Received: 4 September 2024 | Revised: 14 October 2024 | Accepted: 7 November 2024 | Published online: 8 January 2025

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2025, Vol. 00(00) 1–10
DOI: 10.47852/bonviewJCCE52024248

BON VIEW PUBLISHING

Strengthening Security in Clouds Through Cloud Computing Authentication Using Facial Image Forensics

Pranali Dahiwal[1,*] and Vijay Khare[2]

[1] *Vishwakarma Institute of Information Technology, Savitribai Phule Pune University, India*

[2] *Department of Defence and Strategic Studies, Savitribai Phule Pune University, India*

Abstract: Numerous cyber threats may succeed in cloud platforms owing to unproductive authentication methods. Multiple credential authentication is a vital precautionary measure that helps reinforce cloud security for warding off imminent data breaches and illegal access. This kind of authentication approach strongly guarantees that trustworthy clients are only ratified to get cloud services, making it less tiresome for clients and more secure for organizations. This study proposes an Image Forensics-based Dual Credential Authentication (IF-DCA) based on username (i.e., preferably email ID to avoid replica) and client photographs as password. Along with a username, a rapid and acquainted user action of capturing a photograph using a web camera is adequate for this method; hence it does not need any expensive, special hardware devices. During the registration phase, along with the username, the user's face image is stored on the cloud server as a password. When registered clients need to access cloud resources or data, they should log in on the server with their username and their photograph. To improve the authentication approach, this work proposes an Attribute-controlled Conditional Generative Adversarial Network (ACC-GAN) to generate face images of the same user at various age groups. ACC-GAN includes an additional attribute control unit and an age prediction unit to synthesize photo-realistic face images with aging effects. After matching the username and face image with the registered username and ACC-GAN-generated images, the face identification module provides access to the authorized cloud user. The performance of IF-DCA is assessed through the Cross-Age Celebrity Dataset using the MATLAB R2018b/deep learning toolbox. The empirical analysis reveals that the ACC-GAN achieves better performance measures such as 98.40% accuracy, 99.3% sensitivity, 90.0% specificity, and 95.7% precision. The performance of IF-DCA is analyzed using the verification time for username, user image, and complete authentication against the number of users.

Keywords: authentication mechanism, face recognition, generative adversarial network, multi-credential, password

1. Introduction

Recently, the reputation for cloud services has witnessed a dramatic acceleration, thanks to big cloud providers including Alibaba, Microsoft, Google, Oracle, and Amazon [1]. Cloud computing, which was instigated as a mere storage solution, has now become an all-encompassing computing paradigm in quite a short span. In a nutshell, cloud technology is the mainstay of enterprise infrastructure, essentially transforming the way organizations collect, process, store, and communicate data. As security engineers are aware, however, all the technology that grows into prevalent in the cyber world will inexorably become a potential target of malicious attackers, and cloud technology is the same [2]. Of late, the power of cyberattacks to gain illegal access to cloud services and applications is getting undue as the complexity of password hacking methods grows and processing complex calculations becomes very cost-effective.

Incidentally, cloud cyberattacks accounted for 20% of all internet threats in 2020, making this paradigm the third most beleaguered digital technology globally [3]. The Indian Computer Emergency Response Team (CERT-In) divulges that Indian media experienced almost 1.4 million cyber threats in 2022, and among these, threats on the cloud were the maximum. Phishing threats on the cloud servers have now increased by 65% in the six months ending October 2022 related to the preceding year, and 76% of organizations recorded sophisticated phishing attacks in the last year [4]. The cyber threats are also becoming complex and are transcending beyond emails to instant messages and different types of private communication. Currently, we are more vulnerable to account stealing threats that, statistically, as a minimum one of our digital accounts (e.g., social media accounts, email, banks, etc.) will be hijacked or encountered to hack in the next 12-month period [5]. Therefore, it is indispensable to develop more secure and robust access methods to secure systems and data.

Authentication is the process of validating cloud user identities to decide whether he/she is reliable to authorize resources, services, data, and applications. The deficiency of robust and efficient cloud authentication schemes leads to the incidence of some dual

***Corresponding author:** Pranali Dahiwal, Vishwakarma Institute of Information Technology, Savitribai Phule Pune University, India. Email: Pranali.dahiwal@viit.ac.in

© The Author(s) 2025. Published by BON VIEW PUBLISHING PTE. LTD. This is an open access article under the CC BY License (https://creativecommons.org/licenses/by/4.0/).

credential. Data exposure, data altering, account hacking, spoofing identity, repudiation, denial-of-service (DoS), and promotion of access rights are some of the most general attacks in cloud platforms [6]. Traditional cloud authentication mechanisms, including simple text-based passwords (also called one-factor authentication), have exposed susceptibilities to diverse cyberattacks. Indeed, 61% of all attacks encompass credentials, whether hacked through brute force or pinched through social engineering. Hence, a strong user authentication approach is important to develop a safe and sound environment. Integrating text-based passcodes with pictorial passwords in a multiple credential authentication (MCA) can be an efficient method. Some cutting-edge authentication mechanisms (e.g., biometrics) are efficient methods but need extra hardware for effective application.

MCA is intended to provide high-level security against threats by adding complexity for invaders to obtain access to cloud resources and data, even if secret words are cracked by applying attacks or other methods. MCA achieves to do this with a hierarchical method. In this system, a client is asked to submit an amalgamation of two or more factors to authenticate his/her identity, so access can be approved [7]. The number of log-in credentials differs according to the architecture of the security framework and the anticipated security level. MCA methods typically hinge on biometric systems, which are automatic identification of users, according to their activities [8], and biological features including the face, iris, palm print, fingerprint, palm/finger vein, and voice [9]. On the other hand, the application of biological features has its technical hitches, primarily attributed to convenience, which mostly affects the application of the MCA method. Additionally, biometric authentication involves more expensive implementation and is still susceptible to numerous cutting-edge cyber threats including DoS, replay attacks, sensor result capture, presentation attacks, etc.

While numerous research works on validation for cloud platforms have used the concept of MCA, the decisive goal of any security tool is to ensure safe communication by averting conciliation and threats on the current validation approaches. In this context, this research proposes an Image Forensics-based Dual Credential Authentication (IF-DCA) based on usernames and client photographs as passwords. The contributions of this research are four-fold:

We propose an image forensic-based dual-factor authentication using usernames and client photographs as passwords.

The proposed IF-DCA mechanism integrates a text-based authentication for user face images. During the registration phase, the user's face photo is stored on the cloud server as a passcode. Whenever registered clients need to access cloud resources, they should log in on the cloud server with their username and their current photograph. After comparing the username and face image with the registered username and user photographs, the face recognition module in the proposed system offers access to cloud resources to the authorized client.

