Computer Modeling in
Engineering & Sciences

Tech Science Press

DOI: 10.32604/cmes.2023.030640

ARTICLE

Robust Facial Biometric Authentication System Using Pupillary Light Reflex for Liveness Detection of Facial Images

Puja S. Prasad[1], Adepu Sree Lakshmi[1], Sandeep Kautish[2], Simar Preet Singh[3], Rajesh Kumar Shrivastava[3], Abdulaziz S. Almazyad[4], Hossam M. Zawbaa[5] and Ali Wagdy Mohamed[6,7,*]

[1]CSE Department, Geethanjali College of Engineering, Hyderabad, Telangana, 501301, India

[2]Department of Computer Science, LBEF Campus, Kathmandu, 44600, Nepal

[3]School of Computer Science Engineering and Technology (SCSET), Bennett University, Greater Noida, 201310, India

[4]Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, P.O. Box 51178, Riyadh, 11543, Saudi Arabia

[5]CeADAR Ireland's Centre for AI, Technological University Dublin, Dublin, D07 EWV4, Ireland

[6]Operations Research Department, Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, 12613, Egypt

[7]Applied Science Research Center, Applied Science Private University, Amman, 11937, Jordan

*Corresponding Author: Ali Wagdy Mohamed. Email: aliwagdy@gmail.com

Received: 16 April 2023 Accepted: 18 July 2023 Published: 30 December 2023

ABSTRACT

Pupil dynamics are the important characteristics of face spoofing detection. The face recognition system is one of the most used biometrics for authenticating individual identity. The main threats to the facial recognition system are different types of presentation attacks like print attacks, 3D mask attacks, replay attacks, etc. The proposed model uses pupil characteristics for liveness detection during the authentication process. The pupillary light reflex is an involuntary reaction controlling the pupil's diameter at different light intensities. The proposed framework consists of two-phase methodologies. In the first phase, the pupil's diameter is calculated by applying stimulus (light) in one eye of the subject and calculating the constriction of the pupil size on both eyes in different video frames. The above measurement is converted into feature space using Kohn and Clynes model-defined parameters. The Support Vector Machine is used to classify legitimate subjects when the diameter change is normal (or when the eye is alive) or illegitimate subjects when there is no change or abnormal oscillations of pupil behavior due to the presence of printed photograph, video, or 3D mask of the subject in front of the camera. In the second phase, we perform the facial recognition process. Scale-invariant feature transform (SIFT) is used to find the features from the facial images, with each feature having a size of a 128-dimensional vector. These features are scale, rotation, and orientation invariant and are used for recognizing facial images. The brute force matching algorithm is used for matching features of two different images. The threshold value we considered is 0.08 for good matches. To analyze the performance of the framework, we tested our model in two Face antispoofing datasets named Replay attack datasets and CASIA-SURF datasets, which were used because they contain the videos of the subjects in each sample having three modalities (RGB, IR, Depth). The CASIA-SURF datasets showed an 89.9% Equal Error Rate, while the Replay Attack datasets showed a 92.1% Equal Error Rate.

KEYWORDS
SIFT; pupil; CASIA-SURF; pupillary light reflex; replay attack dataset; brute force

1 Introduction

Biometric measures are used to secure the digital world. None of the two people has the same biometric print. The Biometric is a combination of two words, bio and metric. It deals with the physiological and behavioral characteristics of a person. Digitization needs a robust, authentic system due to more use of digital devices as it makes the data readily available to any corner of the world.

Due to increasing cybersecurity risks, decreasing hardware cost, and voluminous data, many organizations prefer to use biometric authentication systems due to their advantages over the traditional way of authentication like signature or password. The goal of such a system is to ensure that the available applications are accessed or used only by a genuine user and not by others. Such services include computer systems, secure access to buildings, cell phones, smartphones, laptops and ATMs. In the lack of a robust individual recognition system, these things remain susceptible to the tricks of an impostor. This is the reason why biometrics has been adopted in many applications. The physiological and behavioral characteristics used for biometric recognition include fingerprint, hand geometry, iris, retina, Face, palmprint, ear, DNA, voice, gait, signature, keystroke dynamics, etc. Because biometrics involve pattern recognition, the organ that gives patterns is mostly used for biometrics. However, the main issues are involved in designing and commissioning a practical biometric system. Table 1 lists the most commonly used biometrics. These identifiers are also called mature biometrics. Table 1 gives some important biometrics that are commonly used.

Table 1: Commonly used biometrics

Some common biometrics	
Physiological	Behavioral
Face	Signature
Iris	Voice
Fingerprint	GAIT
Palm	Keystrokes dynamics

1.1 Face Biometrics

The face is considered one of the most popular biometric models for authentication purposes. The emergence of many face recognition conferences like Audio and Video-Based Authentication international conference, Automatic Face and Gesture Recognition Conference, and systematic empirical Face Evaluations Techniques (FRT), including those reported by Grother et al. 2019 [1], Phillips et al. 2003 [2,3], etc., make it more popular.

