Multimedia Tools and Applications (2023) 82:40039–40063
https://doi.org/10.1007/s11042-023-14453-7

Face presentation attack detection performances of facial regions with multi-block LBP features

Asuman Günay Yılmaz[1] · Uğur Turhal[2] · Vasif Nabiyev[1]

Received: 14 August 2021 / Revised: 28 April 2022 / Accepted: 31 January 2023/
Published online: 30 March 2023

Abstract

Biometric recognition systems are frequently used in daily life although they are vulnerable to attacks. Today, especially the increasing use of face authentication systems has made these systems the target of face presentation attacks (FPA). This has increased the need for sensitive systems detecting the FPAs. Recently surgical masks, frequently used due to the pandemic, directly affect the performance of face recognition systems. Researchers design face recognition systems only from the eye region. This motivated us to evaluate the FPA detection performance of the eye region. Based on this, in cases where the whole face is not visible, the FPA detection performance of other parts of the face has also been examined. Therefore, in this study, FPA detection performances of facial regions of wide face, cropped face, eyes, nose, and mouth was investigated. For this purpose, the facial regions were determined and normalized, and texture features were extracted using powerful texture descriptor local binary patterns (LBP) due to its easy computability and low processing complexity. Multi-block LBP features are used to obtain more detailed texture information. Generally uniform LBP patterns are used for feature extraction in the literature. In this study, the FPA detection performances of both uniform LBP patterns and all LBP patterns were investigated. The size of feature vector is reduced by principal component analysis, and real/fake classification is performed with support vector machines. Experimental results on NUAA, CASIA, REPLAY-ATTACK and OULU-NPU datasets show that the use of all patterns increased the performance of FPA detection.

Keywords Face presentation attack detection · LBP · Face regions · Texture analysis

✉ Asuman Günay Yılmaz
gunaya@ktu.edu.tr

[1] Karadeniz Technical University, Trabzon, Turkey

[2] Bayburt University, Bayburt, Turkey

1 Introduction

Due to the rapid development in technology, today various biometric systems are widely used in real-life applications such as online payment security, e-commerce security, smartphone-based authentication, secure access control, passport, and border controls [30]. Although variability in human-face appearances due to changes in the viewing direction [32], face recognition systems have been among the most studied technologies since the 90s because of their advantages over other biometric features [39]. Thanks to their ease of use, near-perfect performance, and high-security levels, face recognition systems are among the most widely used biometric systems in the market compared to iris and fingerprint recognition systems [15]. However, the increasing use of face authentication systems has made these systems the target of Face Presentation Attacks (FPA). FPA is the general name given to malicious attempts to impersonate another person (impersonation attack) or avoid recognition by the system (obfuscation attack).

The increase in face recognition-based access control systems has raised the need for sensitive systems to detect FPAs. FPA is an attacker's attempt to be authenticated in an identification process by impersonating someone else. Today, it has become easier to access images/videos or detailed information on how to create fake masks, which can be used to spoof facial recognition systems [18]. This also causes an increase in the variety of attacks that the existing systems may encounter. Attacks are no longer limited to theoretical or academic fields but are often carried out against real applications. For example, Apple's iPhone 5S model smartphones with integrated fingerprint reader hardware are spoofed with fake fingerprint samples only one day after being put on the market [14]. Therefore, FPA detection has now become a mandatory part of face recognition systems. Figure 1 shows the complexity of FPA detection problem.