To improve the effectiveness of the proposed authentication approach and decrease false alarms, this work proposes an Attribute-controlled Conditional Generative Adversarial Network (ACC-GAN) to generate face photos of the same person in diverse age groups. ACC-GAN includes an additional attribute control unit and an age prediction unit to synthesize photo-realistic face images with aging effects.

The performance of this model is evaluated through a standard database such as the Cross-Age Celebrity Dataset (CACD) using the MATLAB R2018b/deep learning toolbox software.

This article is arranged as follows: We analyze the related studies about MCA in Section 2. In Section 3, we take a comprehensive look at the GAN model for empowering face recognition. The proposed IF-DCA model using password and image forensics is discussed in Section 4. Then the implementation details and numerical results obtained from experiments are given in Section 5 and Section 6, respectively. We conclude this work in Section 7.

2. Literature Review

MCA has become critical in the authentication of cloud user identity to minimize the jeopardy of illegal access to cloud resources, services, data, and applications [10–12]. Patel et al. [13] suggested a systematic approach for validating users by applying passcode, out-of-band, and biometrics-based access control methods that are appropriate for access control. This approach encompasses a client name, passcode, biometrics features, and a smartphone to get a cyberattacks password (OTP). Kaleem and Arshad [14] proposed an adaptable user validation model using MCA for cloud platforms. The intended model provides a suitable and creative plan by combining the normal username and passcode-based authentication schemes. This model provides effective authentication, which can compete against different types of cyber threats.

Priya and Sumalatha [15] proposed a multilevel security model by applying cohesive MCA methods using security interrogations, biometrics, and OTP on top of passcode-based validation to secure the cloud services and data from intruders. The cohesive approaches provide robust and secure validation owing to the dual-level protection. The biometric validation comprises pictures as a verification parameter and the exacerbation of the attributes is achieved through a face recognition framework with the notion of transfer learning. Hussain et al. [16] proposed a fully secured authentication scheme to alleviate manifold verifications typically required from a specific client. The intended model enables a federated trust between providers and consumers.

Midha et al. [17] suggested a secure MCA protocol for medical applications in a cloud-based Software Defined Network. The authors used a body area network to assess the enactment of the intended approach and guarantee that no illegal user can snip sensitive patient data. The outcomes demonstrate that this approach guarantees safe access to the cloud server regarding identification and spoofing. Prabakaran and Ramachandran [18] developed an MCA scheme for secured financial transactions in a cloud environment. This work employs a cryptography algorithm to achieve secure business dealings by applying a robust MCA scheme using text-based and biometric authentication. The empirical outcomes demonstrate that the projected approach is to be a perfect method for real-time applications. Based on this review, MCA is particularly imperative in cloud platforms, in which data and services are frequently presented by third-party vendors. By applying appropriate MCA, we can decrease the jeopardy of illegal admission to cloud services, even if hackers can get a user's passcode. Providers can select to apply MCA for all clients or only clients who deal with private information. MCA can be realized at the application level or vendor level. In Table 1, MCA-based existing works are mentioned.

3. The Proposed One-Time Model

The proposed IF-DCA model embeds a face identification module (FIM). This module accepts usernames and their photographs as passwords for authentication. The proposed IF-DCA scheme contains four phases: registration, dual credential, authentication, and update. In the registration phase, users can choose their

Table 1
MCA-based existing works

Reference	Method	Feature extraction	Limitation
Patel et al. [13]	Password, biometrics, and out-of-band-based access control	Score matching-based feature extraction	Delay in receiving OTP and higher false alarm rate
Kaleem and Arshad [14]	Choice-based MCA	Score matching-based feature extraction	Leads to latency and overhead
Priya and Sumalatha [15]	MCA with a 2-layer security method	VGG face model	Leads to higher complexity
Hussain et al. [16]	Single sign-on with MCA	Shibboleth	Demands higher computational power
Midha et al. [17]	MCA with hash function	Score matching-based feature extraction	Higher latency
Prabakaran and Ramachandran [18]	MCA with Elliptical Curve Cryptography	Score matching-based feature extraction	Limited users can avail the service at a time

username, and they are allowed to capture his/her face using facial recognition cameras on their gadgets. In the log-in phase, the user must enter the username that was considered during the registration phase and capture their photo. The shortcoming of this approach is that when there is a long elapsed time between two log-in events, clients' biometric features vary (i.e., facial appearance) over time. In the authentication phase, IF-DCA matches the username, and the face recognition module verifies the photograph to provide access to cloud services for the user. In the updation phase, the currently captured photograph is stored for future authentications. To solve the problems related to varying facial features over time, this work employs an ACC-GAN to generate face images of a user at different ages. ACC-GAN includes an additional attribute control unit and an age prediction unit to synthesize photo-realistic face images with aging effects. It creates a database for each registered user.

3.1. Registration phase

The registration process is performed at one time except the client enrolls again. During this phase, the client provides their username (email ID) and captures their photo using facial recognition cameras. Users need to enroll their username to the FIM during this process. FIM verifies the newly entered username against the existing usernames stored in the database. The username should not replicate or match the prevailing client's email IDs. Once verifying the availability of the email ID, the face image must be captured through a camera and transferred to the cloud server as a passcode. The username and image are authorized and transferred to the cloud database. After receiving the username and image, the cloud server searches its records to find out whether the user is new or prevailing. If it is a new user, then the cloud server records the user details and calculates security measures that are exclusive to the user. Security measures are safely stored in a cloud database, making it more challenging for an attacker to access the cloud services.

3.2. Log-in phase

The log-in phase is employed when the client needs to access cloud resources or data. The user passes his/her username on a log-in page which was previously given by the user during the registration phase. For password authentication, the user's face image is captured by a camera. The IF-DCA authenticates the user identities. If the user fails to enroll the right identities, the IF-DCA should start an identity failure procedure to evade username-guessing cyberattacks. Conversely, if the client enrolls a valid username and face image as a password, the IF-DCA will produce a validation message and transfer it to the cloud server. Then, the server allows the user to access cloud data and services.

3.3. Authentication phase

The authentication phase initiates when the verification note is accepted by the server. Then, the server computes the time variation between the message arrival time and the time that the verification note was transferred by the user. This time variation is vital to circumvent replay cyberattacks. The server performs many verification processes to confirm the validity of the message.

3.4. Updating phase

After receiving an authentication message from the server, IF-DCA updates the user database by adding the newly captured photographs to the server database. The shortcoming of this approach is that when there is a long elapsed time between log-in events of a user, their biometric features change over time. In the authentication phase, IF-DCA matches the username, and the face recognition module verifies the photograph to provide access to cloud services for the user. In the image updation phase, the currently captured photograph is stored for future authentications. To solve the problems related to varying facial features over time, this work employs an ACC-GAN to generate face images of a user in different age groups. ACC-GAN includes an additional attribute control unit and an age prediction unit to synthesize photo-realistic face images with aging effects. It creates a database for each registered user.

