Indian Journal of Information Sources and Services
ISSN: 2231-6094 (P), Vol.13, No.2, 2023, pp.7-13
© The Research Publication, www.trp.org.in
DOI: https://doi.org/10.51983/ijiss-2023.13.2.3752

Smart Transaction through an ATM Machine using Face Recognition

Akshay Kumar[1], Pooja Joshi[2], Anju Bala[3], Pravinkumar Sudhakar Patil[4], Dilip Kumar Jang Bahadur Saini[5] and Kapil Joshi[6]

[1&6]Department of CSE, Uttaranchal Institute of Technology, Uttaranchal University, Uttarakhand, India
[2]Department of CSE, HSST, Swami Rama Himalayan University, Uttarakhand, India
[3&5]School of Engineering and Technology, Pimpri Chinchwad University, Pune, Maharashtra, India
[4]Sharad Institute of Technology College of Engineering, Yadrav-Ichalkaranji, Maharashtra, India
E-mail: akshaydhiman685@gmail.com, poojajoshi.baloni@gmail.com, anjub.kathuria@gmail.com, pspatil@sitcoe.org.in, dilipsaini@gmail.com, kapilengg0509@gmail.com

(Received 13 September 2023; Revised 23 October 2023, Accepted 30 October 2023; Available online 6 November 2023)

Abstract - **Automated Teller Machines (ATMs) are a convenient way for people to access their cash and banking services. However, traditional ATMs are susceptible to fraud, such as card skimming and PIN theft. Face recognition technology has the potential to improve ATM security by providing a more secure and convenient way for users to authenticate themselves. This research paper proposes a system for smart transaction through ATM machines using face recognition. To proceed transactions user, need to enter their registered mobile number on which an OTP will be generated subsequently the camera mounted on top of the ATM will capture user's face which will then be compared with the picture associated by the contact number using custom trained YOLO algorithm. Upon successful matching user need to verify OTP received and if verified they can access the ATM functionalities. The dataset used for this model i.e., Contact number and customer photograph will be collected by the bank at the time of customer account opening. The system was evaluated using a dataset of over 10,000 random face images. The results showed that the system was able to achieve an accuracy of over 99% in authenticating users. The system was also able to detect and prevent unauthorized access attempts. The proposed system has the potential to revolutionize the way that people use ATMs. It can make ATMs more secure, convenient, and fraud-resistant.**

Keywords: **Smart ATM, Card Less Transaction, Face Recognition, Artificial Intelligence, Secure Transaction**

I. INTRODUCTION

Technology is advancing rapidly, and with it comes the need for better security measures. While automation has many benefits, it is important to address the ongoing threats of fraud and theft faced by financial institutions, such as banks and ATMs. Adequate security measures are essential to protect both financial institutions and their customers. Face recognition technology offers a promising solution to these challenges, as it provides enhanced security levels and several advantages over traditional authentication methods. The existing ATM model, which relies on a combination of a card and a PIN, is vulnerable to various forms of attack, such as stolen cards, static PINs, card replication, and PIN hacking. The "ATM Security System based on Face Recognition" addresses these challenges by using a combination of facial recognition and OTP to significantly enhance overall security. When users enter the ATM, they must input their registered mobile number, which triggers the generation of an OTP. Simultaneously, webcams automatically capture live images [1], subsequently compared to the stored facial image linked with a mobile number. If a match is detected, the user is prompted to enter their OTP [2]. Upon successful matching, they are granted access to all ATM functions [3]. This incorporation of facial recognition and OTP [4] strengthens security measures and enhances the user experience by providing improved protection against fraudulent activities [5,6].

OpenCV: OpenCV is a popular open-source computer vision library with a wide range of capabilities and applications. One of its most notable applications is in the field of facial recognition [7]. OpenCV's powerful algorithms and advanced image-processing techniques enable developers to build accurate and efficient facial recognition systems [8]. These systems can detect and analyze facial features such as the eyes, nose, and mouth [9], allowing for reliable identification and verification of individuals [10]. With its comprehensive suite of functions and resources, OpenCV empowers researchers and developers to explore the potential of facial recognition technology [11] and develop innovative solutions in areas such as security, access control, and human-computer interaction [12].