Facial Recognition Technology (FRT) is designed to authenticate a person by using facial images without any physical contact. There are two major steps for designing FRT. In the first stage, the enrollment of the subject takes place by extracting features of the facial image using feature extraction or some image processing algorithm and creating a template. In the second stage. When the subject comes for authentication, the facial feature is extracted, and using some pattern-matching algorithm. It is matched with the template. While taking or stealing someone's biometric traits is difficult, it is still possible for fraud to circumvent a biometric system using spoofed or artificial traits. A large number of studies have shown that it is quite possible to construct gluey fingers using lifted fingerprint impressions and utilize them to attack the biometric system. Behavioral biometrics like voice and signatures are more susceptible to such attacks than physiological traits. The facial recognition system is vulnerable to several presentation attacks (direct attacks or spoof attacks). A presentation attack uses Fake faces or facial artifacts for unauthorized access using a facial recognition authentication system. A presentation attack may be dynamic or static in nature. It is applied in two-dimensional as well as three-dimensional images.

In two-dimensional static presentation attacks, an attacker may use a photograph or a flat paper plastic mask as an artifact. On the other side, in a two-dimensional dynamic attack, the fraud perpetrator uses a screen video display or several photographs shown one by one. The three-dimensional presentation attack is also static or dynamic in nature. Static attacks occur using sculpture or 3D print, whereas attacks using robots or well-prepared makeup come under dynamic attack.

1.2 About Pupillary Light Reflex of Pupil

Fig. 1 shows the anatomy of the human eye. Human eyes consist of many parts that can be used in biometrics, such as the iris, pupil, retina, eye movement, etc. The cornea is the front, white, dome-shaped part that is responsible for focusing light when it falls on the eyes. The anterior chamber is present behind the cornea and is filled with a fluid called the aqueous humous. The iris is the colored part of the eyes that is present behind the anterior chamber, having a dark hole at its center called the pupil. When light falls on the eye, the iris muscle constricts and dilates the pupil to control the amount of light entering the eye. The pupillary light reflex is the involuntary reflex that controls the diameter of the pupil. The pupil is one of the most important parts of our vision system. The size of the pupil dilates or gets bigger when the light intensity is dim and it constricts or gets smaller when the intensity of light is brighter.

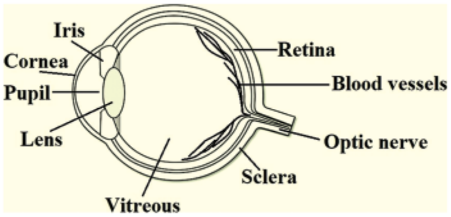

Figure 1: Different parts of the eye

By calculating the diameter of the pupil in different video frames of the subject and comparing the diameter. If we find the difference in diameter in two or more frames, then liveness will be verified as pupillary light reflex changes the diameter of the pupil, and this is not possible with the artifacts as shown in Fig. 2. As the Pupillary light reflex causes the change in the diameter of the pupil, it is now one of the hot research areas in detecting the liveness of the face as well as iris biometric models.

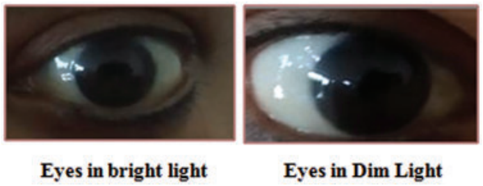

Figure 2: Response of the pupil to light

2 Literature Review

The initial work for facial recognition was found in the 1954's by Bruner and Tagiuri in the handbook of psychology about the perception of people [4]. Some of the earliest works include Darwin's works on the functionality of emotions in the year 1972. According to Darwin, facial expression evolved as the result of certain kinds of emotions and had an important communicative function [5]. Galton works on facial profile-based biometrics in the year 1988.

As the facial recognition system is one of the most popular biometrics among all different biometrics, it is more vulnerable to direct attacks and indirect attacks because of progressive technology. Table 2 depicts the important contribution of the same. A number of researches were going on over the last four decades in order to strengthen the biometric authentication system. Different biometric traits have their own pros and cons. In order to make biometric authentication systems more robust nowadays, two or more biometric traits have been used. Using more than one biometric is called multimodal biometric. Combining iris pattern and pupil is also used for authenticating purposes.

