



# Convolutional Neural Network Method based Security Solution for Facial Recognition in ATM

Kamarunisha M<sup>\*1</sup>, Vimalanand S<sup>2</sup>, Akhtar B<sup>3</sup>, Kiruthika<sup>4</sup>, Dhivyapriya S<sup>5</sup>,  
Aarthi T<sup>6</sup>

<sup>1</sup> Department of Computer Science, Periyar University, Salem, Tamilnadu, India

<sup>2</sup> Achariya Arts and science College, Puducherry, India.

<sup>3,4,5,6</sup>Department of Computer Applications, Dhanalakshmi Srinivasan College of Arts and Science for Women (Autonomous), Perambalur, Tamilnadu, India  
nisharaj6672@gmail.com

**Abstract.** ATM machines have become a standard method for financial transactions, but they have also become vulnerable to fraudulent activities. This study presents a method that utilizes a Convolutional Neural Network (CNN) algorithm to detect ATM hammer use as a preventive measure against robberies. A novel ATM security paradigm is proposed, incorporating One-Time Password (OTP) authentication and face recognition to enhance security and consumer privacy. Face recognition eliminates the risk of fraud and duplicate card use, while OTP serves as a dynamic PIN, reducing the need for users to memorize passwords. The system employs TensorFlow for weapon detection, CNN for user identification, and a vibration sensor for detecting unauthorized machine movement. Additionally, the security framework integrates a stepper motor, buzzer, alert notification system, solenoid valve, siren, and door control mechanism. To further enhance security, the system captures and transmits images of individuals carrying weapons inside ATMs to authorized personnel via email, aiding law enforcement in suspect identification. The performance of the proposed system is evaluated using key metrics such as accuracy, specificity, F1-score, and error rate. The results demonstrate a 99.5% accuracy rate, setting a new benchmark for security in the banking sector.

**Keywords:** Automated teller machine, Convolutional neural network, One time Password, deep learning, Theft prevention

## 1 Introduction

Rapid advancements in science and technology have led to the development of new, highly secure solutions. There are still dangers, though, that could compromise this level of protection. Even though increased automation has generally had a favourable effect, fraud and theft still affect a number of financial institutions, including banks and ATMs. Because the current ATM paradigm only employs a card and PIN, there is a growth in attacks including stolen cards, statically assigned PINs, card duplication, and other dangers. In order to get

around this, a hybrid model that combines traditional elements with extra features. Applications for face recognition can be found in many domains, including criminal identification, human-computer interface, homeland security, and privacy security. [1] The facial recognition technology prevents fraudulent or stolen cards from being used to access accounts. The card alone is insufficient to access an account; a person must be present for the transaction to be completed. A facial recognition technique based on Eigenface is employed. The disadvantage of utilising the eigenface-based approach is that it can occasionally be fooled using fictitious masks or images of the account holder. [2] Facial recognition technology has sparked controversy due to allegations of privacy infringement, frequent misidentification of individuals, promotion of racial profiling and gender norms, and failure to protect individuals' rights. Photographs serve as the foundation for face recognition systems based on deep learning. Primarily used for user authentication and identification, this device detects and measures facial features from an image. Since then, facial recognition systems have been applied more broadly in robotics, smartphones, and other technological fields [3]. Facial recognition software is classified under biometrics, as it utilizes computerized facial recognition to analyse the physiological traits of humans. While iris recognition is considered more accurate than facial and fingerprint recognition, its contactless nature limits its adoption. Facial recognition systems have been widely deployed in advanced human-computer interaction, video surveillance, and automatic image indexing. The advancement of technology introduces tools designed to enhance customer satisfaction. An ATM enables customer money transfers and has been updated with a facial recognition system. The camera-based facial recognition relies on comparing features extracted from image regions with those of the query image [4]. Face recognition, on the other hand, is an essential face detection application. This type of biometric technology goes beyond merely recognizing a person's face. A computer programmer takes a digital image of a person's face (typically from a video frame) and compares it to images recorded in a database of historical data. The requirement for prompt and accurate user identification and verification increases along with the volume of electronic transactions. PINs are used for identification and security purposes in computers, bank accounts, and buildings [5]. This paper aims to improve ATM security by detecting criminal activity and preventing access to ATM resources while wearing masks or wearing lethal or non-lethal weapons. For object detection on OpenCV and Tensor flow platforms, the system makes use of the CNN algorithm and the Keras API. The alarm message is emailed to the bank and other users, protecting both citizens and law enforcement. [6] The system uses a vibration sensor to detect machine movement in case of theft. [7] The automated process doesn't require human involvement, and image processing is chosen over other biometrics due to its passive nature. This system helps prevent crime and ensures the safety of ATM users.

