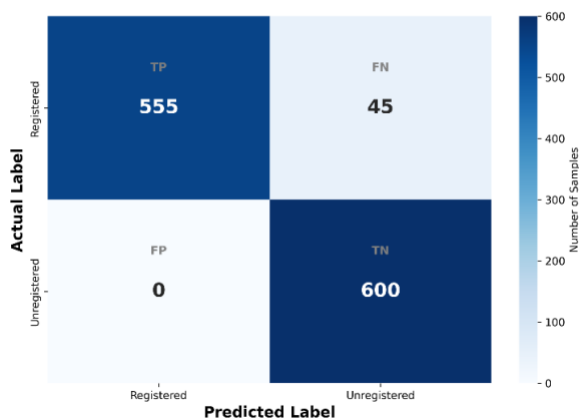


Fig. 8 Confusion Matrix of Face Recognition System

Performance Metrics of Face Recognition System

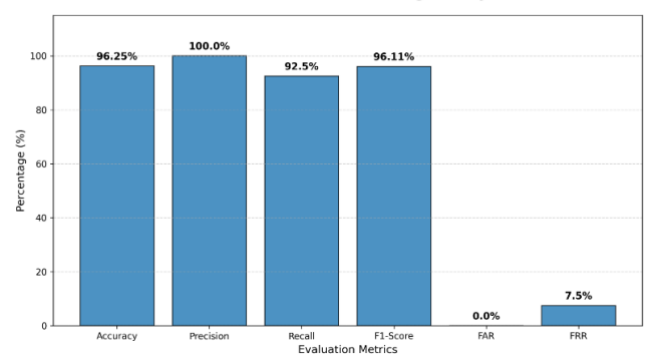


Fig. 7. Performance Metrics of Face Recognition System

majority of authorized users while consistently rejecting all access attempts from unregistered faces.

Furthermore, the confusion matrix analysis provides a comprehensive understanding of the system's classification behavior under realistic operational conditions. The absence of false positives indicates that the system maintains strict access control by preventing unauthorized individuals from being incorrectly recognized as authorized users. Meanwhile, the presence of false negatives reflects a small number of cases where authorized users were not successfully recognized, which may occur due to variations in illumination, facial attributes, or capture distance. This behavior demonstrates that the proposed system prioritizes security by minimizing the risk of false acceptance while maintaining reliable recognition performance for registered personnel.

The absence of False Positives indicates that the system maintains a high level of security, as it does not

affect feature extraction performance. To provide a more comprehensive view of system performance, evaluation metrics including accuracy, precision, recall, F1-score, False Acceptance Rate (FAR), and False Rejection Rate (FRR) were calculated. A comparison of the system's performance metrics is presented in Fig. 8. Based on the evaluation of performance metrics, the proposed face recognition system achieved an accuracy of 96.25%, precision of 100%, recall of 92.50%, and an F1-score of 96.11%. The precision value of 100% indicates that all faces granted access by the system belong to authorized users. Furthermore, a FAR of 0% signifies that the system did not make any false acceptance of unregistered faces. Meanwhile, the FRR of 7.50% reflects a small proportion of legitimate users being rejected, which remains within acceptable limits for high-risk security systems. Overall, the evaluation results from the confusion matrix and performance metrics demonstrate that the proposed face recognition system maintains a good balance

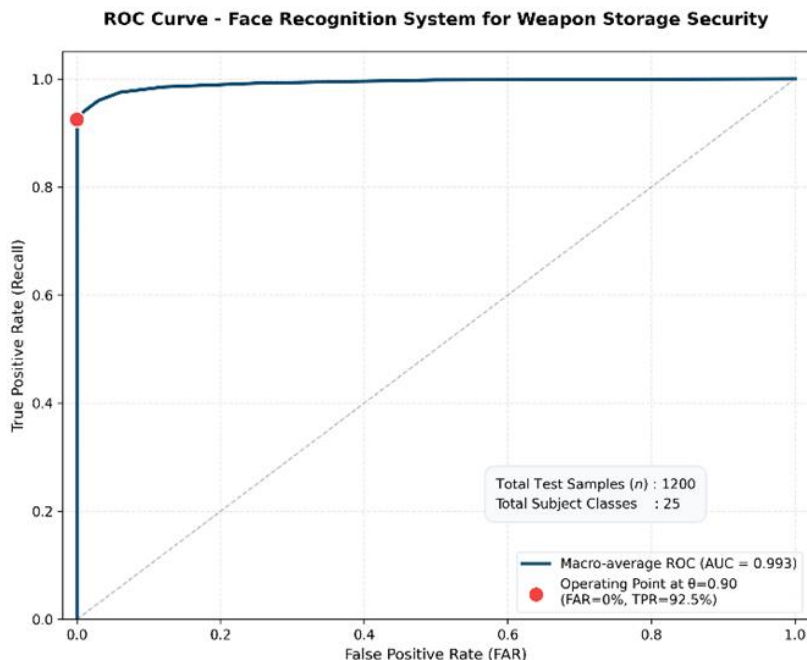


Fig. 9 ROC curve of the proposed face recognition system with AUC = 0.993 and operating point at FAR = 0% and TPR = 92.5%

grant access to individuals whose data is not present in the database. Meanwhile, the presence of False Negatives reflects a small number of recognition failures for registered users, which generally occur under specific testing conditions, such as variations in facial attributes and changes in the image capture distance. These failures were primarily observed under challenging conditions, including low illumination, longer camera distances, and facial camouflage that

between security and recognition reliability. The system effectively prioritizes the prevention of unauthorized access without significantly compromising recognition performance, making it suitable for deployment in high-security environments, such as weapon storage warehouses.