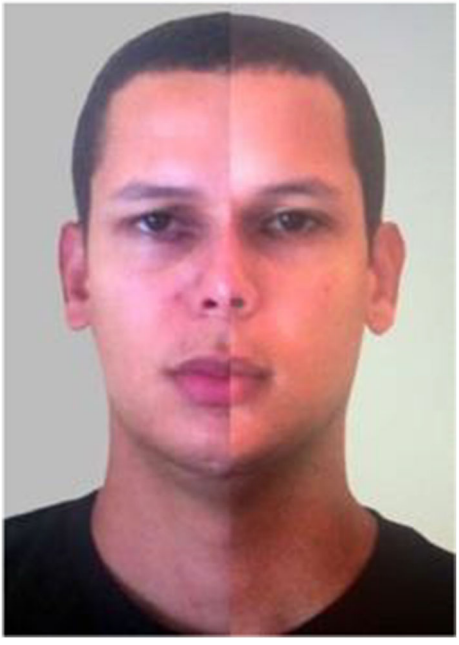


Fig. 1 Sample image with real and fake face parts. Which part is real? (The answer is left) [34]

Surgical masks, which have recently become inseparable parts of our lives due to Covid-19, directly affect the performance of face recognition systems. In cases where the use of masks is mandatory today, and in the future, it is inevitable to develop face recognition systems over the face area outside the mask. Systems that perform recognition over the eye region will be spoofed through the eye region. For this reason, in this study, the FPA performances of various facial regions (wide face region, cropped face region, eye region, nose region, mouth region) were investigated. In the method, multi block uniform local binary pattern (MB-LBP) features were extracted from facial regions and support vector machines (SVM) classifier was used for real/fake classification. Then, principal component analysis (PCA) was used to reduce the size of the features and classified with SVM. Finally, the effects of all LBP features (not only uniform patterns) on FPA detection were investigated.

The main contributions of this study are:

- The FPA detection performances of the wide face, the cropped face and especially the eyes, nose, and mouth regions were examined.
- Uniform patterns are generally used in feature extraction with LBP. In the study, the effects of all LBP patterns on FPA detection were examined.
- In the study, 30 different attack scenarios in 4 datasets were evaluated separately. Experiments were also carried out for the scenarios where all attack types are used together.
- The effect of reducing the size of regional texture features with PCA on FPA detection performance was investigated.

The remainder of the paper is organized as follows: FPA detection studies in the literature are summarized in Section 2. The datasets and the applied method are explained in Section 3. Experimental studies and results are given in Section 4, and the results are discussed in Section 5.

2 Related work

Basically, there are two types of FPA. The increasing amount of face images/videos shared on social media, makes it easier for malicious users to access face images and use them to spoof face authentication systems. Such attacks are called *impersonation attacks*. In *obfuscation attacks*, the attacker uses extreme makeup, plastic surgery, or hiding a specific part of the face not to be recognized by the detection system. Since obfuscation attacks are more troublesome and costly than impersonation attacks, they are used less frequently. For this reason, studies in the literature mainly focused on impersonation attacks. FPA categories and the attack types are shown in Fig. 2.

Widely used FPA's can be grouped under photo attacks and video replay attacks. A photo attack is defined as the presentation of a genuine user's picture to the face authentication system. Attackers usually use several strategies. A printed photo attack is carried out by presenting an image printed on a piece of paper (A3/A4 paper, copper paper, or professional photo paper) to the face recognition system (Fig. 3a). A warped photo attack presents printed photos to the system by skewing them along the vertical and/or horizontal axis to add depth information (Fig. 3b). In cut photo attacks, the mouth, eyes, and/or nose areas in the photo are cut off, and the system is tried to be spoofed by liveness clues (Fig. 3c). In photo attacks, the