## 2 Background

Particularly, Convolutional Neural Networks (CNNs) are increasingly being used for commonplace tasks like character, picture, and audio recognition. In recent years, face recognition has become a major focus of intense study in CNNs. Table 1 demonstrates that the proposed method employs a deep learning algorithm for facial recognition, highlighting its advantages and disadvantages compared to the previous framework. To solve the issue of the masked face recognition procedure, the author [14] provided a dependable technique based on occlusion removal and DL-based features. Deep features were extracted from the regions that were acquired using the VGG-16 approach. Training, however, can be computationally costly and demand a large amount of processing power. The problem was fixed using a DL approach. Because of its great accuracy, it appears to be a suitable method for performing facial recognition procedures. It can take a lot of time, particularly if you're beginning from scratch [15]. Among the finest important areas of the man-machine interface, according to the author [16], is facial expression-based emotion identification. ConvNets was used as a solution to the problem. One disadvantage of the suggested approach is that feature extraction and classification are mutually optimised. An LSTM (Long Short-Term Memory) approach was used to fix the problem. Additionally, it was evaluated on simulated misregistered pictures to assess its robustness, which has not been studied for supervised change detection algorithms but is a significant factor in compromising change detection accuracy [17]. According to the author [18], it is common practice to manually assess behaviours of interest, which is laborious, limited to a few of behaviours, and inconsistent amongst studies. To fix the problem, the DeepEthogram technique was used. It calculates motion, extracts feature from motion and photos, and categorises features into behaviours using CNN. When training the model with CNN, the main issues include overfitting, ballooning gradient, and class imbalance. These problems may make the model less effective. In order to fix the problem, the DenseNet169 technique was used. Nevertheless, the suggested approach is not very accurate or efficient [19]. According to the author [20], computer vision researchers have presented numerous methods due to the importance of classification accuracy, but they still struggle with poor accuracy. The problem was solved by using a Modified Genetic Algorithm (MGA). A multiclass SVM cubic classifier is used for classification, and a non-redundant serial-based technique is used. This layer's features, however, are insufficient for an accurate classification. Low performance on the short initial dataset may therefore be the outcome of some of the currently effective tactics. A large original database can, of course, provide more details of the face photos than a small original database, but obtaining one is frequently more challenging. Furthermore, data labelling such a dataset takes a substantial amount of time and effort, even if a larger core data set can improve the model's precision and the network's ability to extend. Therefore, from a practical perspective, creating DNN-based face recognition techniques on the small original dataset is an interesting topic. Inspired by the aforementioned discoveries and considerations, this research develops a novel method for human face identification using an

Author	Year	Proposed Method	Drawback
T. Sridhar Reddy [8]	2024	Convolutional Neural Network (CNN)	Trespassing and break-ins challenge traditional security, risking privacy through identity theft and financial losses.
Mr. M. Raja Kumar [9]	2023	CNN	This model encourages fraud via stolen cards, weak PINs, inadequate encryption, and employees accessing non-encrypted customer data.
Bhagyashree Kadam [10]	2023	Linear Discriminant Analysis (LDA)	The current ATM model relies on cards and PINs, leading to more stolen card attacks.
Dr. S. Subashini [11]	2024	Support Vector Machines (SVMs)	By relieving tellers of routine tasks, banks can better focus on customer service and complex financial products.
Dr. Harish B G [12]	2023	CNN	However, this system has several limitations, including the risk of card loss, theft, or fraud.
Sahil Bajaj [13]	2022	Local Binary Pattern (LBP)	There is a need to provide more security to these ATMs.

updated dataset and a convolutional neural network. Three points can be made to summarise this paper’s primary contributions. (i) The original tiny dataset is expanded into a large one via many face image alterations. (ii) Face recognition using an innovative CNN is performed using the augmented human face database. This CNN is resistant to changes in image format. (iii) The unique methodology is proven to be superior by means of evaluations with some of the regularly used methods, after numerous tests are carried out on a shared face dataset.