E. ROC Curve Analysis

The ROC curve in Fig. 9 is used to evaluate the discriminative ability of the system in distinguishing

between registered and unregistered faces across various threshold values. Based on testing with 1,200 test samples, the ROC curve demonstrates excellent performance, with an Area Under the Curve (AUC) of 0.993, approaching the ideal value. The curve is positioned near the top-left corner, indicating that the system achieves a high true positive rate at a very low false positive rate. At the operational threshold of 0.90, the system attains a True Positive Rate (TPR) of 92.5% with a False Positive Rate (FAR) of 0%, confirming that the system is effective in preventing unauthorized access while maintaining a high recognition rate for authorized users.

F. Response Time Analysis

System response time testing for the face recognition system was conducted to evaluate the system's performance in controlling the door locking mechanism in real-time. This evaluation aims to ensure that the system is not only accurate but also has a sufficiently fast response time to be deployed in a weapons storage environment, which demands high efficiency and security. The testing was carried out by measuring the time required from the moment a facial image is captured by the camera until the system provides the final decision, either granting or denying door access. Tests were conducted across various distances

between the face and the camera, ranging from 0.25 m to 1 m, and under different facial conditions, including normal faces, with hats, glasses, or military helmets, tilted faces, and faces with markings. In this study, the term "normal faces" refers to facial images of registered users captured under standard conditions, specifically without additional accessories (i.e., hats, glasses, or helmets), with a frontal face orientation, and under adequate lighting conditions. All tests were performed using the same system configuration and hardware. A summary of the system response time results under these conditions is presented in Fig. 10.

Based on the test results, the fastest response times were observed for normal faces and for faces with military helmets, with average response times below 1.2 seconds across all tested distances. This indicates that the system is capable of efficiently processing faces under relatively stable visual conditions. In contrast, increased response times were observed for tilted faces and faces with markings. At a distance of 1 m, the response time for tilted faces reached up to 2.88 seconds, while for faces with markings it increased to 3.22 seconds. This increase is attributed to the additional complexity in the detection and feature extraction processes, where changes in face orientation and distortions in facial texture require more

Analysis of System Response Time Across Distances and Attributes

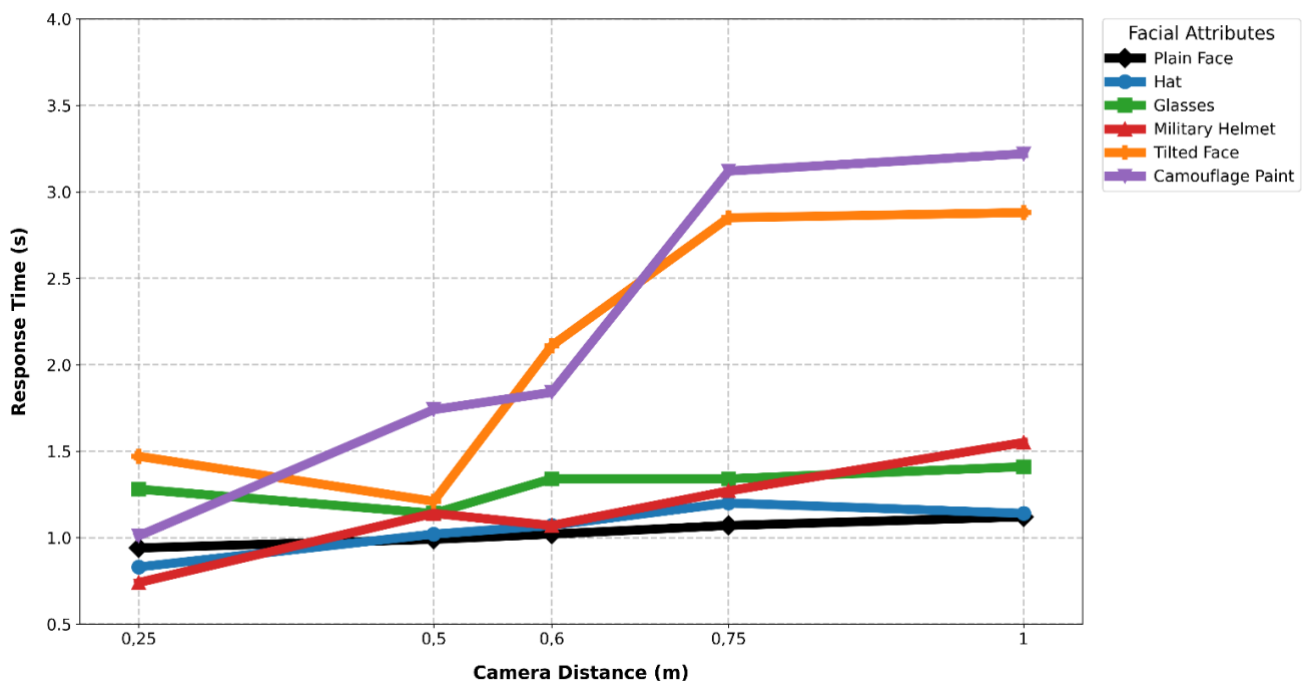


Fig. 10 Analysis of System Response Time Across Distances and Attributes

intensive processing by the MTCNN and FaceNet models. In general, increasing the distance between the face and the camera also impacts system response

time, although the effect is not significant for normal faces or faces with standard accessories. This indicates that the system is still capable of maintaining

stable response performance across the tested operational distances. Overall, the results demonstrate that the proposed face recognition system can achieve an average response time below 3.2 seconds across all testing scenarios. This response time remains within acceptable limits for biometric security systems, particularly in weapons storage applications, where accuracy and security remain the primary priorities over absolute speed.