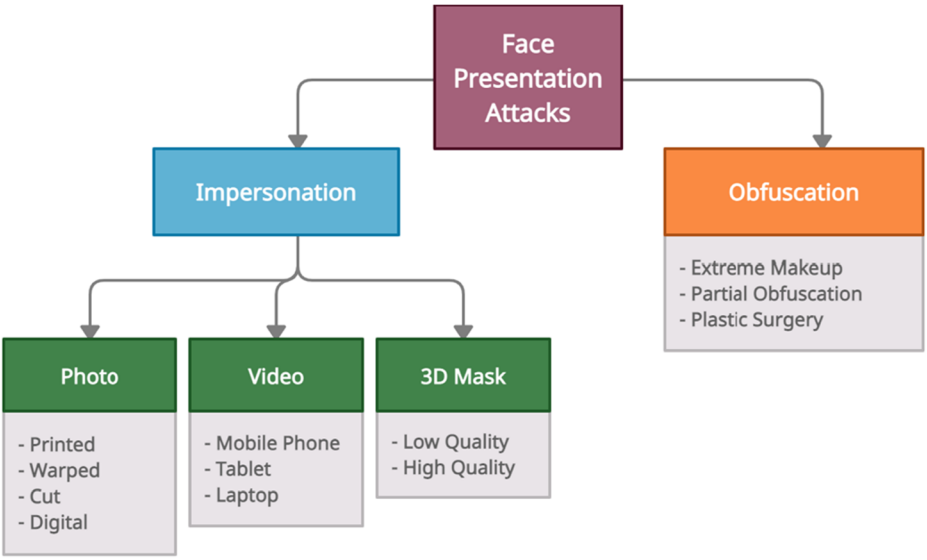


Fig. 2 Face presentation attacks

picture can also be presented to the system by displaying it on the screen of a digital device (Fig. 3d).

Video replay attacks are performed by presenting the person's video to the face authentication system by playing it on smartphones, tablets, or laptops. Compared to static photo attacks, video replay attacks are more complex as they present dynamic information such as eye blinking, mouth movements, and changes in facial expressions [29]. The most significant disadvantage of photo and video replay attacks except warped photo attacks is that they are two-dimensional. On the contrary, 3D masks attacks aim to spoof the system by using low-quality 3D masks made from printed photographs or high-quality 3D masks made of silicon. The high quality of *face-like* 3D structure and its ability to imitate human skin make these attack types more difficult to detect with traditional FPA detection methods [38]. Producing high-quality 3D mask is quite expensive and complex. In addition, this method usually

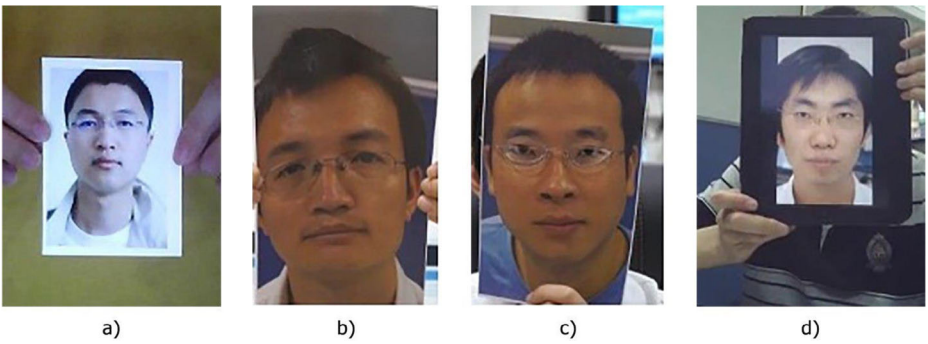


Fig. 3 Types of photo attacks: a) print photo attack, b) warped photo attack, c) photo attack with eye regions cut off, d) digital photo attack

requires user cooperation [17]. For this reason, 3D mask attacks are performed much less frequently than photo or video replay attacks. However, with the proliferation of 3D acquisition sensors, 3D mask attacks are expected to become more common in coming years.

Current FPA detection systems can be divided into four groups: 1) motion analysis, 2) liveness detection, 3) image quality analysis, and 4) texture analysis-based methods.

Motion analysis-based methods depend on optical flow calculated from video sequences. These methods are difficult to emulate and require low user collaboration. However, the need for video sequences with high motion efficiency and high computational complexity are the main disadvantages of these approaches. Anjos et al. proposed a method based on foreground/background motion correlation using optical flow: *i)* The direction of movement for each pixel was obtained by using the horizontal and vertical directions. *ii)* Normalized histograms for the face and background regions were generated, and the distances between the angle histograms of the face and background regions were calculated. *iii)* The average of these values in *N*-frames was used to determine FPA attempts [4].