II. LITERATURE REVIEW

Automated teller machines (ATMs) currently use a combination of a card and a user-generated PIN for authentication. However, this method is vulnerable to fraud, especially if the PIN is stored on the user's mobile phone SIM card. If the entered PIN matches the stored authorization PIN, the user is granted access to all ATM services. However, if there is a mismatch, the authentication process fails, and the user typically has two more attempts to enter the correct PIN. This system is susceptible to fraud, including cases where hackers gain access to the security system. To address these concerns, Mohammed Kabiru *et al.,* proposed a face recognition technique that incorporates image enlargement [13]. Their goal was to improve face recognition accuracy using the Principal Component Analysis (PCA) algorithm. Face recognition has a wide range of applications in the security and entertainment industries. However, recognizing individuals in surveillance systems can be challenging due to poor-quality footage caused by factors such as camera distance and angles. To overcome this challenge, it is important to enhance image resolution for accurate personal identification. Three popular image enlargement methods were examined in the study: bilinear, bicubic, and nearest neighbor. Using the previously described methods, each down sampled image is enlarged once the input image has been down sampled into several resolutions. To identify specific people, the resulting magnified image is checked with the image database.

LhiZiu [14] proposed a new way to recognize faces that is more accurate than previous methods. His system works by first converting all of the images to grayscale and then equalizing the histograms. Next, the system extracts three different types of features from the images: LBP, Gabor, and HOG. These features capture different aspects of the face, such as texture and shape. Once the features have been extracted, the system uses a special algorithm to compare them to a database of known faces. This algorithm takes into account the fact that faces can vary in appearance due to factors such as lighting and expression. As a result, the system can accurately identify faces even in challenging conditions. Lhi Ziu's system was tested on two large datasets of face images, and it achieved a higher accuracy rate than any other single-feature recognition method.

A Face Recognition system using Scale Invariant Feature Transform (SIFT) and Dominant Rotated Local Binary Pattern (DRLBP) for feature extraction was presented by M. Sushama *et al.,* [15]. Numerous mathematical and computational methods, such as SIFT and DRLBP, have been developed to improve facial recognition performance. The authors of this paper provide a novel Artificial Neural Network (ANN)-based method for categorizing human faces. After pre-processing, SIFT is used to extract features from the facial photos. After then, a Back Propagation Network (BPN) is used to detect faces of people. When SIFT and DRLBP are applied together, accuracy is higher than when they are used alone. For face recognition tasks, the combination of these two approaches yields encouraging results. By superimposing training photos, Ravi Kanth Kumar *et al.,* presented a way to improve face recognition [16]. In comparison to conventional methods, their methodology yields an approximate 43 percent gain in accuracy by providing an overlaid version of all relevant photos. The algorithm highlights the superimposition strategy's importance in the face recognition domain. A series of photos is used to train the cascades function of a HAAR feature-based classifier. This novel method shows promise in enhancing the functionality of face recognition algorithms.

Principal Component Analysis (PCA) is one of the techniques used in Maliha Khan and Rani Astya's OpenCV-based face detection and recognition system [17]. Because PCA represents face data economically in a small feature space, it is essential in lowering the amount of storage needed. The process of creating a projection of the self-space for facial identification involves building a wide 1-D pixel vector from the 2-D face image and using PCA to extract the key components. The proper feature space is determined by identifying the eigenvalues and eigenvectors of a matrix that is generated from a set of fingerprint images. Programming was built utilizing OpenCV, HAAR Cascade, Eigen face, Fisher face, LBPH, and Python algorithms in order to implement a real-time face recognition system.

III. METHODOLOGY

To implement it we required a high-quality camera and a Raspberry Pi. The camera will be positioned at the top of the ATM to ensure that the user's face is fully visible. The Raspberry Pi will be used to power the camera and to compare the captured image with a database of individuals associated with bank accounts. The database used for comparison with live images captured during the transaction will be created by the bank at the time of account opening. Customers will be required to provide recent pictures of them when they open a new account. This image will be stored in the bank's database and used to verify the customer's identity when they use ATM. To withdraw money or perform other transactions, customers simply need to enter their phone number and look into the camera. The ATM will then compare the customer's face to the database of customer images. If there is a match, they need to enter the OTP received on their registered mobile number; if it matches the customer will be granted access to their account.