V. Discussion

This study proposes a face recognition system based on the combination of MTCNN as the face detector, FaceNet for feature extraction, and SVM as the classifier, implemented in a weapon storage security system. Based on experimental testing across various operational scenarios, the system demonstrates stable and consistent performance in recognizing registered faces while rejecting unregistered ones. Overall evaluation results show that the system achieved an accuracy of 96.25%, precision of 100%, recall of 92.50%, and F1-score of 96.11%. Furthermore, a False Acceptance Rate (FAR) of 0% confirms that the system successfully prevented unauthorized access, while a False Rejection Rate (FRR) of 7.50% indicates a small number of authorized users being rejected, which remains acceptable in high-risk security environments.

Testing under varying light intensities showed that system performance tends to improve with increasing illumination levels. Under low-light conditions of approximately 10 lux, system accuracy decreased to 86.66%, while at higher illumination levels above 100 lux, accuracy increased to 100%. Although performance decreased under low-light conditions, FAR remained at 0% across all scenarios. This indicates that MTCNN and FaceNet exhibit considerable robustness to varying lighting conditions, although image quality remains a significant factor affecting the quality of the generated face embeddings. In camera distance testing, the system maintained high accuracy across close-to-medium ranges. At distances between 0.25 m and 1.0 m, accuracy remained above 96.66%, while at a distance of 1.5 m, performance decreased to 86.66%. This decrease is attributed to reduced facial resolution captured by the camera, which impacts the feature extraction process by FaceNet. Nevertheless, FAR remained at 0%, indicating that security performance remained stable even under reduced image quality.

Furthermore, testing with variations in facial attributes demonstrated that the system exhibited good robustness against common accessories such as glasses, hats, and military helmets, with accuracy values remaining above 90%. Under these conditions, the system was still able to recognize user identities

with relatively stable confidence values. Conversely, under conditions with face markings or camouflage, the system experienced a more significant performance decline. Accuracy dropped to 61.11% and FRR increased to 38.89%. These findings suggest that significant alterations to the facial surface have a greater impact on system performance than the use of external accessories.

To address this limitation, a failure analysis was conducted to further investigate the causes of recognition failures observed in the experimental results. Under low-light conditions, recognition errors were primarily caused by increased image noise and reduced contrast, which negatively affected the accuracy of face detection by MTCNN. In scenarios involving increased camera distance, reduced facial resolution led to decreased feature discriminability. Meanwhile, camouflage conditions caused distortions in facial texture and key facial features, resulting in lower classification confidence and higher false rejection rates. Evaluation on the unregistered face dataset showed that the system successfully rejected all unauthorized users, achieving a FAR of 0%. The confidence values consistently remained below the system's decision threshold of 0.90. This confirms that the system possesses strong discriminative capability in distinguishing between authorized and unauthorized users, which is essential for weapon-storage security systems.

Based on confusion matrix evaluation on 20% of the dataset, the system achieved an accuracy of 96.25%, precision of 100%, recall of 92.50%, and F1-score of 96.11%. The precision of 100% indicates that all recognized users were correctly identified. Meanwhile, a recall of 92.50% indicates a small number of authorized users being rejected under challenging conditions. This reflects the system's security-oriented configuration. From an operational perspective, the observed FRR of 7.50% indicates that some authorized users may experience access denial under challenging conditions, such as low illumination, increased camera distance, or significant facial variations. However, in high-security environments such as weapon storage facilities, this trade-off is considered acceptable, since preventing unauthorized access is significantly more critical than user convenience. Therefore, a moderate level of false rejection is tolerable compared with the potential risks of false acceptance.

The decision threshold of 0.90 plays a critical role in achieving this balance. The selected threshold successfully achieved a False Acceptance Rate (FAR) of 0% while maintaining an acceptable False Rejection Rate (FRR) of 7.50%. This result indicates that the proposed system effectively prioritizes security by eliminating unauthorized access while maintaining practical usability for authorized personnel. To evaluate

the performance of the proposed system, it was compared with several previous studies employing different classification approaches for face recognition tasks. A summary of methods, testing scenarios, and accuracy levels is presented in Table 8. This comparison provides an overview of performance differences and system characteristics. In addition to accuracy-based evaluation, the operational feasibility of the proposed system was also demonstrated through stable real-time performance and consistent security behavior across multiple testing scenarios. The system maintained reliable recognition performance even under challenging conditions, including low illumination (86.66% accuracy), increased camera distance (86.66% at 1.5 m), and facial camouflage (61.11% accuracy). Despite

performance degradation in these scenarios, the system consistently maintained a FAR of 0%, indicating strong resistance to unauthorized access. This behavior highlights the robustness of the proposed architecture and confirms that the system is more suitable for high-security environments where preventing unauthorized access is prioritized over maintaining maximum recognition accuracy. Furthermore, the stable performance across different operational conditions demonstrates the system's capability to function reliably in real-world deployment scenarios. These findings indicate that the proposed system achieves a balanced trade-off between recognition reliability and strict access security, which is essential for weapon storage warehouse applications.