## 2.1 Research Gap

- However, this system has several limitations, including the risk of card loss, theft, or fraud.
- Additionally, users may face inconvenience in case of card malfunction or forgetting their PINs.

- Losing the ATM card complicates withdrawals. We must carry the card everywhere, limiting access to just one person, which can be challenging.
- Money is easily stolen if the ATM card and PIN are accessed. Losing the card also hinders withdrawals.

### 3 Proposed Methods

The proposed method substitutes high-definition imaging cameras for touch approaches and uses a Convolutional Neural Network (CNN) algorithm to identify criminal conduct. It utilizes vibration sensors to detect machine movement and multiple layers to identify theft. Security is enhanced through features like buzzers and alert notifications. Additionally, the technology employs face recognition and the Internet of Things (IoT) to identify pounding in ATMs. The steps that the suggested method follows are illustrated in Figure 1.

1. Determine whether a person's behavior is normal or abnormal by observing their actions and the visibility of weapons.
2. No action is required if the behavior is normal and no weapon is displayed.
3. If abnormal behavior or a visible weapon is detected, proceed with the detection procedure.

The following steps are taken if a weapon is visible:

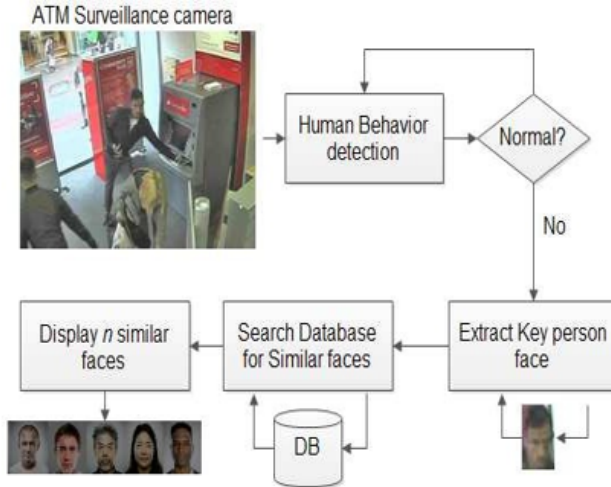
- Extract the important person's face.
- Match their face against the suspect database.
- Extract the closest faces for further examination.
- A human can manually investigate the identified individuals.

The camera captures the behavior, and the face is then captured and searched for a potential suspect. The image acquisition phase (capture) is the first step in the system. Next, the face image is extracted from the overall image, and finally the image is aligned and measured (the angle of the face is adjusted to match the angle of the camera). To achieve better matching, the image's key features are extracted. The process of matching the requested image with the image store is carried out. Lastly, a report was provided by including the photos that most closely resemble the captured image.

## 4 Object Detection

### 4.1 Image processing

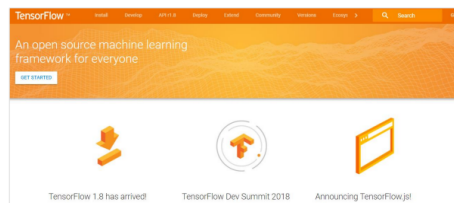
Image processing is a subset of signal processing that involves manipulating images to produce better images or extract valuable data. It is employed to identify actual objects in the world, such as cars, bottles, helmets, etc. An image will be the input, and the output could be the image itself or certain traits or qualities related to it. Additionally, it is employed in the detection of faces, mouths, and eyes. Images, movies, and live stream data are all used in the image processing process. This image processing programme is frequently utilised in surveying, security, and other fields.



**Fig. 1.** Flowchart of the proposed method

## 4.2 Tensor Flow

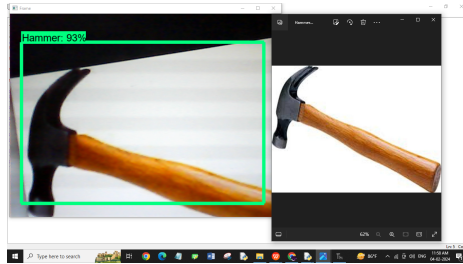
TensorFlow is an open-source predictive modelling framework designed by the Google Brain team that is used for dataflow programming in many applications. It is an open-source numerical calculation library used to build multi-layer, large-scale neural networks. Classification, perception, discovery, and prediction are its primary uses. Thus, we use Tensor flow to implement object detection during this project TensorFlow is an open-source predictive modelling framework



**Fig. 2.** Tensor flow

designed by the Google Brain team that is used for dataflow programming in many applications which is described in figure 2. Some of Tensor flow features are listed below: It has a function that uses multi-dimensional arrays known as tensors to design, optimise, and compute mathematical equations with ease.