4. GAN for Empowering Face Recognition

A generative adversarial network is an efficient promising log-in method with the potential to learn the pattern of specified images and produce analogous record samples. The notion of the GAN model was developed by Goodfellow et al. [19] in 2014. It comprises two autonomous deep networks, called generator or producer (G) and discriminator of differentiator (D). These modules are challenging each other to make each other strong. They can engender pictures with improved quality, identity reliability, and aging exactness related to conventional approaches. The producer provides photographs that bear a resemblance to real user images with a twist rendering those fake samples. A differentiator computes the likelihood of a given photograph that fits into the real database. It acts

as a critic and is improved to find out fake samples from the actual ones. A GAN-based image generation is then used to improve the enactment of the IF-DCA. The producer is employed for producing similar user photos, and the differentiator has a stronger capability to categorize the users' images based on their age. These two modules contest each other during the learning phase so that the producer is attempting to cheat the differentiator, while the differentiator is struggling not to be cheated. This exciting min–max game between two modules allows both to improve their performance.

Given an input face image $i = R^{w\times h\times N}$and several target age groups $t_i = R^N$, in which w and h signify the width and height of an attribute vector, N denotes the number of age groups, correspondingly. To produce a fake photo within the target age group, t_i can be shown in Equation (1):

$$i_t = G(i, t_i) \tag{1}$$

The producer contains an encoding and a decoding module. The encoding unit targets to transform the high-dimensional user image into a low-dimensional embedding space, thus imposing the neural network to extract the most significant attributes. The encoding unit is built with a full convolution neural network (FCNN) as given in Figure 2. To keep the significant data of the image, this work replaces the fully connected module of the FCNN with a convolutional module. In this structure, the number of input modules is optimized to be in line with the size of the image vector after including one-hot coding. This encoder enables the producer to get the appearance learning capacity in the latent vector so that the data is used at the semantic level using conditional implanting or interpolation on the latent space. The decoding unit targets reinstating the latent matrix of the hidden modules to the original size and making the result $i_l = i$. A small-step convolutional design is developed to create the decoding module. A differentiator unit is employed to find out the validity of the input fake image i_t, which is logically dependable on the differentiator in the basic GAN, that is, maxD(i_t, n), in which D is a differentiator to find the quality of the artificial face and n is the original image fitting into the target age group i_t. Figure 1 depicts the general architecture of GAN.

In the GAN learning process, the images are pigeonholed based on age group. After including a suitable noise component, the fake photos are transferred to G and converted into confrontational images for different age groups. The identified labels are employed by D to train the classifier. The losses of D and G are computed from the results of D and the classifiers. Assuming the noise input from the embedding space is n and the producer transfer function is G, the engendered result function becomes (n). Also, i is the actual user image, which denotes a group of one-dimensional vectors based on the designated features and D signifies the transfer function of a differentiator. As an independent component, D performs as a binary classifier that trains to measure the target function. The result of D is described by Equation (2).

$$Y(i,\ G(n)) = D(i) + D(G(n))\begin{cases} 0,\ Y = D(i) \\ 1,\ Y = D(G(n)) \end{cases} \tag{2}$$

The key objective of the producer is to maximize the function $D(G(n))$, while D aims to minimize $D(G(n))$ and maximize $D(i)$ concomitantly. The producer is rewarded with how superior it can produce photos that the differentiator finds hard to categorize. According to the intricacy level with which a differentiator discriminates synthetic (fake) images from real ones, they are either remunerated or reprimanded. The producer and differentiator modules update their corresponding constraints with an effective backpropagation model. It is a min–max game employed by the differentiator to classify real and fake photos flawlessly.

In this model, the mutual effort of these modules is considered. Hence, a differentiator targets to maximize the function (D). The first element of Equation (3) is related to the improved recognition of the real photo, while the second part aids in recognizing counterfeit photos well. To improve the producer constraints to cheat the differentiator cost function, (G) has to be minimized. Equations (3) and (4) display exact depictions of the same.

$$\mathrm{Max_D}\ V(D) = E_{ip_{data}(i)}[\log D(i)] + E_{n\approx p_n(n)}\ [\log(1 - D(G(n))] \tag{3}$$

$$Min_G\ V(G) = E_{n\approx p_n(n)}[\log(1 - D(G(n)))] \tag{4}$$

where $p_{data}(i)$ and $p_n(n)$ are the probability distribution and prior distribution of the real image, correspondingly. $D(G(n))$ is the probability of being decided as a real image after the produced images are recognized by a differentiator. The purpose of the producer is to produce synthetic photos to be decided as a real photo. Furthermore, $D(G(n))$ equals to 1, and $V(D,\ G)$ also decreases. The purpose of D is to judge i as a real photo. The synthetic photos will be decided as fake photos, making $D(i) \approx 1$ and $D(G(n)) \approx 0$. Hence, Equations (3) and (4) define the joint min–max game of the related cost function $V(D,\ G)$. The creation of the learning database in a GAN model is shown in Figure 3. The optimization process sustains with more than a few iterations until the Nash equilibrium is achieved as shown below.

$$\begin{aligned}\mathrm{Min_G Max_D}V(G,D) = {} & E_{ip_{data}(i)}[\log\ D(i)] \\ & + E_{n\approx p_n(n)}[\log(1 - D(G(n)))]\end{aligned} \tag{5}$$

As the producer does not know the dataset, it provides arbitrary guesses. Initially, the differentiator can easily discriminate the producer's feeble forecast from actual photos. As the producer's approximation starts to improve and resemble actual user photos from the database, the differentiator may misclassify the producer's synthetic photos. If this occurs, the differentiator experiences more losses, which allows it to train how to sense fake images from the producer, rather than the actual user photos. The producer retorts to the improved performance of the differentiator by generating improved images. Preferably, the training process halts when the differentiator achieves 50% efficiency.