A. Face Detection and Recognition

To use the new ATM security system, users must first enter their registered mobile number. This will generate a one-time password (OTP), which will be sent to the user's phone. The user will then be prompted to look into the camera installed on the ATM. The system will then compare the user's face to the photograph associated with the linked phone number, instead of searching the entire database. Once the user's face is successfully recognized, they will be prompted to enter the OTP. Upon confirmation, they will be granted access to all ATM services. This additional security measure helps to address the challenges posed by cardless transactions. To further protect against theft, where criminals force users to access their accounts, a simple yet effective method is implemented. If the system detects multiple faces, the account will be temporarily locked. This ensures that transactions can only proceed when the authorized user has exclusive access to the system.

In simpler terms, the new ATM security system works like this (Figure 1).

1. Enter your registered mobile number.
2. Receive an OTP on your phone.
3. Look into the camera.
4. Enter the OTP.
5. Use the ATM.

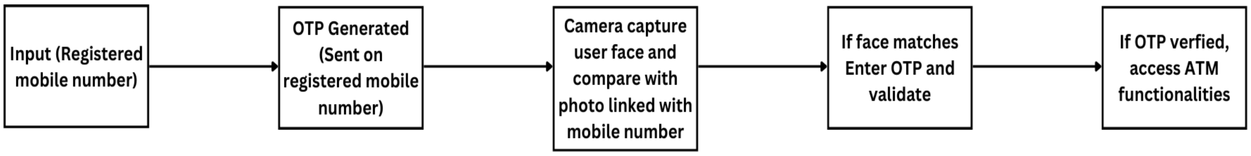
Input (Registered mobile number)
OTP Generated (Sent on registered mobile number)
Camera capture user face and compare with photo linked with mobile number
If face matches Enter OTP and validate
If OTP verfied, access ATM functionalities

Fig. 1 Block diagram of the model

If multiple faces are detected, the account will be locked. This helps to protect against theft and other fraudulent activities.

B. Introduction to YOLO Algorithm

Prior to the publication of YOLOv1, the R-CNN algorithm reigned as the leading approach in target detection. R-CNN exhibits high accuracy in detection but suffers from a two-stage structure that hampers real-time processing, thereby receiving criticism. To overcome this impasse and meet the demand for faster target recognition, Joseph Redmon, Santosh Divvala, Ross Girshick, and others introduced a one-stage target recognition approach in 2016. This method, known as “You Only Look Once” (YOLO), boasted high recognition speed, processing [18] 45 frames per second and enabling instantaneous execution. YOLO achieved this by converting object recognition into a regression function using comprehensive graphical data as network input, thereby obtaining object locations and bounding box groups via a neural network. YOLOv1 adopted a divide-and-rule strategy by dividing an image into a grid of 7x7 cells. Each cell within this grid played a crucial role in predicting the target’s location, specifically the grid cell where its midpoint falls. Contrasting with Faster R-CNN’s utilization of a Region Proposal Network (RPN) to determine regions of interest, YOLOv1 achieved comparable accuracy without the need for an additional RPN network. By partitioning the image into 49 grids, each grid aligned with a target’s area of interest, YOLOv1 eliminated the requirement for designing and developing an extra RPN network, making it a streamlined one-stage network [19].

In 2018, Redmon further improved YOLO with YOLOv2. This iteration replaced the darkNET-19 network for feature extraction with darknet-53 and implemented a feature pyramid network structure to enable multi-scale detection. Additionally, the softmax classification method was replaced with logistic regression, ensuring high accuracy and real-time performance for target detection. With each subsequent YOLO version, such as YOLOv3, the backbone network’s enhancements played a vital role in improving performance. YOLOv3 introduced Darknet-53 and the lightweight Tiny-Darknet. When both precision and speed are desired, Darknet-53 serves as the ideal backbone. However, if faster detection speed is the priority and accuracy can be compromised to some extent, Tiny-Darknet is an excellent choice. In summary, the flexibility of YOLOv3 has made it highly popular in practical engineering applications.

C. Introduction to Facial Recognition Algorithm

The implementation process of YOLO can be outlined as follows:

1. Segmentation: The image is divided into a grid of size S*S cells. Each grid cell plays a crucial role in detecting objects when the center of an object falls within that particular grid.