- It incorporates machine learning and deep neural network programming support.
- TensorFlow leverages GPU computing to automate management; it has a highly scalable property for computation with different datasets. It also has a special feature that optimizes data consumption and RAM usage.
- Using CCTV footage, the composite model, which employs Recurrent Neural Network (RNN) algorithms, identifies weaponry.
- Variations in viewing angles and occlusions from the firearm carrier and other individuals in the vicinity add complexity to the detection process.



**Fig. 3.** Weapon detectors

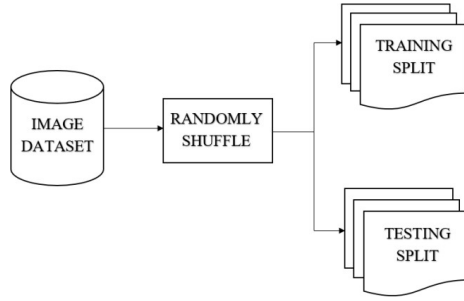
The development of automatic weapon detectors (figure 3) made possible by CNN techniques has improved the status of handgun detection.

## 5 Workflow for Object Detection

Although each Object Detection Algorithm operates differently, they all have a common perspective. Using feature extraction, one can control an image's class by taking features out of the input photos at indications. Using Deep Learning, Open CV, or Mat Lab, for example.

### 5.1 Processing of data

The training process for object detection requires the photos. The photos ought to have the same dimensions, and the training method uses the same dimensions as an input image. To set each variable apart from the others, distinct values were assigned to the stored variables. The features of the trained images are stored in the variable values. Like all deep learning algorithms, the training algorithm receives the photos, divides them into a larger number of parts, and then processes each individual bit of image data. (figure 4) Each segment of the image is processed, and the treated segment is then saved as a ". model" file and transformed into a variable.



**Fig. 4.** Data processing method

## Convolutional Neural Network (CNN)

One of the main subtypes of neural networks used for image identification and categorization is CNNs. CNNs are widely used in many different domains, including face recognition and object identification. A CNN image classification system receives an input image, examines it, and assigns it to one of many groups (such as knife, pistol, steel rod, etc.). A computer interprets an input image as a set of pixels based on the image resolution. The image is represented as  $h \times w \times d$ , depending on the image quality. Unlike Artificial Neural Networks (ANNs), CNN shares characteristics with Recurrent Neural Networks (RNNs) and applies a single filter to various regions of an image. Its hidden layers are convolutional instead of the recurrent connections found in RNNs. As activation functions, pooling and convolution functions are employed.

### Convolution Operation

Applies filters (kernels) to the input image to detect features, as represented in Equation (1):

$$O^l = f(I^{l-1} * K^l + B^l) \quad (1)$$

where:

- $O^l$  is the output of convolution,
- $I^{l-1}$  represents the pixel values of the input image,
- $l$  is the layer index,
- $K^l$  is the convolutional kernel,
- $B^l$  is the bias term.

### ReLU Activation Function

A non-linear activation function that sets all negative values to zero, as shown in Equation (2):

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

## Max Pooling

The max operation is performed over a region (e.g.,  $2 \times 2$ ) to reduce the spatial size of the representation, decreasing computational cost and preventing overfitting. It is formulated in Equation (3):

$$P^l = \max(O^l) \quad (3)$$

## Fully Connected Layer

Following the above operations, a fully connected layer is computed in Equation (4):

$$F^l = W^l O^l + B^l \quad (4)$$

where  $W^l$  is the weight matrix. This layer connects every neuron in one layer to every neuron in the next layer, converting the 2D feature maps into a 1D vector.