Table 2: Related work

Author	Year	Presentation attack method	Presentation attack detection (PAD)
Majeed et al. [6]	2023	Photo	Proposed a PAD model called texture feature analysis using Local Phase Quantization descriptor (LPQ).
Muhammad et al. [7]	2023	Video	Proposed a framework that uses data sampling scheme on multiple video frame.
Liu et al. [8]	2022	Photo	Proposed a model called Pose Independent Face-Anti Spoofing that check the liveness of face using pose and appearance.
Samar et al. [9]	2022	Photo	Proposed a MultimodalTransformer based detection of presentation attack (MFAST).
Shibel et al. [10]	2022	Video	Using RESNET-50 deep learning algorithm.
Chou [11]	2021	Photo and Video	Proposed a framework to analyse speech text and lip motion simultaneously for verification of face liveness.

(Continued)

Table 2 (continued)

Author	Year	Presentation attack method	Presentation attack detection (PAD)
Makowski et al. [12]	2021	Video	Framework that uses CNN to process binocular eye response for liveness detection.
Xu et al. [13]	2021	Photo and Video	Proposed a model to improve existing PAD method by designing two subnetworks that learn motion patterns from images and texture from video.
González-Soler et al. [14]	2021	Short wave infrared images	Propose a framework using CNN for preprocessing multispectral information.
Arashloo [15]	2020	Photo and Video	Designed a framework kernel Fisher null-space facial PAD.
Sun et al. [16]	2020	3D mask attack	Design a polarization MWIR imaging system that capture facial image using polarization imaging.
Sanghvi et al. [17]	2020	Photo	Framework called Mixnet-deep learning based algorithm.
George et al. [18]	2019	Photo	Propose a multi-channel CNN-based approach for presentation attack detection.

The pupil diameter changes with the intensity of light. This pupillary light reflex is used as one trait for authenticating people's identity [19,20]. Biometric authentication techniques are also used for identifying animals like sheep using retinas [21]. The initial work for facial recognition was found in the 1954's by Bruner and Tagiuri in the handbook of psychology about the perception of people [22]. Some of the earliest works include Darwin's works on the functionality of emotions in the year 1972. According to Darwin, facial expression evolved as the result of certain kinds of emotions and had an important communicative function [21]. Galton works on facial profile-based biometrics in the year 1988. Actual work on automatic facial recognition using machines started in the year 1970s by Kelly, in which complex processing tasks were performed on a picture taken from television for the first time [23]. The shallow Method of Face detection and recognition does not use a deep learning method; instead, extracting the feature from an image using handcrafted image descriptors like SIFT, MOPS, GLOH, and LBP [23–27] and combining these local descriptors using pooling mechanism to generate overall face descriptors like Fisher Vectors [28,29]. Face recognition and detection have been done by using different approaches like locality preserving projections LPP [30], and modular PCA [31] in which the face image is divided into sub-images and the PCA method is applied to each sub-image. The improved recognition rate is found under different lighting directions, illumination expressions, and poses trained in different data sets [32,33]. Some facial recognition techniques use HOG feature extraction and fast PCA enhancing accuracy rate and Support Vector Machine used for recognizing faces [34]. The unified LDA/ PCA algorithm combines LDA and PCA and reduces the drawbacks of LDA in facial recognition systems. Learning Based Descriptor [35] reduces the number of issues that occur during the matching and representation of facial images and is highly discriminative, compact, and easy to extract [36]. Content-based facial recognition technique is also gaining popularity [37,38]. Table 3 represents some important method used by early researchers.

Table 3: FRT existing technology

Data sets	Existing algorithm	Features/Recognition rate (in percentage)
M2VTS	[39]	Rule based Frontal face detection-87
	[40]	Super resolution algorithm by using Bayesian estimation-89.5
ORL	[41]	Threshold based LBP-98
	[39]	CNN 98.3
Yale2B	[42]	Neural network with softmax classifier-96
	[19]	Hybrid SIFT 96
	[26]	Particle swarm optimization-92
	[43]	Fusion of covariance matrix and entropy matrix-96
FERET	[22]	Shape and Texture information-94
FACE 94	[44]	SIFT and SURF algorithm-93
	[45]	Artificial neural network-92
	[23]	Holistic Fourier invariant features-96

3 Proposed Methodology

Our proposed algorithm consists of two main steps. In the first steps, we will find the pupil diameter for verifying the liveness of a subject. After verifying the subject is live, the facial recognition algorithm runs to authenticate the person's identity.

Measuring minute fluctuations in pupil diameter change in response to a stimulus is called Pupillometry. The measurement of diameter can be done using digital image processing.

1. Capture facial images under different illumination conditions.
2. Extract the iris portion in order to calculate the diameter of the pupil using Matlab image processing.
3. Find the diameter of the pupil in five different frames under different illumination conditions.
4. Compare the diameter of the pupil in different frames.
5. Extract the facial feature vector using SIFT and MOPS.
6. Train the model to find the embedding function of images.
7. Train the proposed system to obtain the classifier model using SIFT feature vectors.
8. Test the classifier by giving query image feature vector.
9. Calculate the similarity index of the query image and the images of data sets using the classifier.