The actual photo i and its corresponding age group label t_i are considered inputs. The best feature record i^* and the corresponding class t_i^* are outputs after feature selection. The latent vector n and

Figure 1
General architecture of GAN

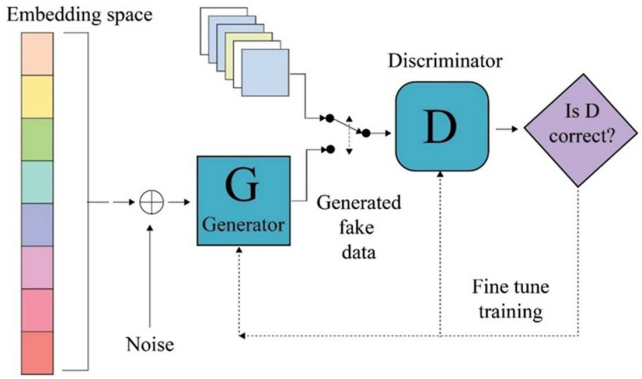


Figure 2
Generator and discriminator structure in GAN

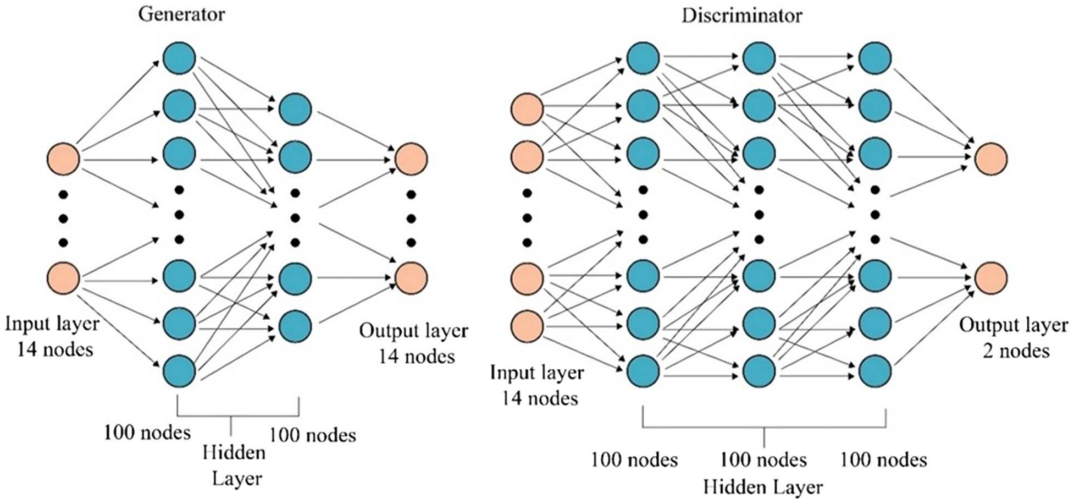
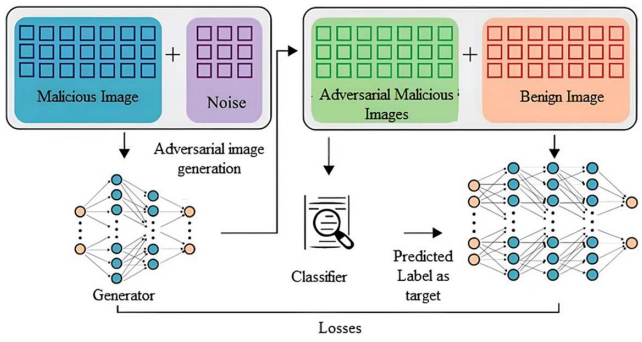


Figure 3
Generation of training dataset in GAN



condition vector t_{fake} are taken as input to the producer to generate the synthetic images i_{fake}. It is categorized by the differentiator. The differentiation loss between actual and generated images $Loss_{if}\ (G)$ and the classification loss $Loss_{class}\ (G)$ are computed. Then, the learning process of the producer is collaboratively achieved, and i* is accepted by the differentiator. The synthetic and actual data decision loss $Loss^{*}_{if}\ (G)$ and the classification loss $Loss^{*}_{class}(G)$ for i^{*} are outputs. In addition, i_{fake} is transferred to the differentiator for further processing. The actual and synthetic data decision loss $Loss^{fake}_{if}\ (D)$ and classification loss $Loss^{fake}_{class}\ (D)$ of i_{fake} are outputs. The learning process of the producer is realized by the loss functions given in Equations (6) and (7).

$$Loss_{if}\ (G) = E_{n\approx p_n(n),t_{fake}\approx P_t}\left[\log\left(1 - D\left(D\left(n, t_{fake}\right)\right)\right)\right] \quad (6)$$

$$Loss_{class}\ (G) = E_{n\approx p_n(n),t_{fake}\approx P_t}\left[L_D\left(t_{fake}||PG\left(n, t_{fake}\right)\right)\right] \quad (7)$$

When the input is random noise n and label t_{fake}, the producer generates images i_{fake} as the output. $Loss_G$ of the producer is computed from the weighted sum of loss functions. It includes the discriminant loss $Loss_{if}\ (G)$ of r_{fake} that is decided by the differentiator as real or as fake data. The classification loss $Loss_{class}\ (G)$ of i_{fake} that is characterized and computed using Equation (8).

$$Loss_G = \Psi Loss_{if}\ (G) + \phi Loss_{class}\ (G) + \eta\ Loss_{sh}\ (G) \quad (8)$$

where $Loss_{sh}\ (G)$ is the confrontation loss that i_{fake} is measured as the input images. The terms Ψ, ϕ, and η are the weights of the loss functions, $Loss_{if}\ (G)$, $Loss_{class}\ (G)$, and $Loss_{sh}\ (G)$, respectively. GAN is more stable in the learning process related to diffusion models, which often demand effective optimization of hyperparameters and noise schedules. GANs typically have faster inference times since they produce images in a single pass through the generator, while diffusion models usually need many iterations to generate a single sample. In some scenarios, GANs can generate high-quality, high-resolution images that are visually attractive and realistic, making them suitable for tasks where sample quality is vital.

4.1. Conditional GAN (deep learning)

To generate images with meticulous semantics, the labels of the fake images must be controlled. Although GAN can produce new arbitrary trustworthy images for a specific dataset, it is difficult to regulate the class of images that are created apart from attempting to recognize the multipart relationship between the generated images and the latent vector. The C-GAN facilitates a controlled cohort of images. Image cohort can be controlled by a class label to control the targeted synthetic data of a specific form. GAN can be enhanced by integrating with a conditional model. In this integrated model, the producer and differentiator are trained on some extra statistics ϖ. These statistics are supplementary data, such as class tags or data from other processes. In this work, the conditioning process is achieved by adding ϖ into both the producer and differentiator as an additional input. In the producer, the previous input noise $p_n\ (n)$ and ϖ are pooled in a hidden depiction, and the adversarial learning framework permits considerable suppleness in how this anonymous depiction is collected. In the differentiator, i and ϖ are supplied as inputs to a differentiator. The target function of a dual-factor min–max game for this C-GAN is described in Equation (9).

$$\begin{aligned} Min_G \mathrm{Max}_D\, V(G,D) = {} & E_{rp_{data}(r)}[\log\ D(r|\varpi)] \\ & + E_{n\approx p_n(n)}[\log(1 - D(G(n|\varpi)))] \end{aligned} \quad (9)$$

4.1.1. Attribute-controlled C-GAN (ACC-GAN)

Attribute-controlled C-GAN includes an attribute control unit and an age classifier to the C-GAN architecture, allowing the real

user image to produce new photos that maintain the semantic data and pose better pictorial impacts. The objective of ACC-GAN is to produce an image of advanced years that follows the target age group from the actual photo. In the face generation process, the attribute control unit excerpts the related attribute vectors of the encoding and decoding modules, correspondingly, and makes them follow to loss limits. Simultaneously, L2 loss is also employed to limit the semantic attributes excerpted by the pre-trained model [20]. The differentiator is employed to classify whether the synthetic image is correct or not.