Liveness detection-based methods try to detect physiological signs of life in videos such as eye blinks, facial expressions. However, these methods require high user collaboration, different devices, and video sequences. They are also time-consuming and computationally complex. Alotaibi and Mahmood proposed a face liveness detection method that uses nonlinear diffusion to obtain depth information and preserve boundary locations. Then, the convolutional neural network is used to extract the distinctive and high-level features of images [2].

Since the image quality characteristics of real accesses and illegal attacks are different, the methods based on image quality analysis use attributes such as color diversity, blur, edge information and chromatic moment. These methods are easy to implement, have low computational costs, and do not require user collaboration. However, their performance mostly depends on the quality of images. Weng et al. proposed FPA detection method based on Image Distortion Analysis (IDA). In the method, four different IDA features (mirror reflection, blur, color moment, and color diversity) were obtained from face images, and the real/fake decision was made by using SVM classifier [40]. Galbally et al. used 25 image quality features (mean frame error, peak signal to noise ratio, maximum difference, average difference, etc.) to distinguish between real and fake faces. Images were classified as real or fake by linear and second-order discriminant analysis [19].

Texture analysis-based methods use the differences between texture patterns (print errors, image blur, etc.) of real and fake faces to identify FPA interferences. These approaches are easy to implement and do not require user collaboration. However, they need suitable feature vectors to distinguish between real and fake faces. Also, low-quality images or videos that produce low texture information may reduce the performance. Tan et al. used the Lambertian reflection model to distinguish between real and fake faces [36]. Määttä et al. used LBP features to analyze facial texture in FPA detection. The features extracted with multiscale LBP operators were classified with SVM to capture the differences between real and fake faces [27]. In another study, the FPA detection performances of both texture-based (LBP, Gabor) and gradient-based (Histogram of Gradients-HoG) face descriptors were examined [28]. Agarwal et al. applied discrete wavelet transform to image sequences, extracted block based Haralick texture features (correlation, contrast, entropy, difference variance, total mean, etc.) and FPA detection is performed using SVM classifier [1]. Zhao et al. proposed a new texture descriptor (Volume Local Binary Count- VLBC) to represent dynamic features. In the method, for a center pixel in any *t* frame, *P* neighboring pixels equally positioned at radius distance *R* in *t-1*,

t, and *t* + *1* frames are used together to extract features [46]. Boulkenafet et al. extracted SURF (Speeded-Up Robust Features) features from different color spaces (HSV, YCbCr) and used Fisher vector coding to embed the feature vectors in a high-dimensional space more suitable for linear classification [9]. In another study, Boulkenafet et al. proposed a color texture analysis based FPA prevention technique. To calculate texture features from luminance and chromaticity channels of multiple color spaces (HSV and YCbCr), multiple texture descriptors (LBP, LPQ, BSIF, and SID) were utilized [8]. Arashloo and Kittler proposed an anomaly based FPA detection approach. In this approach, training data comes only from positive classes, while test data comes from both positive and negative classes. Dynamic features were extracted from video sequences using different texture descriptors (LBP-TOP, LPQ-TOP) [5].

Sthevanie and Ramadhani used LBP and GLCM (Gray-Level Co-Occurrence Matrix) features together in FPA detection. Four different test scenarios were used in the study. The best results were obtained by applying LBP and GLCM matrices to the eye and nose regions [35]. Khurshid et al. developed a system that detects real-time FPA over textural features. First RGB images were converted to gray level and YCbCr color space. Then LBP features were generated from these images, and CoALBP (Co-occurrence of the Adjacent Local Binary Patterns) features were generated from only the gray level image. Finally, the feature set was classified by SVM [23].