2. Bounding Box Prediction: Each grid predicts B bounding boxes, and for each box, five parameters (x, y, width, height) along with a confidence score are estimated. Class Identification: Additionally, each grid cell assigns a category to the detected object, denoted as C classes, providing sector-specific information.

In general, YOLO utilizes a grid structure with SS cells, where each grid is responsible for detecting B-bounding boxes and producing C network-like outputs. As a result, the overall output tensor size for each grid is SS*(5B+C) Coordinate prediction loss, confidence prediction loss, and category prediction loss are the three parts of the loss function. The coordinate prediction loss measures the discrepancy between the predicted and ground truth bounding box coordinates. The confidence prediction loss penalizes incorrect confidence scores, with specific adjustments made for grid cells that do not contain any targeted objects.

The category prediction loss quantifies the difference between predicted and actual object categories, employing a squared difference or error. To account for variations in bounding box sizes, the square root of the width (W) and height (H) is taken during error calculation. This adjustment aims to mitigate the issue of predicting large bounding boxes compared to smaller ones. Given the higher significance of positioning accuracy over classification accuracy, a penalty of λ coord = 5 is applied to positioning errors. During training, grid cells that do not contain any objects reduce the confidence scores of bounding boxes within those cells to zero. This adjustment prevents the model from assigning high confidence to frames without targets, which could lead to model instability. For such frames, the loss associated with confidence prediction is reduced to λ noobj = 0.5.

The convolution layer plays a vital role in convolutional neural networks (CNNs). It performs calculations on pixel matrices of an image, generating activation maps that store different attributes of the provided image. By applying feature detectors, convolution identifies patterns in the data matrix and generates various versions of the image. Through back propagation training, the convolutional model determines the minimum error for each layer. Based on the minimum error settings, the depth and characteristics are assigned as depicted in Figure 1, showcasing the working principle of convolution. The convolution process involves an image matrix and a set of feature detectors, producing an activated figure or features. The values in the data and features located at the same positions (i.e., values greater than or equal to 1) are preserved, while the remaining values are discarded. The size of the feature detector varies depending on the type of CNN. Lastly, the detection neural network structure is described.

An introduction to the dataset is provided, including the number of pictures and face categories, as well as the display of facial features within the dataset. Facial recognition algorithms can be divided into geometric methods, which focus on luminosity-based distinguishing features, and statistical methods, which extract values from image data for matching and template comparison. These algorithms can further be categorized as feature-based models or holistic models, with the former focusing on specific facial markers and their spatial parameters, while the latter treats human faces as a single entity. Convolutional neural networks (CNNs) are a breakthrough in the field of artificial intelligence, specifically within deep learning. CNNs have gained popularity and achieved impressive results in various areas, such as computer vision, natural language processing (NLP), and image classification tasks. CNNs consist of multiple layers, including convolution and pooling layers, which learn to identify different visual features. The Passthrough layer, similar to the Shortcut layer in the ResNet network, takes the high-resolution feature map from the front and connects it to the lower-resolution feature map from the rear, doubling its dimensions. This Passthrough layer extracts every two feature maps from the front layer, resulting in new feature maps of 26x26 and 512, which are then connected to the 13x13 feature maps of 1024. These forms feature maps of 13x13x3072, which are subsequently convolved for prediction based on the feature maps. Figure2 illustrates this process.

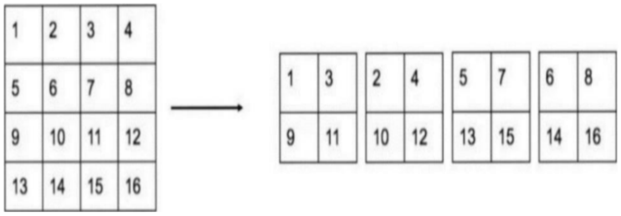

Fig. 2 Convolution Operation

YOLO utilizes a fully connected layer to forecast values after extracting key features using a convolutional grid. As seen in Figure 8, the network design is based on the GooLeNet model, which has 24 convolutional layers and 2 fully linked layers. Both the convolutional and fully connected layers receive the application of the Leaky ReLU Activation function, which uses 1x1 convolutions for channel reduction and 3x3 convolutions for feature extraction. However, the last layer uses a linear activation function. The output of the frame network is a tensor with a size of 30x7x7, as discussed previously and illustrated in Figure 3. Within each cell of the grid, the top 20 attributes represent class probability values. The following two elements correspond to bounding box confidence, and multiplying those yields the category.