Table 8. Comparison of Performance and Operational Characteristics

Study	Method	Device/Platform	Accuracy	Real-time	Security Level
[52]	LBP + PCA + KNN	Desktop PC	91.0%	Y	Moderate
[53]	Random Forest	Desktop PC	96.1%	Y	Moderate
[54]	MTCNN + Softmax CNN	Embedded/PC	95.0%	Y	Moderate
This Study	MTCNN+FaceNet+SVM	Edge (Mini PC)	96.25%	Y	High

Table 8 compares the classification methods, device platforms, accuracy, real-time capability, and system security levels. The security level classification is primarily based on the availability and reporting of the FAR, which reflects the system's ability to prevent unauthorized access. The security level classification does not imply relative superiority, but rather reflects the availability of verifiable security metrics reported in each study. Based on the comparison, the study by Gaur et al. [52], which employed a combination of LBP, PCA, and KNN, achieved an accuracy of 91.0% on a desktop platform with real-time capability. However, its security level is categorized as moderate, as no explicit FAR or unauthorized access rejection metrics were reported, and therefore, its security performance cannot be quantitatively verified. Similarly, the Random Forest-based approach proposed by Amirgaliyev et al. [53] demonstrated an increased accuracy of 96.1%, but it is also categorized as having moderate security due to the absence of explicitly reported FAR or equivalent security evaluation metrics.

The deep learning-based approach used by Prasanna et al. [54], which combined MTCNN and Softmax CNN, achieved an accuracy of 95.0% with real-time support on embedded and PC platforms. However, the system still relied on softmax-based classification and did not explicitly report security-oriented evaluation metrics such as FAR, making its

resistance to unauthorized access difficult to verify. In this study, the combination of MTCNN, FaceNet, and SVM was implemented on an edge device (Mini PC) and achieved an accuracy of 96.25%, comparable to or exceeding previous studies. The main advantage of the proposed system lies in its security level, demonstrated by a FAR of 0%, indicating that no unauthorized access was accepted during evaluation. This allows the system to be categorized as having a high security level. Moreover, the edge implementation shows that the system can operate in real time without relying on cloud computing, thereby enhancing both reliability and data security. Therefore, this comparative analysis indicates that the proposed system is not only competitive in terms of accuracy but also provides a more clearly verifiable level of security performance and readiness for deployment in real operational environments. This classification ensures a more transparent and objective comparison of system security across different studies.

Despite the experimental results demonstrating stable performance across the conducted testing scenarios, this study has several limitations that should be noted. The dataset used consists of a relatively small number of subjects (25 personnel) and was collected in a controlled environment tailored to the context of weapon storage access, and thus does not fully represent broader real-world variations such as

extreme lighting conditions, heavy occlusion, and significant appearance changes. To mitigate these limitations, data were collected across multiple sessions under varied conditions and with diverse facial attributes, although cross-environment generalization still requires further validation. Additionally, the current system does not yet integrate liveness detection or anti-spoofing mechanisms to protect against presentation attacks, and computational evaluation was performed on a single edge device configuration. Future research will focus on using larger and more diverse datasets, testing broader attack scenarios, benchmarking across multiple devices, and integrating presentation attack detection modules to enhance the robustness and generalization capability of the system.

VI. Conclusion

This study successfully designed and implemented a face recognition-based access control system for weapon storage facilities by integrating MTCNN for face detection, FaceNet for feature extraction, and SVM for identity classification, evaluated under real operational conditions, including variations in lighting, camera distance, and facial attributes. The experimental results demonstrated reliable performance with an accuracy of 96.25%, precision of 100%, recall of 92.50%, and F1-score of 96.11%, while achieving a False Acceptance Rate (FAR) of 0% and a False Rejection Rate (FRR) of 7.50%, indicating strong security performance suitable for high-risk environments. The system, implemented on an edge-based mini PC, also achieved real-time operation with a stable response time, demonstrating its feasibility for deployment in weapon storage security applications. Compared with previous studies, the proposed system achieves competitive accuracy while improving security performance and operational readiness. For future work, enhancements will focus on integrating liveness detection or anti-spoofing mechanisms, expanding dataset diversity and data augmentation strategies, implementing adaptive thresholding for dynamic FAR-FRR optimization, evaluating performance across multiple edge devices, and exploring multimodal approaches using infrared or thermal sensors to improve robustness under extreme operational conditions.

Acknowledgment

The authors would like to thank Telkom University for its support in facilities, funding, and research resources. Appreciation is also extended to the Center of Excellence STAS RG Telkom University for its technical support and facilitation of data collection and equipment.

Thank you to all parties involved in the testing process of this research.

Funding

This research received no external funding

Data Availability

The data are not publicly available due to privacy and security restrictions.

Author Contribution

Conceptualization, A.A. and G.A.M.; methodology, A.A.; software, A.A.; validation, A.A. and M.R.A.; formal analysis, A.A.; investigation, A.A.; data curation, A.A.; writing original draft preparation, A.A.; writing, review and editing, G.A.M. and M.R.A.; supervision, G.A.M.; project administration, G.A.M.; funding acquisition, G.A.M. All authors have read and agreed to the published version of the manuscript.

Declarations

Ethical Approval

This study involved the voluntary participation of registered personnel. All participants provided informed consent prior to data collection, and all facial data were used strictly for research purposes while maintaining confidentiality and privacy.

Consent for Publication Participants.

Consent for publication was given by all participants

Competing Interests

The authors declare no competing interests.

References

- [1] P. Zhou and W. Zhang, "Research on Computer Network Information Security and Protection Strategy Based on Deep Learning Algorithm," *Proc. - 2020 Int. Conf. Adv. Ambient Comput. Intell. ICAACI 2020*, pp. 181–184, Sep. 2020, doi: 10.1109/ICAACI50733.2020.00046.
- [2] K. Wysocki, "Protecting Critical Infrastructure: Methods and Techniques," in *Civil Protection Systems and Disaster Governance: a Cross-Regional Approach*, Springer Nature, 2024, pp. 41–60. doi: 10.1007/978-3-031-60167-5_3.
- [3] H. Martinez, F. J. Rodriguez-Lozano, F. León-García, J. M. Palomares, and J. Olivares, "Distributed Fog computing system for weapon detection and face recognition," *J. Netw. Comput. Appl.*, vol. 232, p. 104026, Dec. 2024, doi: 10.1016/J.JNCA.2024.104026.
- [4] M. R. Alfarisi, P. Telsoni, P. Aji, G. A. Mutiara, and Periyadi, "Development of a Web-Based Weapon Rack Security System Utilizing RFID