## Softmax Layer

The softmax function converts the network's outputs into probabilities:

$$S^l = \frac{e^{O^l}}{\sum e^{O^l}} \quad (5)$$

## Loss Function

The loss function measures the error between predicted and true labels:

$$L = - \sum y \log(\hat{y}) \quad (6)$$

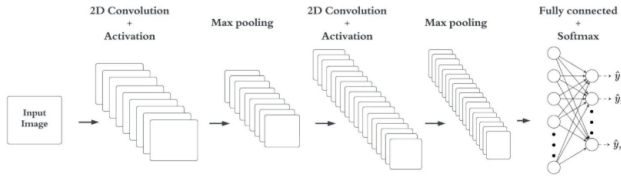
## Weight Update Rule

Updating the weights using the learning rate  $\eta$ :

$$W^{l+1} = W^l - \eta \frac{\partial L}{\partial W^l} \quad (7)$$

where:

- $\eta$  is the learning rate,
- $\frac{\partial L}{\partial W^l}$  is the gradient of the loss function with respect to the weights.



**Fig. 5.** Convolution Neural Network

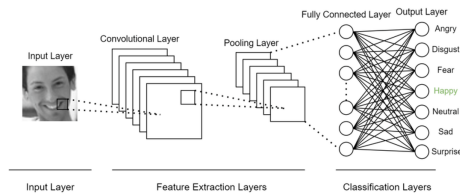
### Key CNN Components

- **Convolution:** An input image is processed by applying a filter (kernel) over it.
- **Pooling:** Selecting the region's maximum value is known as max pooling, and vice versa.

In order to extract the appropriate features from the data, CNN automatically learns the filter. (Figure 5) In contrast to ANN, it is able to record spatial characteristics, or the arrangement of pixels.

### 5.2 Face recognition technique

**CNN facial recognition model** The CNN approach used in this work improves the precision of facial recognition in photos. The input data, network breadth, and full connection layer of the model differ from the standard LeNet-5 approach, despite the model's general structure being comparable.



**Fig. 6.** CNN face recognition model

### CNN Model Architecture

**Figure 6:** The proposed CNN consists of two convolutional layers (C1 and C2) and two pooling layers (S1 and S2), arranged in an alternating manner as follows:

$$C1 \rightarrow S1 \rightarrow C2 \rightarrow S2$$

The outcome of this architecture is a 40-dimensional vector representing 40 distinct faces identified through multi-label classification using the sigmoid function.

## CNN Input and Feature Maps

The CNN model receives a normalized facial image as input from a single feature map in the input layer. The first convolutional layer (C1) consists of six feature maps, where each neuron applies a  $5 \times 5$  convolution kernel that is initialized randomly.

## Pooling and Convolution Layers

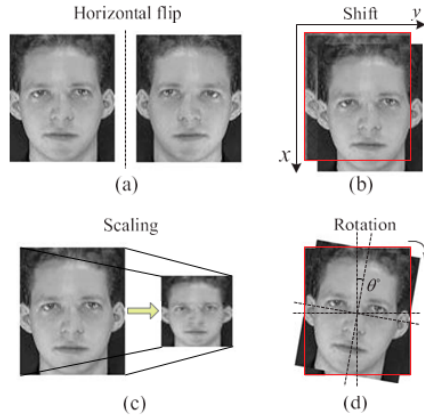
- The first pooling layer (S1) utilizes the output from the previous convolutional layer (C1) to generate six feature maps.
- In order to prevent overlap in the receptive fields, the C1 layer employs a mean convolution kernel that links each element in the feature map.
- The second convolutional layer (C2) and the second pooling layer (S2) inherit the same feature mapping strategy and computation as their preceding layers.

## Fully Connected Layer

Finally, a fully connected single-layer perceptron connects the output layer to the S2 layer, facilitating classification.

## 5.3 Dataset and its enhancement

The 400 photos of faces shot by 40 different people make up the Olivetti Research Laboratory's (ORL) face collection. Ten images of each person's face in various settings are added to the collection. Each image is stored in the BMP and PGM file formats, with a dimension of 92 by 112 pixels. Compared to other face datasets like the MIT or Yale face datasets, the popular ORL face dataset is easier to label. However, not enough images exist to train the DNN to identify faces correctly. To tackle this problem, illustrates the use of four distinct data augmentation techniques—horizontal shift, flip, rotation, and scaling to increase the dataset's image size. As demonstrated by Figure 7, the dataset can actually be greatly enhanced by adjusting the augmentation methods' parameters. Following the aforementioned methods, the dataset is enhanced by 1000 times in this study. Before the photographs are entered into the face recognition system, they are first resized, normalised, and tagged. It is foreseeable that the enlarged dataset will increase the system's resilience while simultaneously lowering the likelihood of over-fitting.