The remaining eight elements represent the bounding box coordinates (x, y, width, height). It’s worth noting that the arrangement of the confidence degrees and bounding box coordinates can be separated rather than following the (x, y, width, height) order. This separation is primarily for ease of calculation. In other words, the 30 elements within each cell can be arranged randomly but separating them simplifies the extraction of each part. To clarify, the determined value of the network is a two-dimensional tensor P with dimensions [height, 7x7x30].

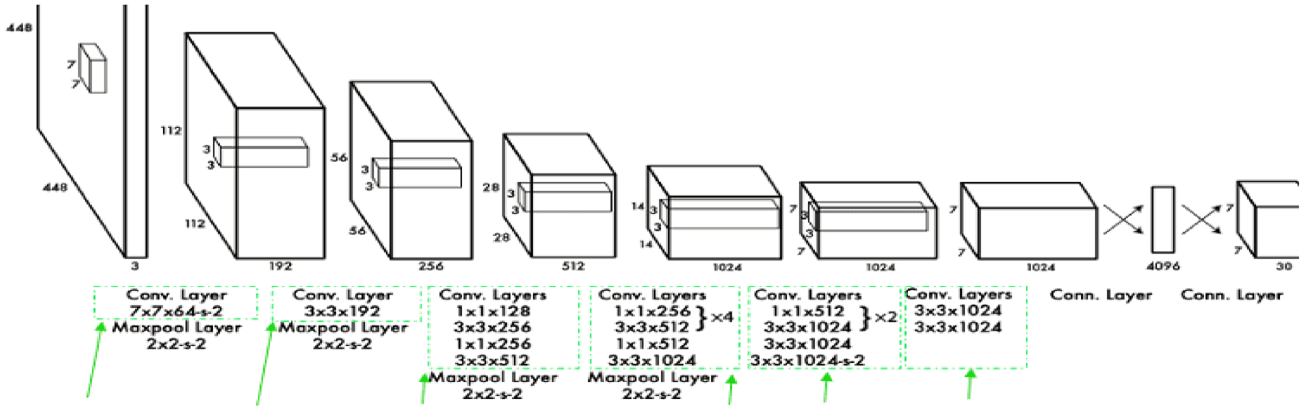
448
7
7
448
3
112
3
3
112
192
56
3
3
56
256
28
3
3
28
512
14
3
3
14
1024
7
3
3
7
1024
7
7
1024
4096
7
7
30
Conv. Layer 7x7x64-s-2 Maxpool Layer 2x2-s-2
Conv. Layer 3x3x192 Maxpool Layer 2x2-s-2
Conv. Layers 1x1x128 3x3x256 1x1x256 3x3x512 Maxpool Layer 2x2-s-2
Conv. Layers 1x1x256 3x3x512 }x4 1x1x512 3x3x1024 Maxpool Layer 2x2-s-2
Conv. Layers 1x1x512 3x3x1024 }x2 3x3x1024 3x3x1024-s-2
Conv. Layers 3x3x1024 3x3x1024
Conn. Layer
Conn. Layer

Fig. 3 YOLO Detector framework [12]

In the training process, each cell in YOLO predicts multiple bounding boxes but only one category related to the target. If there is indeed a target within a specific cell, only the bounding box with the highest Intersection over Union (IOU) score compared to the ground truth is considered responsible for predicting that target. The other bounding boxes in the cell are treated as non-targets. To address this issue, a solution is to specialize the bounding boxes within each cell and account for targets of various sizes and aspect ratios. This approach improves the overall performance of the model. It is important to note that the YOLO algorithm has a limitation in that it can train only one bounding box per cell. Therefore, if there are multiple targets within a single cell, the algorithm can only focus on predicting one of them. This limitation should be taken into consideration. Additionally, it's worth mentioning that when there is no corresponding target bounding box in a cell, the error term only accounts for the confidence score, and the coordinate error is not calculated. The error in classification can only be determined when a target is present within a cell; otherwise, it cannot be determined. Before discussing the detection method of the YOLO algorithm, it is important to mention the non-maximum suppression (NMS) algorithm.