- Technology and Real-Time Data Logging," *Int. J. Comput. Methods Exp. Meas.*, vol. 13, no. 1, pp. 157–163, Mar. 2025, doi: 10.18280/ijcmem.130117.
- [5] B. Indrawan, D. G. Alzamora, M. S. Rahmatullah, and A. Wikarta, "Facial Recognition-Based Automatic Door Security System Integrated with Internet of Things for Smart Home Actualization," *Lect. Notes Mech. Eng.*, pp. 360–368, 2023, doi: 10.1007/978-981-19-0867-5_43.
- [6] Y. Grover, P. Sharma, and N. Kansal, "Revolutionizing Face Recognition with IIoT: Smarter, Faster, Connected," *ICCECE 2025 - Int. Conf. Comput. Electr. Commun. Eng.*, 2025, doi: 10.1109/ICCECE61355.2025.10941049.
- [7] L. Chetty, A. Odowa, A. C. Avenido, I. Hussein, and Y. Elakkad, "Performing facial recognition using ensemble learning," in *Advanced Interdisciplinary Applications of Machine Learning Python Libraries for Data Science*, IGI Global, 2023, pp. 89–123. doi: 10.4018/978-1-6684-8696-2.ch004.
- [8] H. M. Al-Dabbas, R. A. Azeez, and A. E. Ali, "Machine Learning Approach for Facial Image Detection System," *Iraqi J. Sci.*, vol. 64, no. 10, pp. 5428–5441, 2023, doi: 10.24996/ijjs.2023.64.10.44.
- [9] D. Wu, L. Li, W. Tian, H. Xian, and C. Tian, "Biometric identification on the cloud: A more secure and faster construction," *Inf. Sci. (Ny.)*, vol. 669, p. 120553, May 2024, doi: 10.1016/J.INS.2024.120553.
- [10] M. Omara, M. Fayez, H. Khalid, and S. Ghoniemy, "A Transfer Learning Approach for Face Liveness Detection," *Proc. - 11th IEEE Int. Conf. Intell. Comput. Inf. Syst. ICICIS 2023*, pp. 122–127, 2023, doi: 10.1109/ICICIS58388.2023.10391203.
- [11] P. Bhavya Sri, K. Navya, K. Saiteja, P. Keerthi, S. Hariharan, and V. Kukreja, "Harnessing Security Improvements Using FaceNet Approach for Face Recognition System," *Proc. - 2024 13th IEEE Int. Conf. Commun. Syst. Netw. Technol. CSNT 2024*, pp. 306–312, 2024, doi: 10.1109/CSNT60213.2024.10545894.
- [12] P. Ahlawat, N. Kaur, C. Kaur, S. Kumar, and H. K. Sharma, "Deep Learning Based Face Recognition System for Automated Identification," *Commun. Comput. Inf. Sci.*, vol. 1930, pp. 60–72, 2024, doi: 10.1007/978-3-031-48781-1_6.
- [13] K. Ramalakshmi, B. J. Jingle, C. P. Shirley, V. Suvishik, P. Joyce Beryl Princess, and K. Vidhya, "Facial Recognition System with LBPH Algorithm: Implementation in Python for Machine Learning," *2nd Int. Conf. Intell. Cyber Phys. Syst. Internet Things, ICoICI 2024 - Proc.*, pp. 1681–1686, 2024, doi: 10.1109/ICoICI62503.2024.10696268.
- [14] S. E. Bakan and M. Yildirim, "An Access Control with Face Recognition Based on LBPH Algorithm," *Proc. - 2025 IEEE 7th Glob. Power, Energy Commun. Conf. GPECOM 2025*, pp. 991–996, 2025, doi: 10.1109/GPECOM65896.2025.11062016.
- [15] C. Ung, P. Mantini, and S. K. Shah, "Minimizing Number of Distinct Poses for Pose-Invariant Face Recognition," *Proc. Int. Jt. Conf. Comput. Vision, Imaging Comput. Graph. Theory Appl.*, vol. 2, pp. 447–455, 2025, doi: 10.5220/0013186400003912.
- [16] X. Wang and W. Zhang, "Anti-occlusion face recognition algorithm based on a deep convolutional neural network," *Comput. Electr. Eng.*, vol. 96, Dec. 2021, doi: 10.1016/j.compeleceng.2021.107461.
- [17] M. Hao, F. Yuan, J. Li, and Y. Sun, "Facial expression recognition based on regional adaptive correlation," *IET Comput. Vis.*, vol. 17, no. 4, pp. 445–460, Jun. 2023, doi: 10.1049/cvi2.12179.
- [18] S. G. Aydin and H. S. Bilge, "Optimal hardware implementation for end-To-end CNN-based classification," *2023 Int. Conf. Innov. Trends Inf. Technol. ICITIIT 2023*, 2023, doi: 10.1109/ICITIIT57246.2023.10068601.
- [19] D. Santhakumar, S. Nagaraju, S. Sivamani, B. S. N. Prasad, V. Veeresh, and V. Nithianantharaj, "Facial Recognition and Image Processing in Security and Surveillance Systems Using Deep Learning Algorithms and Multi-task Cascade Deep Convolutional Neural Networks," *3rd Int. Conf. Adv. Comput. Commun. Mater. ICACCM 2024*, 2024, doi: 10.1109/ICACCM61117.2024.11059009.
- [20] M. Mohana and P. Subashini, "Analysing the performance of Viola-Jones and multi-task convolution neural networks face detection algorithms using real-time video sequences," *Int. J. Comput. Vis. Robot.*, vol. 15, no. 3, pp. 286–311, 2025, doi: 10.1504/IJCVR.2025.146293.
- [21] M. K. Roy, P. Dwibedi, A. Singh, R. P. Chakraborty, and M. K. H. Mondal, "MTCNN and FACENET-Based Face Detection and Recognition Model for Attendance Monitoring," *Smart Innov. Syst. Technol.*, vol. 376, pp. 525–539, 2024, doi: 10.1007/978-981-99-7711-6_42.
- [22] D. R. Ghaida, V. V. Siedharta, U. Hakim, K.