**Fig. 7.** Four methods for data augmentation

## 6 Experimental Results and Analysis

This section includes various experiments to assess the created facial recognition method's performance; the method's superiority can be confirmed by comparing it to certain commonly used approaches. MATLAB 2018a is used to conduct the experiments using a PC equipped with an Intel Core i7-6700 CPU. It is evident that as the number of training examples increases, face recognition accuracy increases as well, remaining steady for varying numbers of test samples. The

**Table 2.** Simulation Parameters

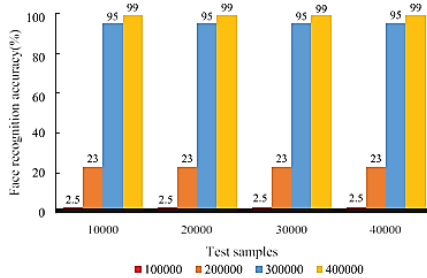
Simulation	Value
Dataset Name	ORL Dataset
No of Dataset	4000
Training	2987
Testing	1013
Language	MATLAB 2018

face recognition network is trained using various sample sizes to evaluate the impact of the enhanced dataset. (Table 3) This indicates that a large number of sample characteristics may be added to the augmented dataset, improving network training and producing a high degree of face recognition accuracy. The suggested method's outcomes are contrasted with those of a few other face recognition techniques, such ANN (artificial neural network), that are based on the ORL face dataset and have been published in the literature. Table 4 lists their face recognition accuracy to further substantiate the superiority of the deployed

**Table 3.** Comparison of Accuracy of Face Recognition

No of Images	SVM	VGG-16	DenseNet169	MGA	CNN
1000	72	76	80	85	89
2000	75	82	86	89	91
3000	79	88	90	93	95
4000	82.3	93.43	94	96	97.23

methodology. As previously stated, the face recognition network exhibits robustness against varying quantities of test samples, hence enabling the maintenance of a constant test sample in subsequent tests. In the meanwhile, more epochs will help improve accuracy when dealing with various training data. That is, there is a positive correlation between the number of training samples and epochs and the accuracy of face recognition. In fact, additional epochs and training data in the studies can potentially support this result. MSE is typically applied to the neural network’s cost function. A lower mean square error (MSE) indicates strong network model accuracy. Several experiments are conducted using various numbers of training samples and epochs. In the meanwhile, several test



**Fig. 8.** Accuracy of face recognition using various numbers of test and training samples

sample counts are used in the neural network test; for additional information, refer to Figure 8. Figure 9 illustrates how, in every situation of training samples, the MSE will drop as the number of training iterations increases. Furthermore, the MSE results can support the findings of the earlier studies. Based on the ORL face dataset’s face recognition accuracy as 97.3%, it is clear that previous approaches as VGG-16, MGA, DenseNet-169 and SVM-based techniques outperform CNN-based methods. Nonetheless, the proposed approach, which combines CNN with the enhanced face dataset, is able to achieve a higher level of face recognition accuracy in comparison to earlier techniques. Consequently, it may be claimed that the enhanced face dataset’s abundance of available attributes may increase the accuracy of face recognition. Figure 10 shows that the F1 score