NMS is not exclusive to YOLO but is commonly used in various detection methods to address the issue of multiple detections of the same target. In the context of face detection, for example, we often want to output only the best prediction box. To achieve this, the NMS algorithm is employed as follows: Identify the box with the highest confidence among the detected boxes. Calculate the Intersection over Union (IOU) between this box and all other boxes. If the IOU exceeds a predefined threshold indicating a high degree of overlap), remove the box with the lower confidence. Repeat this process for the remaining boxes until all detection frames are examined. The YOLO prediction process relies on the NMS algorithm to refine the final detection results. Now, let's discuss the basic strategy for obtaining the results of the bounding boxes. Initially, each predicted box is assigned a confidence score based on its category. After further processing, we obtain a set of prediction boxes with corresponding confidence values. The size of this set is determined by the formula mentioned earlier.

Typically, a confidence threshold is set, and boxes with confidence scores below this threshold are discarded. The remaining boxes with high confidence scores are then used as the final prediction boxes, forming the detection results. It is worth noting that there is a debate about whether NMS should be applied separately for each category or uniformly for all prediction boxes. While some sources suggest a separate application for each category, other implementations treat all boxes equally. The probability of objects from different categories appearing in the same location is generally low. The basic strategy described above is straightforward. However, in the case of the YOLO algorithm (according to the C source code), a slightly different approach is employed. The main difference lies in performing NMS first and then determining the categories of each box. Initially, the confidence values below the threshold are set to 0 for the 98 boxes. NMS is then applied separately to the confidence values. NMS does not eliminate the boxes but rather sets their confidence values to 0, and the category of each box is determined. Detection results are output for boxes with non-zero confidence values. This strategy is not as straightforward, but it is the approach adopted by the YOLO main code. The YOLO paper acknowledges that the NMS algorithm greatly influenced the evolution of YOLO, suggesting that this strategy works well for YOLO. However, in my own testing of ordinary picture detection, both strategies yield the same results.

D. OTP Working

Upon entering the number, the computation of the one-time password is initiated. The process involves several steps for data preparation and computation, as outlined below:

E. Data Preparation

1. The token captures the system time and generates a Date Time String in the format of YYYYMMDDHHMM, such as "200409281405" for September 28, 2004, at 2:05 PM. This Date Time String serves as the <Seed> following the RFC 2289 standard.
2. A Time Remaining value is calculated as the minimum of {20, 60 - seconds}, representing the remaining duration in which the one-time password will be displayed.
3. The Hash Count is computed as 24 - hour, representing the <sequence integer> according to the RFC 2289 standard. For instance, if the current time is 2:05 PM, the Hash Count would be 24 - 14 = 10.

F. One-Time Password Computation

1. Initial Step:
    a. The initial Digest is obtained by performing an MD5-hash on the concatenation of the following:
        i. <Passphrase> composed of
            1. Serial Number (excluding the null-terminating character)
            2. Secret
            3. Bias.
        ii. <Seed> consisting of
            1. Date Time String.
    b. The initial Digest is then folded into 8 bytes, where byte 0 is XOR'd with byte 8, byte 1 with byte 9, and so on, resulting in a 64-bit value.

2. Computation Step:
    a. In this step, the Hash Count number of MD5 hashes is performed on the folded result obtained from the previous operation.

b. The input to this step is the 64-bit output generated in the initial step.
c. The output of this step is a 128-bit digest representing the final MD5 hash.

3. Display Step:

a. The 128-bit digest output is folded into 8 bytes, where byte 0 is XOR'd with byte 8, byte 1 is XOR'd with byte 9, and so on.
b. For each of the 8 bytes, the corresponding one-time password value is calculated. This is done by multiplying the byte value by 39, dividing the result by 1000, and truncating the decimal point.
c. Let xi represent the value of the i-th byte, and ni represents the i-th numeric one-time password digit at the corresponding position.

The leftmost significant digits, based on the specified one-time password length, are shown as the displayed number. This displayed number represents the one-time password for the current minute.

The display duration is determined by the Time Remaining value, which indicates the number of seconds the number should remain visible. Once the total displayed time reaches the Time Remaining value, the token should turn off the display.

Additionally, if the user presses the button before the Time Remaining duration is completed, the token should also have the capability to switch off the display.