- Mutijarsa, A. I. Septiana, and Y. Rosmansyah, "SVM-Classified FaceNet and Eigenface Models under Lighting and Occlusion Variations for Face Recognition," *Proc. - 2024 Int. Semin. Appl. Technol. Inf. Commun. Smart Emerg. Technol. a Better Life, iSemantic 2024*, pp. 415–420, 2024, doi: 10.1109/iSemantic63362.2024.10762684.
- [23] N. Raj and D. Sadhya, "Securing Biometric Data over Cloud via Shamir's Secret Sharing," *Commun. Comput. Inf. Sci.*, vol. 1376 CCIS, pp. 282–292, 2021, doi: 10.1007/978-981-16-1086-8_25.
- [24] N. Santos, B. Ghita, and G. L. Masala, "Medical Systems Data Security and Biometric Authentication in Public Cloud Servers," *IEEE Trans. Emerg. Top. Comput.*, vol. 12, no. 2, pp. 572–582, Apr. 2024, doi: 10.1109/TETC.2023.3271957.
- [25] R. Shanker, S. P. Chouhan, S. Bhardwaj, A. Kumar, and M. Bhattacharya, "Fast and Efficient Real-Time Facial Recognition System Using Raspberry Pi and IoT," *2024 IEEE 8th Int. Conf. Inf. Commun. Technol. CICT 2024*, 2024, doi: 10.1109/CICT64037.2024.10899457.
- [26] N. Srinu, C. S. S. Kumar, M. Senthil, and D. B. Babu, "Smart Attendance Recording System Utilizing CNN and Edge Computing Techniques," *Proc. 5th Int. Conf. Soft Comput. Secur. Appl. ICSCSA 2025*, pp. 551–558, 2025, doi: 10.1109/ICSCSA66339.2025.11170847.
- [27] A. A. Nair, R. Adithyan, A. Unni, and S. Nalinakshan, "RFID Door Lock Access Control Systems: Trends, Technologies and Applications," *3rd Int. Conf. Intell. Data Commun. Technol. Internet Things, IDCIoT 2025*, pp. 906–912, 2025, doi: 10.1109/IDCIOT64235.2025.10914928.
- [28] V. Harish, D. Sam Chrisvin, and R. Thottungal, "Overview of RFID Security and its Applications," *2021 Int. Conf. Adv. Electr. Electron. Commun. Comput. Autom. ICAECA 2021*, 2021, doi: 10.1109/ICAECA52838.2021.9675754.
- [29] H. Mistareehi, J. Owen, and A. Aboualy, "Biometric Authentication in Identity and Access Management," *Int. Conf. Ubiquitous Futur. Networks, ICUFN*, pp. 650–652, 2025, doi: 10.1109/ICUFN65838.2025.11169800.
- [30] S. Bajpai and D. Sharma, "Moving towards 3D-biometric," in *Digital Image Security: Techniques and Applications*, CRC Press, 2024, pp. 76–92. doi: 10.1201/9781003468974-4.
- [31] S. M. Kadhim, J. K. S. Paw, Y. C. Tak, and S. T. Abd Al-Latief, "Robust Security System: A Novel Facial Recognition Optimization Using Coronavirus-Inspired Algorithm and Machine Learning," *Iraqi J. Comput. Sci. Math.*, vol. 6, no. 2, 2025, doi: 10.52866/2788-7421.1260.
- [32] M. H. Abul Hasanat, M. Fofana, M. Tamim, R. Basheer, and R. Singh, "Facial Recognition and Detection: Overcoming Pose Variations and Enhancing Accuracy with Machine Learning Implementation," *Lect. Notes Networks Syst.*, vol. 1399 LNNS, pp. 164–176, 2025, doi: 10.1007/978-3-031-91005-0_16.
- [33] T. Honcharenko, S. Dolhopolov, I. Sachenko, I. Achkasov, A. Fesan, and S. Paliy, "Automated Face Recognition System Using Convolutional Neural Network," *SIST 2025 - 2025 IEEE 5th Int. Conf. Smart Inf. Syst. Technol. Conf. Proc.*, 2025, doi: 10.1109/SIST61657.2025.11139261.
- [34] S. Z. Liu, S. Ma, H. Q. Chen, L. Z. Cui, and J. Ding, "Combining KNN with AutoEncoder for Outlier Detection," *J. Comput. Sci. Technol.*, vol. 39, no. 5, pp. 1153–1166, Sep. 2024, doi: 10.1007/s11390-023-2403-y.
- [35] N. E. I. Karabadji, A. Amara Korba, A. Assi, H. Seridi, S. Aridhi, and W. Dhifli, "Accuracy and diversity-aware multi-objective approach for random forest construction," *Expert Syst. Appl.*, vol. 225, Sep. 2023, doi: 10.1016/j.eswa.2023.120138.
- [36] D. Vetrithangam, S. Palit, A. Mehta, G. Saranya, D. Joseph, and A. Pathak, "Machine Fault Diagnosis Using Random Forest with Recursive Feature Elimination and Cross Validation," *J. Mach. Comput.*, vol. 5, no. 3, pp. 1700–1711, Jul. 2025, doi: 10.53759/7669/jmc202505134.
- [37] E. D. Handoyo, S. Santoso, and R. A. Nathasya, "Performance of Face Recognition Machine Learning Algorithms in Attendance Recording System with Limited Training Data," *J. Innov. Image Process.*, vol. 7, no. 3, pp. 707–724, Sep. 2025, doi: 10.36548/jiip.2025.3.008.
- [38] C. El Morr, M. Jammal, H. Ali-Hassan, and W. El-Hallak, "Support Vector Machine," in *International Series in Operations Research and Management Science*, vol. 334, Springer, 2022, pp. 385–411. doi: 10.1007/978-3-031-16990-8_13.
- [39] N. Yan and Z. Xu, "Research on Image Classification of RBF Kernel SVM," *Proc. 2024 Acad. Conf. China Instrum. Control Soc. ACCIS 2024*, pp. 68–71, 2024, doi: 10.1109/ACCIS62068.2024.10948583.
- [40] W. C. Cheng, H. C. Hsiao, Y. Q. Hong, and D. Y. Wang, "Masked face recognition based on facenet and genetic algorithm," *Int. J. Appl. Sci.*