Retaining the semantic data of the input image is a vital constraint in the face aging procedure. Conversely, only applying the confrontational loss to generate the synthetic image and target image may not sufficiently conserve the semantic data. To resolve this problem, we present the semantic data protection process to control the image cohort procedure using the attributes excerpted from the model. The attribute control unit acquires the attribute vectors of the encoding/decoding modules individually and relates the equivalent attribute vectors. This requires the encoding/decoding modules to be balanced when scheming the framework. The benefit of this is that the excerpted related attribute vectors are of equal dimension, and they can be supplied into the pre-trained model, guaranteeing that the ultimate result is a linear matrix for parameter mapping using L2 normal form. Therefore, to make the equivalent attribute vectors of equal dimensions, L_{layer} is computed using Equation (10)

$$L_{layer} = \sum_{i \in l} ||f_{en}^{k} - f_{de}^{k}||^2 \tag{10}$$

In the above equation, the term l is the number of layers, and f_{en}^{i} and f_{de}^{i} denote the attribute vectors of the encoding unit and the decoding unit at the kth layer, correspondingly. To achieve semantic reliability, the perceptual loss is presented to reduce the difference in semantics of the input and the resultant image of the producer. This loss can be calculated using Equation (11).

$$L_{id} = \sum_{i \in p_i(i)} ||h_{id}^{i^*} - h_{id}^{i}||^2 \tag{11}$$

where $h_{id}^{i^*}(.)$ denotes attributes excerpted from a particular attribute module during the training process with image i as input. The difference between the paired attribute vectors can preserve the semantic data between the input and fake images.

4.1.2. Adversarial training

Adversarial training is a technique for generating synthetic images and storing them in a learning database to guarantee that the deep networks learn the latent confrontational features. Adversarial training has not yet been analyzed entirely in deep network-based image generation. This new method of learning will increase the generality and dependability of the image generation by learning the semantic features of the user photographs. A scarcity associated with the adversarial training is that it only offers durability against the adversarial images it was trained on, and the classifier will still be circumvented by new adversarial perturbations. To address this problem, we develop an ACC-GAN-based method for confrontational training of the IF-DCA model for excluding unrelated images. The intended IF-DCA integrates the ACC-GAN model. The IF-DCA is not only trained on the real user images but also the generated images. By this process, we can improve the efficiency of the IF-DCA in face recognition. The pseudo-code of this training is given in Algorithm 1.

Algorithm 1: AC-CGAN	
Input: Training image database and attributes The noise *n* for the synthetic image generation	
Output: The optimized producer and the differentiator	
1:	Initialize the producer, the differentiator;
2:	for the number of training iterations do
3:	for G-steps do
4:	Generator engenders the adversarial data based on loss functions;
5:	Apprise the constraints of the generator based on the loss function
6:	end for
7:	for D-steps do
8:	D classifies the training database as real and fake images;
9:	D classifies the training database;
10:	Apprise the constraints of D using loss functions
11:	end for
12:	end for

5. Implementation

The intended IF-DCA with ACC-GAN model is realized on an Intel Core i7-4790 processor with 16GB RAM, 3.6 GHz, and Windows 10 operating system through the MATLAB R2018b/deep learning toolbox. The hyper-parameter settings for the experimentation carried out in this work are shown in Table 2.

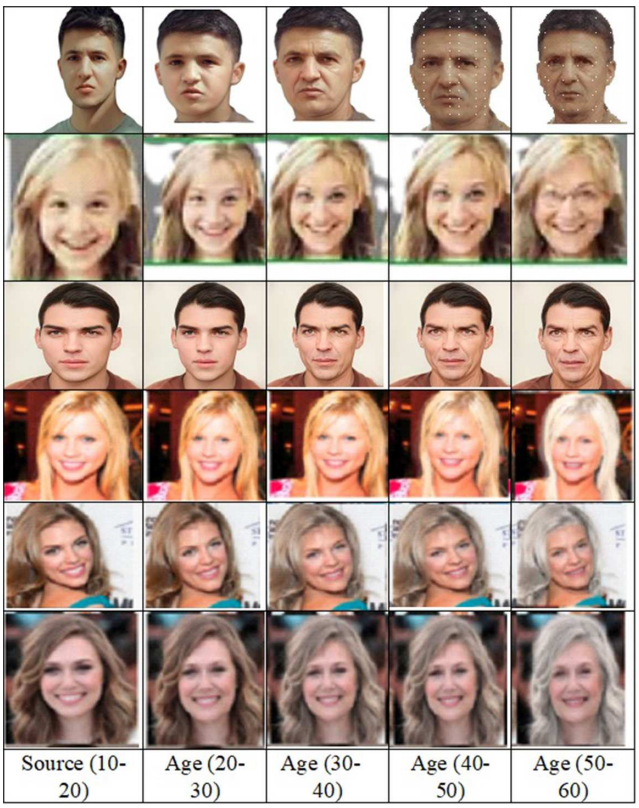
Table 2
ACC-GAN model parameter settings

Parameter	Settings
Batch size	128
Training rate	0.001
Iterations	100
Latent dimension	13
Dropout rate	0.5
Noise dimension	32
Optimizer	Adam
Weight initialization	Xavier initializer
Loss function	Categorical cross-entropy

5.1. Dataset preparation

To evaluate the effectiveness of the proposed IF-DCA with the ACC-GAN face recognition model, this study employs the CACD database [21]. This database encompasses around 160,000 face images with changes in expression, illumination, and pose captured from 2000 personalities aged from 16 to 62. Each photograph is marked based on age, however, not very precisely. This work first applies target detection to standardize the face alignment and then executes different image improvement methods, such as fine-tuning, illumination, flipping, and angle rotation.

Figure 4
The generated aged faces by ACC-GAN



After applying preprocessing methods, we select around 146794 images with a size of 400×400 pixels for analysis. This entire dataset is divided into two fragments. The face images are divided into five age groups: 10–20 (8656 images), 20–30 (36662 images), 30–40 (38736 images), 40–50 (35768 images), and 50-60 (26972 images) years old. To realize more precise outcomes, the 10-fold cross-validation (10 f-CV) method is used where the whole database is split into 10 parts. For each fold, one part is employed for testing, and the other slices are employed for training the classification algorithm. Now, the average value of all trials is considered for assessment. Figure 4 shows the general architecture of GAN.