3.1 Recognition of Pupillary Light Reflex

In this step, we detect the pupil inside the captured image, and the size is calculated using the segmentation process.

3.1.1 Detection and Localization

Detection and Localization are the two important steps, in which detection confirms the presence og the pupil inside the frame and detection gives its position. The Hough transform is used to find the boundary between the iris and the pupil. Pupil diameter has been calculated using MATLAB Image Processing toolbox that provides several algorithm environment tools for processing images, analyzing images, visualization, as well as developing algorithms.

The different Image Processing steps involved are for finding the pupil dynamics are:

- Preprocessing
- Segmentation
- Data Processing.
- Feature extraction

The transform technique is modified to be sensitive towards a dark circular shape instead of any other light irregular shape. Using gradient and sensitivity, this proposed algorithm becomes robust as it detects the pupil even after the eyelashes half cover it.

3.1.2 Artifacts Removal

Two types of noises are detected in the raw linear pupil radii signal. The first one is Noise generated at the time of pupil detection because of the blinks called pupil detection error, and the second one is Pupil segmentation error that arises due to eye motion, Non-circular pupil, Partially covered pupil, and off-axis gaze. Segmentation error is very difficult to identify, and it is only observed when a sudden change in radii is marked compared to the previous neighboring frame, whereas detection error we can rectify during the modeling of pupil dynamics. For building the classification model, we use the support vector machine as it is considered one of the best classifiers that perform very well in low-dimensional feature vectors.

3.2 Proposed Facial Recognition Technology

The captured image contains different types of noises. The specular noises produced over the eye's surface due to infrared illumination have been removed using the different noise removal functions of the image processing tools of Matlab. The RGB images are converted into binary images and the circular region is calculated for the iris and the pupil. The Hough transform is used for extracting the features as it finds the instances of objects within a certain class of shapes.

All facial images have certain content that describes the images and differentiates them from other facial images, and this is called Image features. Edges, corner or interest points, BLOB, and ridges provide rich information about image content. Image features are used as input for the facial recognition process following major steps:

- Detection of face using a camera either solo or from a crowd.
- Analyzing the geometry of the face using a certain algorithm. Geometry includes the depth of eyes, the distance between eyes and eyebrows, the distance between chin and forehead, the contour of nose, lips, ears, etc.
- After analyzing the face, the mathematical form called face print is generated that contains the digital information of the facial features.
- Finally, finding the match of the given person. Matching Image is an important task for a facial recognition system.

After getting the feature descriptor, the main task is to use the key points of the feature descriptor for matching purposes. The main goal of matching is that it would be able to perfectly match the same images if they are taken from different angles, different viewpoints, different scales, and different camera parameters. Panorama stitching of two different images of the same scene or object and finding key points for matching purposes is the main task of the matching algorithm.

3.2.1 SIFT Algorithm

Scale Invariant Feature Transform is a feature extractor algorithm that is invariant to scaling and rotation. SIFT is invariant to rotation, viewpoint, and illumination and gives good results. Continuous improvement is going on SIFT to enhance the performance as well as accuracy. The number of variants introduced with making certain changes in the steps or pipeline of the SIFT algorithm. MOPS, also called Multi-Scale Oriented Patches Descriptor is one of the variants of the SIFT feature descriptor in which patches around the key points is rotated according to the dominant gradient orientation, and after that, the histogram and descriptor are computed. Multi-Scale Oriented Patches consist of normalized patches oriented via blurred local gradient and the features are positioned at Harris Corner. This is useful because by rotating the patch to its dominant gradient orientations all key points have canonical orientation same [20]. The SIFT algorithm mainly have following four steps:

I) Constructing scale space extrema
II) Localizing key point
III) Estimating orientation
IV) Keypoint descriptor.

3.2.2 Scale Space Extrema Construction

Using the Difference of Gaussian (DOG), the scale space extrema step locates an intriguing point (invariant to scale and orientation). DOG findings that approximate the Laplacian of the Gaussian. Discovering the value of the pixel that is the maximum or minimum value in the surrounding scale images and around its spatial region is referred to as discovering the extrema in scale space. Extrema Construction first creates a scale space for an image for scale space. Scale-space is created by taking an image and convolving it with the Gaussian. Using the k-time Gaussian operator, convolve a second image of the same scale.

All these groups of images form an octave. An octave consists of the number of images depending on the value of k. Again, subsample the image and repeat the same process. The scale space group of blurred images of different scales. Blurring is referred to as a technique in which the convolution of the Gaussian operator with every pixel of the image takes place to output blurred images [46].

To find the extrema, the Laplacian idea is used to find a Gaussian difference by taking the difference between successive Gaussian images and constructing this set for every octave.