- Eng., vol. 20, no. 3, 2023, doi: 10.6703/IJASE.202309_20(3).005.
- [41] S. Ammar, T. Bouwmans, and M. Neji, "Face Identification Using Data Augmentation Based on the Combination of DCGANs and Basic Manipulations," *Inf.*, vol. 13, no. 8, Aug. 2022, doi: 10.3390/info13080370.
- [42] A. Arora and S. A. Dhondiyal, "Advancements in Face Detection: A Comparative Study of Multitask Cascaded Convolutional Networks (MTCNN) and Leading Algorithms," *Lect. Notes Networks Syst.*, pp. 531–554, 2026, doi: 10.1007/978-981-96-7140-3_36.
- [43] H. K. Nayem, L. Akter, and M. E. Islam, "Deep Learning and Machine Learning Based Automated Person Identification with Intruder Detection for CCTV Surveillance," *2024 27th Int. Conf. Comput. Inf. Technol. ICCIT 2024 - Proc.*, pp. 1069–1074, 2024, doi: 10.1109/ICCIT64611.2024.11021986.
- [44] S. M. Dinesh and A. R. Kavitha, "Development of Algorithm for Person Re-Identification Using Extended Openface Method," *Comput. Syst. Sci. Eng.*, vol. 44, no. 1, pp. 545–561, 2022, doi: 10.32604/csse.2023.024450.
- [45] A. Zhalgas, B. Amirgaliyev, and A. Sovet, "Robust Face Recognition Under Challenging Conditions: A Comprehensive Review of Deep Learning Methods and Challenges," *Appl. Sci.*, vol. 15, no. 17, Sep. 2025, doi: 10.3390/app15179390.
- [46] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June-2015, pp. 815–823, Oct. 2015, doi: 10.1109/CVPR.2015.7298682.
- [47] A. Jayesh and N. Singh, "Implementation and Evaluation of a FaceNet Inspired Model for Facial Recognition," *2025 Int. Conf. Cogn. Comput. Eng. Commun. Sci. Biomed. Heal. Informatics, IC3ECSBHI 2025*, pp. 1428–1434, 2025, doi: 10.1109/IC3ECSBHI63591.2025.10990681.
- [48] Z. Jun, "The Development and Application of Support Vector Machine," *J. Phys. Conf. Ser.*, vol. 1748, no. 5, Jan. 2021, doi: 10.1088/1742-6596/1748/5/052006.
- [49] M. Mittal, H. M. Al-Jawahry, N. Varshney, S. P. Kumar, J. Michaelson, and R. A. Reddy, "Improving Support Vector Machine Performance with Advanced Kernel Methods," *Proc. Int. Conf. Contemp. Comput. Informatics, IC3I 2024*, pp. 1749–1754, 2024, doi: 10.1109/IC3I61595.2024.10828664.
- [50] Z. Nesma, B. Belkacem, and B. Ahcene, "Optimizing hyperparameters of linear SVM classifiers using the support method," *2023 Int. Conf. Decis. Aid Sci. Appl. DASA 2023*, pp. 446–449, 2023, doi: 10.1109/DASA59624.2023.10286806.
- [51] I. Markoulidakis and G. Markoulidakis, "Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis," *Technologies*, vol. 12, no. 7, 2024, doi: 10.3390/technologies12070113.
- [52] S. Gaur, N. Tiwari, S. Vyas, and M. Pandey, "Enhancing Facial Recognition Accuracy through KNN Classification with Principal Component Analysis and Local Binary Pattern," *Int. J. Electr. Electron. Res.*, vol. 12, no. 3, pp. 791–798, 2024, doi: 10.37391/ijeer.120309.
- [53] Y. N. Amirgaliyev, Z. A. Buribayev, Z. M. Melis, and A. S. Ataniyazova, "On one approach to recognizing fuzzy images of faces based on an ensemble," *Proc. - 25th Int. Conf. Circuits, Syst. Commun. Comput. CSCC 2021*, pp. 15–20, 2021, doi: 10.1109/CSCC53858.2021.00011.
- [54] D. S. J. D. Prasanna, Ragupathi, and P. Kumar, "Enhanced Face Recognition Performance Through Convolutional Neural Networks," *2nd IEEE Int. Conf. IoT, Commun. Autom. Technol. ICICAT 2024*, pp. 220–224, 2024, doi: 10.1109/ICICAT62666.2024.10922908.
- [55] A. K. Tiwari, A. Dwivedi, A. Dubey, A. Kushwaha, D. K. Dubey, and H. Tiwari, "Face Recognition Attendance Management System Using Python," *Proc. - IEEE 2024 1st Int. Conf. Adv. Comput. Commun. Networking, ICAC2N 2024*, pp. 727–730, 2024, doi: 10.1109/ICAC2N63387.2024.10895357.
- [56] A. Potdar, P. Barbhaya, and S. Nagpure, "Face Recognition for Attendance System using CNN based Liveliness Detection," *2022 Int. Conf. Adv. Comput. Commun. Mater. ICACCM 2022*, 2022, doi: 10.1109/ICACCM56405.2022.10009024.
- [57] H. Wang, T. Cai, Y. Wei, and J. Cai, "A KNN Algorithm Based on Mixed Normalization Factors," *Commun. Comput. Inf. Sci.*, vol. 2146 CCIS, pp. 388–394, 2024, doi: 10.1007/978-981-97-4393-3_31.
- [58] A. Kanan and A. Taha, "Cloud-Based Reconfigurable Hardware Accelerator for the KNN Classification Algorithm," *Proc. - 2022 14th IEEE Int. Conf. Comput. Intell. Commun. Networks, CICON 2022*, pp. 308–312, 2022, doi: 10.1109/CICON56167.2022.10008343.
- [59] X. Ma, T. Yang, J. Chen, and Z. Liu, "k-Nearest Neighbor algorithm based on feature subspace," *Proc. - 2021 Int. Conf. Big Data*