5.2. Performance indicators

To evaluate the performance of ACC-GAN models, this study employs accuracy, sensitivity, specificity, precision, and ρ-values as performance indicators. These measures (except ρ-values) are essential to be higher to improve the efficiency of the ACC-GAN. The efficiency of the proposed model is computed in terms of ACC. The accuracy of the ACC-GAN is computed by Equation (12). In this equation, true positive (T^+) denotes the number of persons who are correctly classified as authorized persons; false negative (F^-) is the number of illegal users who are wrongly classified as authorized persons; true negative (T^-) is the number of users who are correctly identified as unauthorized users; and false positive (F^+) denotes the number of images who are wrongly classified as an unauthorized person. Sensitivity and specificity represent the ability of the ACC-GAN to differentiate between authorized and illegal users. Precision is the ratio of T^+ of a specific label to the total number of unauthorized users classified as the relevant class. The ratio of F^+ and F^- are also important measures to assess the enactment of the intended model. These indicators are defined by Equations (12)–(15).

$$\text{Accuracy} = \frac{(T^- + T^+)}{(T^- + T^+ + F^- + F^+)} \tag{12}$$

$$\text{Sensitivity} = \frac{T^+}{T^+ + F^-} \tag{13}$$

$$\text{Specificity} = \frac{T^-}{T^- + F^+} \tag{14}$$

$$\text{Precision} = \frac{T^+}{T^+ + F^+} \tag{15}$$

Wilcoxon's rank sum test is conducted to determine whether the ACC-GAN model provides a significant enhancement compared to other existing approaches or not. This nonparametric test is carried out by analyzing the effects of the intended ACC-GAN and relating it with other image generation models at a 5% significance level. The p-values less than 5% signify that there is a notable difference at a level of 5%. The p-values greater than 5% signify that there is no noteworthy difference between the related values. From the results, it can be concluded that in most of the trials, the p-values are < 5%, which proves that the enhancement obtained by our ACC-GAN models is statistically significant.

6. Results and Discussion

This study evaluates to which degree the IF-DCA model is capable of authenticating user images. This work assesses the efficiency of the proposed model by evaluating (i) the image generation capability of ACC-GAN and relating its results with some recently proposed classifiers, such as original GAN [19], Deep face [22], FaceNet [23], Pixie [24], Deep Convolutional Generative Adversarial Net (DCGAN) [25], and GPT-4o [26]; and (ii) the execution time of IF-DCA for username verification, user image verification, and total authentication [27].

6.1. Performance of ACC-GAN

The experimental results gained by the proposed model on the CACD database for different folds are given in Table 3. From this table, it is found that the ACC-GAN model has achieved better mean value of performance measures such as 98.40% accuracy, 99.3% sensitivity, 90.0% specificity, and 95.7% precision. The proposed ACC-GAN model also achieves less than 5% of ρ-value (1.6%), which designates that the results gained by ACC-GAN are significant. Furthermore, it realizes a reduced standard deviation for performance measures such as 0.7% accuracy, 0.3% sensitivity, 2.6% specificity, 2.8% precision, and 0.8% p-value.

Figure 5 shows the complete results obtained by the ACC-GAN model for different folds with respect to the mean value of the performance measure.

Table 4 and Figure 6 show the mean value of performance measures obtained by all the models including ACC-GAN. From these results, it is observed that the basic GAN model has achieved 87.60% accuracy, 96.2% sensitivity, 74.20% specificity, 87.90% precision, and 4.5% ρ-value. By applying a 9-layer deep network, Deep face achieves better performance than basic GAN. This model includes around 120 million factors through some locally linked modules without exchanging parameters, instead of the normal

Table 3
Results of ACC-GAN for different folding

Fold	Accuracy	Sensitivity	Specificity	Precision	*p*-value
#1	0.981	0.991	0.865	0.951	0.016
#2	0.982	0.991	0.883	0.932	0.019
#3	0.968	0.990	0.912	0.893	0.028
#4	0.991	0.992	0.940	0.978	0.029
#5	0.981	0.991	0.914	0.974	0.024
#6	0.982	0.998	0.935	0.975	0.022
#7	0.988	0.991	0.885	0.985	0.020
#8	0.992	0.994	0.912	0.961	0.003
#9	0.990	0.997	0.890	0.945	0.019
#10	0.985	0.991	0.867	0.976	0.010
Mean	0.984	0.993	0.900	0.957	0.019
S.D	0.007	0.003	0.026	0.028	0.008

Figure 5
Evaluation metrics obtained by ACC-GAN

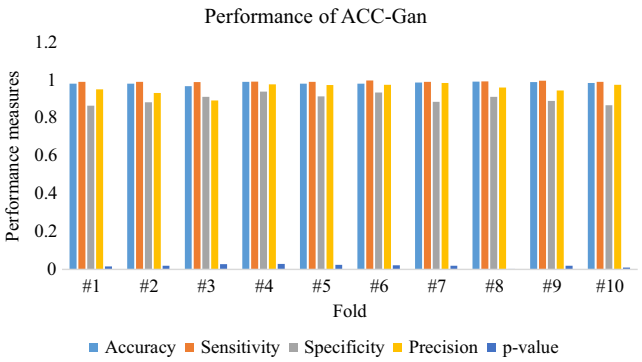


Table 4
Mean value of performance measures in ACC-GAN

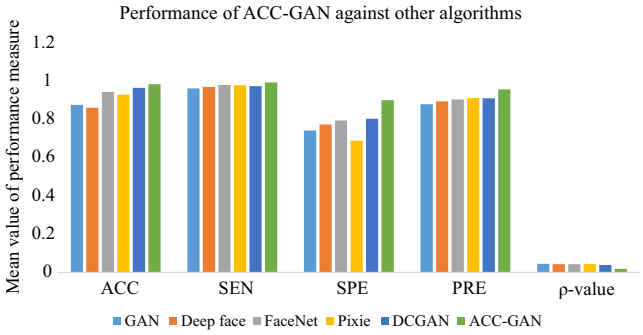
Algorithm	ACC	SEN	SPE	PRE	*ρ*-value
GAN	0.876	0.962	0.742	0.879	0.045
Deep face	0.861	0.970	0.773	0.894	0.043
FaceNet	0.944	0.980	0.794	0.904	0.043
Pixie	0.930	0.979	0.689	0.912	0.044
DCGAN	0.965	0.974	0.804	0.910	0.039
GPT-4o	0.970	0.960	0.885	0.920	0.035
ACC-GAN	0.984	0.993	0.900	0.957	0.019

convolution modules. The Deep face model has achieved 86.1% accuracy, 97.00% sensitivity, 77.3% specificity, 89.40% precision, and 4.3% *ρ*-value.