G. One-Time Password Computation

The verification of the one-time password can be performed using the DSSS Authentication Server. The DSSS Authentication Server incorporates a single drift window parameter, which can be configured to accommodate clock drift in the tokens. This parameter represents the allowable drift in minutes (both forward and backward) for the token since its last login. For instance, if the drift window parameter is set to 3, the server will permit a time drift of up to 3 minutes ahead and 3 minutes behind the token's last login time. Special procedures are necessary for initializing the token and the first login. When a token is registered on the server, an initialization string is generated specifically for that token. The initialization string can be transmitted to the token user through various means such as Pin Mailer, SMS, Bluetooth, infra-red, or email. During the first login, the server allows a wider time drift from the actual server time, which is calculated as 5 times the drift window parameter. For instance, if the drift window parameter value is 3, the server permits the token time to deviate up to 15 minutes before or after the server time. Additionally, the server computers securely stores a secret Bias value in the database. The Bias value can range from 0 to 255, ensuring that the probability of another token (with the same serial number) generating the same one-time passwords from the same initialization string is only 1 in 256 (0.4%).

IV. RESULTS AND DISCUSSION

In this paper, we've attempted to propose a solution to fraudulent transactions using an automated teller machine (ATM) that only allow withdrawing cash if the account holder is physically present by confirming there phone number and scanning there face in real time. As a result, it reverses instances of unauthorized transactions that are done without the consent of the true owner. Additional face recognition adds even more security.

The implementation of smart ATMs using facial recognition technology has significant benefits for the banking industry and its customers. By providing a more secure and convenient way to perform transactions, smart ATMs reduce the risk of fraudulent activities while enhancing the overall banking experience. As technology continues to evolve and improve, we can expect to see further developments in this area, such as integration with mobile devices, advanced security features, customized user experiences, and increased accessibility. In conclusion, the implementation of smart ATMs using facial recognition technology has significant benefits for the banking industry and its customers. By providing enhanced security, convenience, accessibility, and personalization, smart ATMs offer a glimpse into the future of banking and how technology can be used to improve our everyday lives. Exciting advancements in smart ATM technology are continuing to shape the future of banking.

The incorporation of artificial intelligence (AI) and machine learning algorithms into intelligent automated teller machines (ATMs) is a highly auspicious advancement that has the potential to promptly identify and avert fraudulent behaviors. These algorithms can detect possible fraudulent activity and send out alerts or extra security measures by analyzing transaction patterns, consumer behavior, and other variables. Moreover, smart ATMs can also help banks reduce operational costs by automating several tasks, such as cash withdrawals and deposits, balance inquiries, and account transfers. This automation can free bank staff to focus on more complex tasks, such as customer service and financial advice. Smart ATMs can also help banks gather valuable customer data and insights, such as transaction histories and spending patterns. This data can be used to develop customized and personalized products and services that meet the unique needs and preferences of individual customers. Another potential benefit of smart ATMs is their ability to offer a range of non-financial services, such as bill payment, ticket purchases, and even video conferencing with bank representatives. These additional services can provide customers with greater convenience and flexibility, while also generating additional revenue streams for banks. As smart ATM technology continues to evolve, we can expect to see further integration with other emerging technologies, such as block chain and the Internet of Things (IoT). These technologies can further enhance the security, efficiency, and convenience of smart ATMs, while also enabling new use cases and applications. They have the potential to transform the way we interact with banks, by providing secure, convenient, and personalized services that meet the needs and preferences of individual customers.

V. CONCLUSION

Smart ATMs incorporating facial recognition technology offer numerous advantages for both the banking industry and its customers. These innovative systems provide enhanced security and convenience, effectively mitigating the risk of fraudulent activities and elevating the overall banking experience. As technology advances, we can anticipate even more exciting progress in this field. This may include integration with mobile devices, the implementation of advanced security features, personalized user experiences, and improved accessibility. Smart ATMs utilizing facial recognition technology provide a promising glimpse into the future of banking, showcasing the transformative potential of technology to enhance our daily lives.

REFERENCES

[1] Jiang, C., Ma, H., Li, L. (2022). IRNet: An Improved RetinaNet Model for Face Detection. In 2022 *7th International Conference on Image, Vision and Computing (ICIVC),* 129-134.