3.2.3 Keypoint Localizattion

This step is for finding the exact location of minima and maxima ie whether it is present either between the two coordinates or between the two scales.

3.2.4 Determination of Exact Location and Scale

Taylor Series Expansion is used to find the exact location. To determine the location of the extrema, the derivatives compute the first derivative and second derivative simply by finite differences to find where the actual extrema exists. The low contrast point, as well as the edge point, need to be removed.

3.2.5 Edge Elimination

Very similar to Harris Corner, Lowe [47] proposed using Hessian of D finding curvature or sharp changes in different direction. Hats Eigen value of Hessian also give good estimate for corner estimation.

$$H = \begin{bmatrix} D_{X_X} & D_{X_Y} \\ D_{X_Y} & D_{X_Y} \end{bmatrix} \tag{1}$$

$$Tr(H) = D_{x_x} + D_{y_y} = \alpha + \beta$$

$$Det(H) = D_{x_x} D_{y_y} - D^2_{x_y} = \alpha\beta \tag{2}$$

Evaluate ratio:

$$Tr(H)^2/Det(H) = (\alpha + \beta)^2/\alpha\beta = (r\beta + \beta)^2/r\beta^2 \tag{3}$$

$$Tr(H)^2/Det(H) = (r + 1)^2/rwherer = \alpha/\beta \tag{4}$$

The $Tr(H)^2/Det(H)$ is minimum when r = 1,when r = 1 then α and β are close to each other or equally high that is useful for finding corner point. SIFT proposes that reject a keypoint if $Tr(H)^2/Det(H)$ > a threshold value.

At the last of this step exact location and scale at every extrema point as well as also selected some stable keypoint by rejecting edges and low contrast point [47].

3.3 Orientation Estimation

For rotation invariance orientation estimation is required. This step is used to find the orientation of extrema. For this use scale of point to choose appropriate image:

$$\hat{I}(x, y) = G(x, y, \sigma) * I(x, y) \tag{5}$$

After that using finite differences compute gradient magnitude and orientation:

$$m(x, y) = \sqrt{(\hat{I}(x + 1, y) - \hat{I}(x - 1, y))^2 + (\hat{I}(x, y + 1) - \hat{I}(x, y - 1))^2} \tag{6}$$

$$\theta(x, y) = tan^{-}1((\hat{I}(x, y + 1) - \hat{I}(x, y - 1))/(\hat{I}(x + 1, y) - \hat{I}(x - 1, y))) \tag{7}$$

where gradient magnitude is $m(x, y)$ at every point and $\theta(x, y)$ is orientation gradient. The result of this step is magnitude and orientation for every point in a image.

3.4 Histogram Creation

To create a histogram, select the area surrounding the key point and consider ă the orientation and magnitude of all the other key topics. Histogram input is weighted using gradient magnitude and the Gaussian function to increase performance. The histogram will be less dense if certain spots are farther away. Instead of a single point, local factors determine the ă peak.

If another peak is within 80 percent of max peak then this also as keypoint with different direction [48].

3.5 Key-Point Descriptor

Descriptor gradient information is utilized to locate critical points. Take the 16 ∗ 16 window size with the nearby key point found for this. This 16 ∗ 16 window should be divided into four 4 ∗ 4 quadrants. Create a histogram of the position of the points inside these quadrants and align it with eight bins. Compared to the closer point, the further point makes up less of the histogram. Create a gradient orientation histogram using 8 histogram bins for each 4 by 4 quadrant. The raw version key point Descriptor is formed at the end of this 128 non-negative vector. 16 histograms with 8 values each produced 128 non-negative vectors. The 128-dimensional feature vector is created simply by gradient-orienting the area surrounding the key point. In order to decrease the effects of contrast, normalization to unit length is done for 128-d vector. The values are clipped to 0.2 and the resulting vector is once again normalized to unit length to make the descriptor robust to different photometric variations.

3.6 Results and Discussion

In this proposed system, two popular antispoofing image and video data sets Casia-surf, Replay attack dataset are used to perform test on our proposed algorithm.

There are four decisions every biometric system gives. (1) Authorized a legitimate person called True Positive. (2) Authorized an illegitimate person called False Positive. (3) Deny an illegitimate person called True Negative. (4) Deny a legitimate person called False Negative. Our proposed method is evaluated in terms of performance using two different classifiers: decision trees and random forests. Table 4 analysis reveals that decision tree classifiers are more accurate than random forest classifiers in the M2VTS, ORL, and Face 94 databases, respectively, while table analysis also reveals that the accuracy of random forest classifiers is higher in the Yale 2B, FERET, and Face 94 databases. Results for M2VTS, ORL, and FERET are encouraging.