FaceNet employs a deep convolutional network that unswervingly learns a pattern from images and stores it in a dense Euclidean space in which distances reflect the amount of face resemblance. FaceNet provides 94.4% accuracy, 98.0% sensitivity, 79.4% specificity, 90.4% precision, and 4.3% *ρ*-value. The face recognition performance of Pixie, a camera-based dual-factor verification solution for gadgets, is better than other models. It achieves 93.0% accuracy, 97.9% sensitivity, 68.9% specificity, 91.2% precision, and 4.4% *ρ*-value.

By integrating data augmentation methods with the FaceNet model, DCGAN provides better performance. DCGAN increases the generalization ability and dimension of the learning dataset.

Figure 6
Mean value of measures in ACC-GAN and other models



DCGAN provides 96.5% accuracy, 97.4% sensitivity, 80.4% specificity, 91.0% precision, and 3.9% *ρ*-value. GPT-4o demonstrates high performance across multiple datasets. It excels in tasks that require few-shot learning and also provides notable improvements in multimodal tasks compared to its predecessors. It achieves 97% accuracy, 96% sensitivity, 88.5% specificity, 92.0% precision, and 3.5% *ρ*-value. However, the model shows variability and faces limitations in handling complex and ambiguous inputs, particularly in audio and vision capabilities.

From these results, it is found that the ACC-GAN model has achieved better mean values of performance measures such as 98.40% ACC, 99.3% SEN, 90.0% SPE, and 95.7% PRE. The proposed ACC-GAN model also achieves less than 5% of *ρ*-value (1.6%), which designates that the results gained by ACC-GAN are significant. The proposed model exhibits better results with 1.4% accuracy, 3.3% sensitivity, 1.6% specificity, 3.8% precision, and 8.4% *ρ*-value than GPT-4o. The SD value of the performance measure gained from the CACD database by each classifier is listed in Table 5. While considering the SD value of the evaluation metrics, the proposed model gains minimum SD values with 0.7% accuracy, 0.3% sensitivity, 2.6% specificity, 2.8% precision, and 0.8% *ρ*-value. From Figure 7, it can be observed that the SD of the ACC-GAN is smaller than all other intrusion detection models regarding the evaluation metrics. Hence, the ACC-GAN delivers much more reliable outcomes for detecting cyberattacks than the others. Therefore, the ACC-GAN model is considered a very viable model for detecting cyberattacks.

Table 5
SD value of measures in ACC-GAN

Algorithm	ACC	SEN	SPE	PRE	*ρ*-value
GAN	0.031	0.007	0.027	0.029	0.014
Deep face	0.017	0.006	0.029	0.038	0.017
Face Net	0.008	0.012	0.034	0.028	0.016
Pixie	0.019	0.003	0.023	0.028	0.018
DCGAN	0.011	0.005	0.021	0.025	0.020
GPT-4o	0.008	0.005	0.025	0.024	0.021
ACC-GAN	0.007	0.003	0.026	0.028	0.008

6.2. Performance of IF-DCA

The performance of IF-DCA is analyzed using the verification time for username, user image, and complete authentication against the number of users. As shown in Figure 8, the authentication time

Figure 7
Results of ACC-GAN in terms of SD values

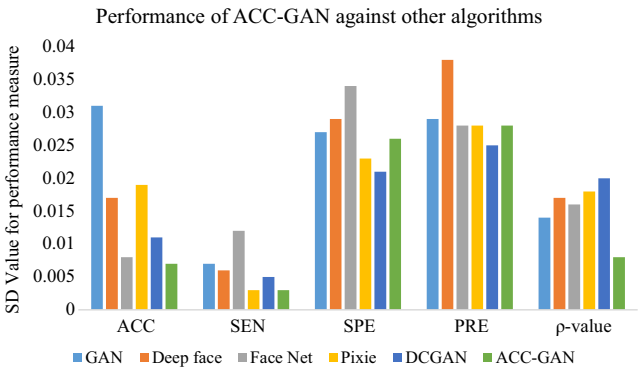
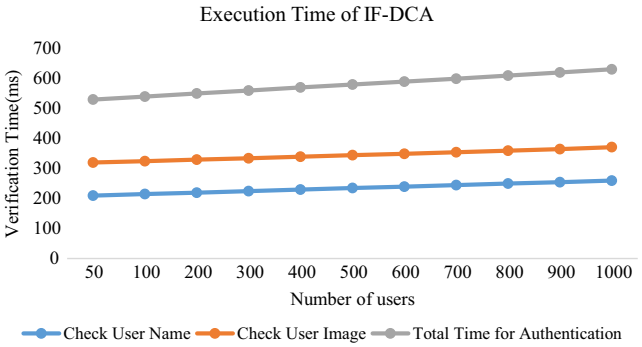


Figure 8
Verification time for user name, user image



for the first factor (username verification) increases linearly with the increasing number of users. The verification time for checking a valid username was 210.25 ms for 50 users, while it was 260.52 ms for 1000 users. For checking the user's facial image, the proposed algorithm takes 320.96 ms for 50 users and 372.11 ms for 1000 users. The overall authentication time also increases linearly with the increasing number of users. For 50 users, the total verification time is 531.21 ms, and 632.63 ms for an environment with 1000 users.

7. Conclusion

Traditional authentication approaches, such as simple text-based passwords, have exhibited susceptibilities to various kinds of cyberattacks on cloud platforms. MCA is a vital security mechanism to reinforce security for warding off illicit data access in C-GAN. This mechanism guarantees that only authorized users can access the cloud services, data, and applications. This paper proposes an image forensic-based authentication mechanism based on username and client photograph as password. During the registration phase, the user registers their username and their face image as a password. When registered clients need to access cloud resources, they should log in on the cloud server with their username and their photograph. To improve the authentication approach and decrease false alarms, this work proposes ACC-GAN to generate face images of the same person at different ages. ACC-GAN includes an additional attribute control unit and an age prediction unit to synthesize photo-realistic face images with aging effects. After matching the username and face image as a password with the registered username and ACC-GAN-generated images, the face recognition module provides access to cloud services to the authorized user. The performance of this model is evaluated through the CACD database. The extensive experimental analysis divulges that the ACC-GAN achieves better results than other existing models. The effectiveness of the IF-DCA is measured in terms of verification time for username, user image, and complete authentication against the number of users.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The Cross-Age Celebrity Datasets that support the findings of this study are openly available at https://paperswithcode.com/dataset/cacd.

Author Contribution Statement

Pranali Dahiwal: Conceptualization, Software, Formal analysis, Resources, Writing – original draft, Visualization, Project administration. **Vijay Khare:** Methodology, Validation, Investigation, Data curation, Writing – review & editing, Supervision.