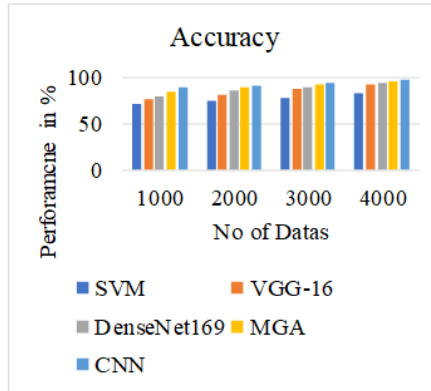


Fig. 9. Face recognition accuracy 2

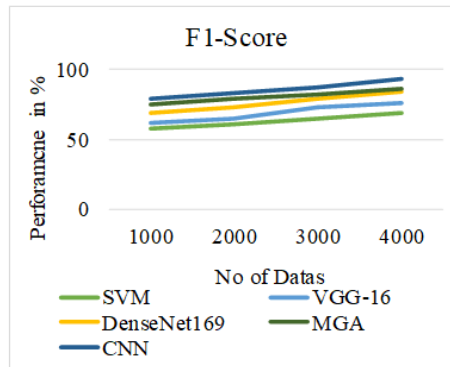


Fig. 10. Analysing the F1 score

of the previous methods MGA is 86%, SVM is 69.2%, DenseNet-169 is 84.2%, VGG-16 is 76.2% and the analysis F1 score of the CNN model is 93.6%. The CNN has a higher score rather than previous methods. A high F1 score indicates a method's ability to achieve high specificity. A high F1 score means the model can accurately identify the objects. The network is now trained with varying sample sizes and epochs; Figure 10 summarises the outcomes. It is evident that when more training samples from various epochs are used, the accuracy of face recognition will rise. Figure 11 shows that the specificity performance of the

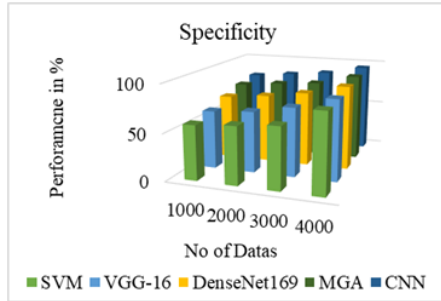


Fig. 11. Analysing the Specificity

previous methods MGA is 92%, DenseNet-169 is 89%, VGG-16 is 85%, SVM is 83.4% and the analysis specificity of the CNN model is 95.3%%. The CNN has a better precision rather than previous methods. Figure 12 shows that the error

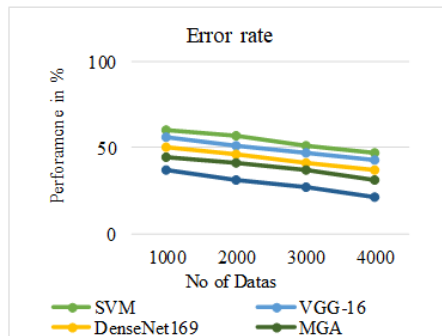


Fig. 12. Analysis of the error rate

rate performance of the previous methods MGA is 31%, VGG-16 is 43%, SVM is 47%, DenseNet-169 is 37% and the analysis error rate of the CNN model is 21%.

Lower error rates mean the proposed method is more accurate in identifying objects, reducing the chances of misdiagnosis and the number of false positives.

## 7 Conclusion

This research provides an open-source, multifunctional ATM security system based on the CNN model and IoT technologies. It ensures a safer and more dependable ATM environment by detecting warm things entering ATM cabins and offering a tool to arrest fraudsters. This work uses transformations such as flip, shift, scaling, and rotation to provide a novel method for face identification on a tiny dataset. The CNN implementation then makes advantage of the enhanced dataset. Tests demonstrate the method's superiority and efficacy over other facial recognition techniques. More difficult issues like picture recognition, signal processing, and fault detection will be the main focus of future research. The results obtained from the ORL face dataset indicate a face recognition accuracy of 97.3%. This proposed CNN method demonstrates that earlier methodologies, including VGG-16, MGA, DenseNet-169, and SVM-based techniques, perform better than recent CNN-based ones.