Table 4: Datasets details used in the proposed work for facial recognition purpose

Database	Number of classes	Total number of images
M2VTS Messer et al.	60 classes and 10 samples from each class	600
ORL Database	120 classes and 10 samples from each class	1200
Yale2B Database	230 classes and 20 samples from each class	4600
FERET Database	32 classes and 64 samples from each class	2048
Face 94	80 subjects and 10 samples from each class	800

Small data sets were the primary cause of Face 94's performance decline. When it comes to true positive rates, the random forest classifier performs better with Yale2B data sets and the decision tree performs better with M2VTS data sets. The performance is shown in Table 5 as an equal error rate. The classifier's false positive rate across all datasets is displayed in Tables 6 and 7. The high AUC in Table 8 indicates that the suggested feature set is effective for the facial recognition system. The performance evaluation in terms of execution time is shown in Table 9. The proposed algorithm is compared with different existing algorithms and significant improvements are found for M2VTS, YALE 2B, and Face 94 data sets, and performance diminishes in the case of FERET and ORL.

Table 5: Performance evaluation (equal error rate)

Decision tree recognition accuracy					
Feature extraction method	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	84	79.5	67	68	51.8
REPLAY ATTACK (32)	86.5	80.5	73.6	71	67.7
Random forest classifier					
Feature extraction method	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	78	72.5	77	82	61.5
REPLAY ATTACK	82	78.5	78.5	81.5	77.7

Table 6: Performance table true positive rate for random forest classifier

Random forest classifier					
Feature extraction method	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	78	72.5	97	72	81.5
REPLAY ATTACK	82	78.5	98.5	71.5	82.7
Decision tree classifier					
Feature extraction method	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	98	72.5	87	82	61.5
REPLAY ATTACK	92	78.5	88.5	81.5	77.7

Table 7: Performance table false positive rate for random forest classifier

Random forest classifier					
Antispoofing dataset	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	0.02	0.03	0.03	0.02	0.02
REPLAY ATTACK	0.07	0.03	0.02	0.03	0.03
Decision tree classifier					
Antispoofing dataset	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	0.01	0.02	0.01	0.02	0
REPLAY ATTACK	0.01	0.02	0.05	0.05	0.01

Table 8: Area under curve

Decision tree classifier					
Antispoofing dataset	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	94	99.5	87	88	81.8
REPLAY ATTACK	96.5	91.5	83.6	83	87.7
Random forest classifier					
Antispoofing dataset	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	79	82.5	97	92	91.5
REPLAY ATTACK	82	88.5	98.5	91.5	97.7

Table 9: Evaluation time in seconds (decision tree)

Decision tree classifier					
Antispoofing dataset	M2VTS	ORL	Yale2B	FERET	Face 94
CASIA-SURF	0.04	0.17	0.23	0.24	0.28
REPLAY ATTACK	0.19	0.16	0.16	0.25	0.18

4 Conclusion

A strong facial authentication method is suggested in this study. Two-way authentication is suggested in this architecture. By measuring the change in pupil diameter under various lighting conditions during the image-capturing process, the liveliness of the facial images is first validated. Following the confirmation of liveness, the facial recognition system integrates SIFT and MOPs, two distinct feature descriptors, into a deep neural network architecture for recognition and detection. The accuracy, area under the curve, false positive rate, and true positive rate of two distinct classifier decision trees and random forests are measured. The Yale2B, M2VTS, FERET, ORL, and FACE 94 datasets are used for the training, which has been demonstrated to be computationally effective. By analyzing the results, it is found that the proposed algorithm is an efficient and acceptable method for facial recognition. To sum up, this technique appears to be a strong contender for liveness detection, and it has a lot of potential for real-world application.

5 Future Scope

Our proposed algorithm has three limitations and opens the path for future work in this area. First, this algorithm does not include the measurement of elderly people as more elusive changes in their pupil size. Second, the dynamics feature measurement takes time, and the devices designed to capture iris images do not allow additional time during capturing. The third one is drug or alcohol ingestion, which also alters pupil dynamics.

Acknowledgement: The authors present their appreciation to King Saud University for funding the publication of this research through Researchers Supporting Program (RSPD2023R809), King Saud University, Riyadh, Saudi Arabia.

Funding Statement: The research is funded by Researchers Supporting Program at King Saud University (RSPD2023R809).