References

[1] Rajendran, P., Maloo, S., Mitra, R., Chanchal, A., & Aburukba, R. (2023). Comparison of cloud-computing providers for deployment of object-detection deep learning models. *Applied Sciences, 13*(23), 12577. https://doi.org/10.3390/app132312577

[2] Habibzadeh, H., Nussbaum, B. H., Anjomshoa, F., Kantarci, B., & Soyata, T. (2019). A survey on cybersecurity, data privacy, and policy issues in cyber-physical system deployments in smart cities. *Sustainable Cities and Society, 50,* 101660. https://doi.org/10.1016/j.scs.2019.101660

[3] Al lelah, T., Theodorakopoulos, G., Reinecke, P., Javed, A., & Anthi, E. (2023). Abuse of cloud-based and public legitimate services as command-and-control (C&C) infrastructure: A systematic literature review. *Journal of Cybersecurity and Privacy, 3*(3), 558–590. https://doi.org/10.3390/jcp3030027

[4] Dawood, M., Tu, S., Xiao, C., Alasmary, H., Waqas, M., & Rehman, S. U. (2023). Cyberattacks and security of cloud computing: A complete guideline. *Symmetry, 15*(11), 1981. https://doi.org/10.3390/sym15111981

[5] Riggs, H., Tufail, S., Parvez, I., Tariq, M., Khan, M. A., Amir, A., &, Sarwat, A. I. (2023). Impact, vulnerabilities, and mitigation strategies for cyber-secure critical infrastructure. *Sensors, 23*(8), 4060. https://doi.org/10.3390/s23084060

[6] Tabrizchi, H., & Kuchaki Rafsanjani, M. (2020). A survey on security challenges in cloud computing: Issues, threats, and solutions. *The Journal of Supercomputing, 76*(12), 9493–9532. https://doi.org/10.1007/s11227-020-03213-1

[7] Otta, S. P., Panda, S., Gupta, M., & Hota, C. (2023). A systematic survey of multi-factor authentication for cloud

infrastructure. *Future Internet, 15*(4), 146. https://doi.org/10.3390/fi15040146

[8] Mekruksavanich, S., & Jitpattanakul, A. (2021). Biometric user identification based on human activity recognition using wearable sensors: An experiment using deep learning models. *Electronics, 10*(3), 308. https://doi.org/10.3390/electronics10030308

[9] Jain, A. K., Nandakumar, K., & Ross, A. (2016). 50 years of biometric research: Accomplishments, challenges, and opportunities. *Pattern Recognition Letters, 79,* 80–105. https://doi.org/10.1016/j.patrec.2015.12.013

[10] Sinigaglia, F., Carbone, R., Costa, G., & Zannone, N. (2020). A survey on multi-factor authentication for online banking in the wild. *Computers & Security, 95,* 101745. https://doi.org/10.1016/j.cose.2020.101745

[11] Patil, D. H., Asbe, V. S., Chavan, M. S., Birajdar, P. L., & Joshi, G. A. (2019). A survey on private cloud storage security using multifactor authentication. *Journal of Xi'an University of Architecture & Technology, 11*(8), 7–11.

[12] Yang, Y., Zhang, B., Guo, D., Du, H., Xiong, Z., Niyato, D., & Han, Z. (2024). Generative AI for secure and privacy-preserving mobile crowdsensing. *IEEE Wireless Communications,* 1–10. https://doi.org/10.1109/MWC.004.2400017

[13] Patel, S. C., Jaiswal, S., Singh, R. S., & Chauhan, J. (2018). Access control framework using multi-factor authentication in cloud computing. *International Journal of Green Computing, 9*(2), 1–15. https://doi.org/10.4018/IJGC.2018070101

[14] Kaleem, M., & Arshad, M. J. (2017). A customizable client authentication framework (CCAF) based on multi-factor for cloud computing application. *International Journal of Computer Science and Telecommunications, 8*(3), 18–25.

[15] Priya, K. D., & Sumalatha, L. (2020). Trusted hybrid multi-factor authentication for cloud users. *i-manager's Journal on Cloud Computing, 7*(1), 12–20. https://doi.org/10.26634/jcc.7.1.16670

[16] Hussain, M. I., He, J., Zhu, N., Sabah, F., Zardari, Z. A., Hussain, S., & Razque, F. (2021). AAAA: SSO and MFA implementation in multi-cloud to mitigate rising threats and concerns related to user metadata. *Applied Sciences, 11*(7), 3012. https://doi.org/10.3390/app11073012

[17] Midha, S., Verma, S., Kavita, Mittal, M., Jhanjhi, N. Z., Masud, M., & AlZain, M. A. (2023). A secure multi-factor authentication protocol for healthcare services using cloud-based SDN. *Computers, Materials & Continua, 74*(2), 3711–3726. https://doi.org/10.32604/cmc.2023.027992

[18] Prabakaran, D., & Ramachandran, S. (2022). Multi-factor authentication for secured financial transactions in cloud environment. *Computers, Materials & Continua, 70*(1), 1781–1798. https://doi.org/10.32604/cmc.2022.019591

[19] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., . . . , & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems.*

[20] Gao, F., Ma, F., Wang, J., Sun, J., Yang, E., & Zhou, H. (2018). Semi-supervised generative adversarial nets with multiple generators for SAR image recognition. *Sensors, 18*(8), 2706. https://doi.org/10.3390/s18082706

[21] Chen, B. C., Chen, C. S., & Hsu, W. H. (2014). Cross-age reference coding for age-invariant face recognition and retrieval. In *Computer Vision – ECCV 2014: 13th European Conference,* 768–783. https://doi.org/10.1007/978-3-319-10599-4_49

[22] Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. In *IEEE Conference on Computer Vision and Pattern Recognition,* 1701–1708. https://doi.org/10.1109/CVPR.2014.220

[23] Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. In *IEEE Conference on Computer Vision and Pattern Recognition,* 815–823. https://doi.org/10.1109/CVPR.2015.7298682

[24] Azimpourkivi, M., Topkara, U., & Carbunar, B. (2017). Camera based two factor authentication through mobile and wearable devices. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1*(3), 35. https://doi.org/10.1145/3131904

[25] Ammar, S., Bouwmans, T., & Neji, M. (2022). Face identification using data augmentation based on the combination of DCGANs and basic manipulations. *Information, 13*(8), 370. https://doi.org/10.3390/info13080370

[26] Shahriar, S., Lund, B. D., Mannuru, N. R., Arshad, M. A., Hayawi, K., Bevara, R. V. K., & Batool, L. (2024). Putting GPT-4o to the sword: A comprehensive evaluation of language, vision, speech, and multimodal proficiency. *Applied Sciences, 14*(17), 7782. https://doi.org/10.3390/app14177782

[27] Gomathi, N., & Wagh, M. B. (2022). Improved rider for vehicular adhoc NETwork routing via neural network. *Evolutionary Intelligence, 15,* 1517–1530. https://doi.org/10.1007/s12065-021-00602-0

How to Cite: Dahiwal, P., & Khare, V. (2025). Strengthening Security in Clouds Through Cloud Computing Authentication Using Facial Image Forensics. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE52024248