## References

1. Abousamra, R.; Hosam, O. Quantitative Classification of Cognitive Behaviors for Industrial Projects' Managers in the MENA Region. *8th International Conference on Information Technology Trends (ITT)*, IEEE, 2022. doi:10.1109/itt56123.2022.9863957.
2. Hosam, O.; Abousamra, R. Enhancing Deep Training of Image Landmarking with Image CAPTCHA. *8th International Conference on Information Technology Trends (ITT)*, IEEE, 2022. doi:10.1109/itt56123.2022.9863967.
3. Dhanusree S, Saeeda Aseema M, Santhiya R, Mathumitha M. "Face Biometric Authentication System for ATM Using Deep Learning." *International Journal of New Innovations in Engineering and Technology*, ISSN: 2319-6319, Volume 24, Issue 1, March 2024.
4. Binusha S, Susai Mary Susila A. "OPENCV-BASED ATM SECURITY SOLUTION WITH FACIAL RECOGNITION CAPACITY." *International Research Journal of Modernization in Engineering Technology and Science*, Volume:06, Issue:08, August-2024. DOI:10.56726/IRJMETS61262.
5. M. Srinivasa Sessa Sai, Ranjih Kumar Gatla. "Development of Facial Detection System for Security Purpose Using Machine Learning." *E3S Web of Conferences* 564, 07002 (2024). DOI:10.1051/e3sconf/202456407002.
6. BinDarwish, Abdulaziz, Salim Alhammadi, and ABDULLAH SALEHI. "Crime Detection and Suspect Identification System." (2023).
7. Lu, Peng, Baoye Song, and Lin Xu. "Human face recognition based on convolutional neural network and augmented dataset." *Systems Science & Control Engineering* 9.sup2 (2021): 29-37.
8. T. Sridhar Reddy, B. Sujatha, V. G. V. Prasanna Kumar. "Facial Recognition Reinvented: Deep Learning Based Security Alert System." July 2024. DOI:10.2991/978-94-6463-471-6\_117.

9. M. Raja Kumar, Medavarapu Tejaswi, Kurukuri Haritha. "Deep Learning and Image Processing Based ATM Security and Identifying the Face." *Journal of Science & Technology (JST)*, 8(4), 46–52 (2023). DOI:10.46243/jst.2023.v8.i04.pp46-52-01.
10. Bhagyashree Kadam, Kishor Dukr, Vaibhav Kumbhar, Shivraj Sakunde, Akash Jadhav. "Real-Time Face Recognition in ATM for Security System." *International Journal of Research Publication and Reviews*, Vol 4, no 5, pp. 6055-6058, May 2023.
11. Dr. S. Subashini, S.A. Nanthitha Sri, S. Ponsree, R.R.Riduvashini. "Multi-Factor Secure ATM Access With Face Recognition Using Deep Learning." *IJCRT*, Volume 12, Issue 11, November 2024, ISSN: 2320-2882.
12. Dr. Harish B G, Chetankumar G S, Akhila N, Rachana S P, Pavan S Hatti, Lakshmi D M. "Deep Learning-Based Card-Less ATM Using Fingerprint and Face Recognition Techniques." *JETIR*, July 2023, Volume 10, Issue 7. www.jetir.org (ISSN-2349-5162).
13. Sahil Bajaj, Sumit Dawda, Pradnya Jadhav, Rasika Shirude. "Card-less ATM Using Deep Learning and Facial Recognition Features." *Research Article*, Volume 12, Issue 4 (2022).
14. Hariri, W. "Efficient Masked Face Recognition Method During the COVID-19 Pandemic." *SIViP 16*, 605–612 (2022). DOI:10.1007/s11760-021-02050-w.
15. KH Teoh et al. "J. Phys.: Conf. Ser. 1755 012006" (2021). DOI:10.1088/1742-6596/1755/1/012006.
16. Chowdary, M.K., Nguyen, T.N., Hemanth, D.J. "Deep Learning-Based Facial Emotion Recognition for Human–Computer Interaction Applications." *Neural Comput & Applic* 35, 23311–23328 (2023). DOI:10.1007/s00521-021-06012-8.
17. Liu, T., Yang, L., Lunga, D. "Change Detection Using Deep Learning Approach with Object-Based Image Analysis." *Remote Sensing of Environment*, 256, 112308 (2021). DOI:10.1016/j.rse.2021.112308.
18. James P. Bohoslav, Nivanthika K. Wimalasena, Kelsey J. Clausing, Yu Y. Dai, David A. Yarmolinsky, Tomás Cruz, Adam D. Kashlan, M. Eugenia Chiappe, Lauren L. Orefice, Clifford J. Woolf, Christopher D. Harvey. "DeepEthogram, a Machine Learning Pipeline for Supervised Behavior Classification from Raw Pixels." *eLife 10:e63377* (2021). DOI:10.7554/eLife.63377.
19. Zhang, Q., Yang, Q., Zhang, X., Bao, Q., Su, J., Liu, X. "Waste Image Classification Based on Transfer Learning and Convolutional Neural Network." *Waste Management*, 135, 150-157 (2021). DOI:10.1016/j.wasman.2021.08.038.

**Open Access** This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