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Puja S. Prasad, Ali Wagdy Mohamed; data collection: Adepu Sree Lakshmi, Hossam M. Zawbaa; analysis and interpretation of results: Sandeep Kautish, Abdulaziz S. Almazyad; draft manuscript preparation: Simar Preet Singh, Rajesh Kumar Shrivastava. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: In this paper, we used open access data, which is available in open repositories for researchers.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Grother, P., Ngan, M., Hanaoka, K. (2019). *Face recognition vendor test (FRVT): Part 3, demographic effects*. Gaithersburg, MD, USA: National Institute of Standards and Technology.
2. Phillips, P. J., Grother, P., Micheals, R. (2011). Evaluation methods in face recognition. In: *Handbook of face recognition*, pp. 551–574. London: Springer.
3. Phillips, P. J., Grother, P., Micheals, R., Blackburn, D. M., Tabassi, E. et al. (2003). Face recognition vendor test 2002. *2003 IEEE International SOI Conference*, Newport Southeastern Rhode Island, USA, IEEE.
4. Bruner, J., Tagiuri, R. (1954). The perception of people. In: lmdzey, G. (Ed.), *Handbook of social psychology*, vol. 2. NY, USA: John Wily & Sons, Inc.
5. Kaiser, S., Wehrle, T. (2001). Facial expressions as indicators of appraisal processes. *Appraisal Processes in Emotion: Theory, Methods, Research, 1,* 285–300.
6. Majeed, Q., Fathi, A. (2023). A novel method to enhance color spatial feature extraction using evolutionary time-frequency decomposition for presentation-attack detection. *Journal of Ambient Intelligence and Humanized Computing, 14(4),* 3853–3865.
7. Muhammad, U., Oussalah, M. (2023). Face anti-spoofing from the perspective of data sampling. *Electronics Letters, 59(1),* e12692.
8. Liu, A., Wan, J., Jiang, N., Wang, H., Liang, Y. (2022). Disentangling facial pose and appearance information for face anti-spoofing. *2022 26th International Conference on Pattern Recognition (ICPR)*, Montreal, QC, Canada, IEEE.
9. Samar, A. R., Farooq, M. U., Tariq, T., Khan, B., Beg, M. O. et al. (2022). Multi-modal face anti-spoofing transformer (MFAST). *2022 19th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, Islamabad, Pakistan, IEEE.
10. Shibel, A. M., Ahmad, S. M. S., Musa, L. H., Yahya, M. N. (2022). Deep learning detection of facial biometric presentation attack. *LIFE: International Journal of Health and Life-Sciences, 8(2),* 61–78.
11. Chou, C. L. (2021). Presentation attack detection based on score level fusion and challenge-response technique. *The Journal of Supercomputing, 77,* 4681–4697.

12. Makowski, S., Prasse, P., Reich, D. R., Krakowczyk, D., Jäger, L. A. et al. (2021). Deepeyedentificationlive: Oculomotoric biometric identification and presentation-attack detection using deep neural networks. *IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(4),* 506–518.
13. Xu, Y., Wang, Z., Han, H., Wu, L., Liu, Y. (2021). Exploiting non-uniform inherent cues to improve presentation attack detection. *2021 IEEE International Joint Conference on Biometrics (IJCB)*, Shenzhen, China, IEEE.
14. González-Soler, L. J., Gomez-Barrero, M., Kolberg, J., Chang, L., Pérez-Suárez, A. et al. (2021). Local feature encoding for unknown presentation attack detection: An analysis of different local feature descriptors. *IET Biometrics, 10(4),* 374–391.
15. Arashloo, S. R. (2020). Unseen face presentation attack detection using sparse multiple kernel fisher null-space. *IEEE Transactions on Circuits and Systems for Video Technology, 31(10),* 4084–4095.
16. Sun, P., Zeng, D., Li, X., Yang, L., Li, L. et al. (2020). A 3D mask presentation attack detection method based on polarization medium wave infrared imaging. *Symmetry, 12(3),* 376.
17. Sanghvi, N., Singh, S. K., Agarwal, A., Vatsa, M., Singh, R. (2021). Mixnet for generalized face presentation attack detection. *2020 25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy, IEEE.
18. George, A., Mostaani, Z., Geissenbuhler, D., Nikisins, O., Anjos, A. et al. (2019). Biometric face presentation attack detection with multi-channel convolutional neural network. *IEEE Transactions on Information Forensics and Security, 15,* 42–55.
19. Mohanraj, V., Vimalkumar, M., Mithila, M., Vaidehi, V. (2016). Robust face recognition system in video using hybrid scale invariant feature transform. *Procedia Computer Science, 93,* 503–512.
20. Brown, M. A. (2005). *Multi-image matching using invariant features*. University of British Columbia: Citeseer.
21. Cinbis, R. G., Verbeek, J., Schmid, C. (2011). Unsupervised metric learning for face identification in TV video. *2011 International Conference on Computer Vision*, Barcelona, Spain, IEEE.
22. Ahonen, T., Hadid, A., Pietikäinen, M. (2004). Face recognition with local binary patterns. *European Conference on Computer Vision*, Berlin, Heidelberg, Springer.
23. Lai, J. H., Yuen, P. C., Feng, G. C. (2001). Face recognition using holistic fourier invariant features. *Pattern Recognition, 34(1),* 95–109.
24. Yan, S., Chang, S., Wang, J., Azhar, S. (2020). Using pupil light reflex for fast biometric authentication. *Proceedings of the ACM Turing Celebration Conference*, Hefei, China.
25. Lu, C., Tang, X. (2015). Surpassing human-level face verification performance on lfw with gaussianface. *Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin, Texas, USA.
26. Krisshna, N. A., Deepak, V. K., Manikantan, K., Ramachandran, S. (2014). Face recognition using transform domain feature extraction and PSO-based feature selection. *Applied Soft Computing, 22,* 141–161.
27. Wolf, L., Hassner, T., Maoz, I. (2011). Face recognition in unconstrained videos with matched background similarity. *CVPR 2011*, Colorado Springs, USA, IEEE.
28. Simonyan, K., Parkhi, O. M., Vedaldi, A., Zisserman, A. (2013). Fisher vector faces in the wild. *BMVC 2013–Electronic Proceedings of the British Machine Vision Conference 2013*, vol. 2. UK.
29. Li, X. Y., Lin, Z. X. (2017). Face recognition based on HOG and fast PCA algorithm. *The Euro-China Conference on Intelligent Data Analysis and Applications*, Málaga, Spain, Springer.
30. He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H. J. (2005). Face recognition using laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(3),* 328–340.
31. Li, L., Liu, S., Peng, Y., Sun, Z. (2016). Overview of principal component analysis algorithm. *Optik, 127(9),* 3935–3944.