[2] Qi, D., Tan, W., Yao, Q., Liu, J. (2023). YOLO5Face: Why reinventing a face detector. In *Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings,* Part V, pp. 228-244. Springer Nature Switzerland, Cham.

[3] Chen, W., Huang, H., Peng, S., Zhou, C., Zhang, C. (2021). YOLO-face: a real-time face detector. *The Visual Computer, 37*, 805-813.

[4] Shi, M. A., Gao, Y. (2021). Lightweight real-time face detection method based on improved YOLOv4. In *2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI),* 273-277.

[5] Zhou, Y., Wang, Z., Xie, G., Liao, J., Huang, W., & Xiao, C. (2022). A novel YOLO V5 framework for infrared image recognition. In 2022 *IEEE 5th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE),* 353-360.

[6] Sharma, S., Raja, L., Bhatnagar, V., Sharma, D., Bhagirath, S. N., &Poonia, R. C. (2022). Hybrid HOG-SVM encrypted face detection and recognition model. *Journal of Discrete Mathematical Sciences and Cryptography, 25*(1), 205-218.

[7] Deshmukh, A., *et al.*, (2019). Face Recognition Using OpenCv Based On IoT for Smart Door. *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM),* Amity University Rajasthan, Jaipur-India.

[8] Ismael, Khansaa Dheyaa, & Stanciu Irina. (2020). Face recognition using Viola-Jones depending on Python. *Indonesian Journal of Electrical Engineering and Computer Science, 20*(3), 1513-1521.

[9] Puthea, Khem, Rudy Hartanto, &Risanuri Hidayat. (2020). The Attendance Marking System, based on Eigenface Recognition using OpenCV and Python. *Journal of Physics: Conference Series, IOP Publishing, 1551*(1),

[10] Arya, Zankruti, & Vibha Tiwari. (2020). Automatic Face Recognition and Detection Using OpenCv' Haar Cascade and Recognizer for Frontal Face. ECE Department Medi-Caps University Indore (MP), India.

[11] Zhu, Zhiguo, and Yao Cheng. (2020). Application of attitude tracking algorithm for face recognition based on OpenCV in the intelligent door lock. *Computer Communications, 154,* 390-397.

[12] Kushal, M., *et al.,* (2020). ID Card Detection with Facial Recognition using Tensor flow and OpenCV. *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), IEEE.*

[13] Halidu, M. K., Bagheri-Zadeh, P., Sheikh-Akbari, A., & Behringer, R. (2019). PCA in the context of Face Recognition with the Image Enlargement Techniques, *2019 8th Mediterranean Conference on Embedded Computing (MECO),* Budva, Montenegro, 1-5, DOI: 10.1109/MECO.2019.8760162.

[14] Liu, Z., Jiang, D., Li, Y., Cao, Y., Wang, M., & Xu, Y. (2019). Automatic Face Recognition Based on Sparse Representation and Extended Transfer Learning, *in IEEE Access, 7,* 2387-2395, DOI: 10.1109/ACCESS.2018.2883288.

[15] Sushama, M. & Rajinikanth, E. (2018). Face Recognition Using DRLBP and SIFT Feature Extraction, *2018 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, 994-999, DOI: 10.1109/ICCSP.2018.8524427.

[16] Sinha, A. S., Rahman, A. U., Kumar R. K. & Sanyal, G. (2019). Enhancing Face Recognition through Overlaying Training Images. *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP),* Gangtok, India, 1-4, DOI: 10.1109/ICACCP.2019.8882933.

[17] Khan, M., Chakraborty, S., Astya, R. and Khepra, S. (2019). Face Detection and Recognition Using OpenCV. *2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS),* Greater Noida, India, 116-119, DOI: 10.1109/ICCCIS48478.2019.8974493.

[18] Saravanan, P., Pandey, A., Joshi, K., Rondon, R., Narasimharao, J., & Imran, A. A. (2023, May). Using machine learning principles, the classification method for face spoof detection in artificial neural networks. *In 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE),* 2784-2788. IEEE.

[19] Kumar, A., Pathak, N., Kirola, M., Sharma, N., Rajakumar, B., & Joshi, K. (2023, April). AI based mouse using Face Recognition and Hand Gesture Recognition. *In 2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1), IEEE,* 1-6.