- Anal. Comput. Sci. BDACS 2021*, pp. 225–228, Jun. 2021, doi: 10.1109/BDACS53596.2021.00056.
- [60] W. Feng, C. Ma, G. Zhao, and R. Zhang, "FSRF: An Improved Random Forest for Classification," *Proc. 2020 IEEE Int. Conf. Adv. Electr. Eng. Comput. Appl. AEECA 2020*, pp. 173–178, Aug. 2020, doi: 10.1109/AEECA49918.2020.9213456.
- [61] L. Bo and H. Chunlan, "EL-RFSVM: An Ensemble Learning Framework Based on Support Vector Machine and Random Forests for Labour Resource Allocation," *2022 IEEE 2nd Int. Conf. Electron. Technol. Commun. Information, ICETCI 2022*, pp. 536–541, 2022, doi: 10.1109/ICETCI55101.2022.9832255.
- [62] X. An et al., "Killing Two Birds with One Stone: Efficient and Robust Training of Face Recognition CNNs by Partial FC," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2022-June, pp. 4032–4041, 2022, doi: 10.1109/CVPR52688.2022.00401.
- [63] A. Khalifa and A. Al-Hamadi, "A Survey on Loss Functions for Deep Face Recognition Network," *Proc. 2021 IEEE Int. Conf. Human-Machine Syst. ICHMS 2021*, Sep. 2021, doi: 10.1109/ICHMS53169.2021.9582652.

Author Biography

Anisa was born in Sumedang, Indonesia, in 2005. She is currently pursuing an Associate degree (D3) in Computer Technology at Telkom University in Bandung, Indonesia, from 2023 to 2026. During her academic period, she has been actively involved in research in the fields of artificial intelligence and machine learning, particularly in computer vision and intelligent security systems. From 2025 to 2026, she participated in an internship program at the Center of Excellence for Smart Technology Research Group (STAS RG) at Telkom University, where she contributed to research and development projects related to intelligent systems and AI-based applications. Her technical competencies include machine learning model development, deep learning implementation, image processing, and face-recognition system development for access-control applications. Her research interests include artificial intelligence, machine learning, computer vision, biometric authentication, and intelligent security systems. She actively participates in academic and applied research focusing on AI implementation for real-world security solutions.



Giva Andriana Mutiara gained a Bachelor of Engineering from Institut Teknologi Nasional Bandung, a Master of Electrical Engineering from Institut Teknologi Bandung, Indonesia, and a Doctorate from Universiti Teknikal Melaka Malaysia (UTeM) in 2022. She began as a lecturer at Telkom University's School of Applied Sciences in 2007, became an Assistant Professor in 2011, and is now an Associate Professor as of 2022. She has published over 30 research articles, most in top-ranked Scopus-indexed journals from IEEE, Springer, and Elsevier, and over 15 conference presentations, most at top-ranked international conferences. She has also written three books and holds more than 15 intellectual property rights, including three patents. Her research interests include embedded systems, wireless sensor networks, smart technologies, computer vision, and the Internet of Things. She is a member of the Association for Computing Machinery (ACM) and the director of Telkom University's Center of Excellence for Smart Technology and Applied Sciences, the Rapid Research Generator (STAS-RG). She also directs numerous smart technology research projects in domains such as tourism, military, and animal husbandry, with involvement in a variety of industries.



Muhammad Rizqy Alfarisi was born in 1991 in Bandar Lampung, Indonesia. He earned both his Bachelor of Engineering (B.Eng.) and Master of Engineering (M.Eng.) from Institut Teknologi Bandung (ITB), where he developed a rigorous technical foundation in Computer Technology and Computer Science. His academic specialization primarily focuses on the Internet of Things (IoT) and Machine Learning (ML).

Since 2020, he has served as a faculty member at the School of Applied Science, Telkom University. In 2022, he expanded his research portfolio by joining the STAS-RG Laboratory as a dedicated researcher. His current scholarly inquiries are concentrated on Smart Systems, the application of advanced Machine Learning algorithms, and the optimization of IoT architectures. His academic contributions are substantiated by a robust collection of certified intellectual property rights and publications indexed in reputable global databases.