32. Kuruvayil, S., Palaniswamy, S. (2022). Emotion recognition from facial images with simultaneous occlusion, pose and illumination variations using meta-learning. *Journal of King Saud University–Computer and Information Sciences, 34(9),* 7271–7282.
33. Yin, X., Liu, X. (2017). Multi-task convolutional neural network for pose-invariant face recognition. *IEEE Transactions on Image Processing, 27(2),* 964–975.
34. Monika, Kumar, M., Kumar, M. (2021). XGBoost: 2D-object recognition using shape descriptors and extreme gradient boosting classifier. *Computational Methods and Data Engineering*, pp. 207–222. Delhi-NCR, India, Springer.
35. Boussaad, L., Boucetta, A. (2022). Deep-learning based descriptors in application to aging problem in face recognition. *Journal of King Saud University–Computer and Information Sciences, 34(6),* 2975–2981.
36. Caro, M. A. (2019). Optimizing many-body atomic descriptors for enhanced computational performance of machine learning based interatomic potentials. *Physical Review B, 100(2),* 024112.
37. Chhabra, P., Garg, N. K., Kumar, M. (2020). Content-based image retrieval system using ORB and SIFT features. *Neural Computing and Applications, 32,* 2725–2733.
38. Bansal, M., Kumar, M., Kumar, M., Kumar, K. (2021). An efficient technique for object recognition using shi-tomasi corner detection algorithm. *Soft Computing, 25,* 4423–4432.
39. Kotropoulos, C., Pitas, I. (1997). Rule-based face detection in frontal views. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4. Munich, Germany, IEEE.
40. Sezer, O. G., Altunbasak, Y., Ercil, A. (2006). Face recognition with independent component-based super-resolution. *Visual Communications and Image Processing 2006*, vol. 6077. San Jose, California, USA, SPIE.
41. Meng, J., Gao, Y., Wang, X., Lin, T., Zhang, J. (2010). Face recognition based on local binary patterns with threshold. *2010 IEEE International Conference on Granular Computing*, San Jose, California, USA, IEEE.
42. Zhang, Z., Li, J., Zhu, R. (2015). Deep neural network for face recognition based on sparse autoencoder. *2015 8th International Congress on Image and Signal Processing (CISP)*, Shenyang, China, IEEE.
43. Wang, S., Liu, P. (2015). A new feature extraction method based on the information fusion of entropy matrix and covariance matrix and its application in face recognition. *Entropy, 17(7),* 4664–4683.
44. Vinay, A., Hebbar, D., Shekhar, V. S., Murthy, K. B., Natarajan, S. (2015). Two novel detector-descriptor based approaches for face recognition using sift and surf. *Procedia Computer Science, 70,* 185–197.
45. Réda, A., Aoued, B. (2004). Artificial neural network-based face recognition. *First International Symposium on Control, Communications and Signal Processing*, Hammamet, Tunisia, IEEE.
46. Lindeberg, T. (2013). A framework for invariant visual operations based on receptive field responses. *2013: Fourth International Conference on Scale Space and Variational Methods in Computer Vision*, Schloss Seggau, Graz Region, Austria.
47. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60(2),* 91–110.
48. Nixon, M., Aguado, A. (2019). *Feature extraction and image processing for computer vision*. Elsevier Science: Academic Press.