Zhang and Xiang used a combination of DWT (Discrete Wavelet Transform), LBP, and DCT (Discrete Cosine Transform) features to evaluate whether a video is real or not. DWT-LBP attributes were obtained from LBP histograms of DWT blocks in each frame. DWT-LBP-DCT features were produced by performing vertical DCT operation on DWT-LBP features. These features were used to train a SVM classifier for FPA detection [43]. Shu et al. proposed FPA detection method based on the chromatic ED-LBP texture feature. In the study, neighbor pixel mismatches in a face image were considered and coded with LBP. The feature histograms in different color channels were calculated separately on each image band. Then, with the help of chromatic ED-LBP histograms and a two-level spatial pyramid, the local structure information of face region in the input image was extracted. Finally, ED-LBP histograms from different color spaces were classified using SVM [33].

3 Material and method

In this section, the datasets, detection and normalization of facial regions, feature extraction, dimensionality reduction and classification techniques are explained in detail.

3.1 Datasets

3.1.1 NUAA

The NUAA [36] is the first dataset for print photo attacks and consists of images obtained from 15 subjects using a webcam. During the taking of these images, the subjects were asked not to blink and to pose from the front with neutral facial expressions. The attacks were carried out using photographs. The dataset is divided into two separate subsets for training and testing. Example images from NUAA dataset are given in Fig. 4.

Fig. 4 Attack examples of the NUAA dataset

3.1.2 CASIA-FASD

CASIA-FASD [44] is a FPA detection dataset that includes printed photo and video replay attacks. CASIA-FASD consists of three types of attacks: *i)* Warped photo attack (imitates paper mask attacks), *ii)* Printed photo attack with cropped eye areas, and *iii)* Video replay attacks (includes signs of liveness such as blinking, mouth, and head movements).

The data was collected for three different imaging quality for all attack types: low, normal, and high. High-quality videos have a resolution of 1280 × 720 pixels, and normal/low-quality videos have a resolution of 640 × 480 pixels. The dataset is divided into two subgroups as training and testing. The training and test sets include real and fake images taken from 20 and 30 subjects, respectively. Sample images from CASIA-FASD dataset are given in Fig. 5.

3.1.3 REPLAY-ATTACK

The REPLAY-ATTACK dataset [12] consists of real and fake access videos of 50 subjects. The videos were taken with the MacBook Air 13″ built-in camera in two distinct lighting conditions: *i)* Controlled: images that use fluorescent lamps for illumination and have a uniform background; *ii)* Uncontrolled: images that use daylight for illumination and have a non-uniform background. High-resolution photos and videos were obtained under the same conditions with the iPhone 3GS, and Canon PowerShot SX150 IS devices. These recordings

Fig. 5 Attack examples of the CASIA dataset

were used to create three different types of attacks: *i)* Print Photo Attack (showing the high-resolution photos printed on A4 paper to the camera), *ii)* Mobile Attack (showing the high-resolution photos and videos to the camera using the iPhone 3GS screen), *iii)* High-Resolution Attack (showing the high-resolution photos and videos to the camera using the iPad screen).

Attack types are divided into two subgroups according to the presentation method of the images/videos to the camera: *i) Hand:* Attacks carried out by holding the input device, *ii) Fixed:* Attacks performed by positioning the input device on a fixed support. The dataset divided into three separate subgroups for training, development, and testing. Sample images from REPLAY-ATTACK dataset are given in Fig. 6.

3.1.4 OULU-NPU

The OULU-NPU face presentation attack detection database [10] of 5940 real access and attack videos of 55 subjects. The videos were recorded using the front cameras of six mobile devices (Samsung Galaxy S6 edge, HTC Desire EYE, MEIZU X5, ASUS Zenfone Selfie, Sony XPERIA C5 Ultra Dual and OPPO N3) in three sessions with different illumination conditions and background scenes. In the database print and video-replay presentation attack types are considered. The attacks were created using two printers and two display devices. The videos were divided into three subject-disjoint subsets for training (20 subjects), development (15 subjects) and testing (20 subjects). There are 4 test protocols for the evaluation of the generalization capability of FPA detection methods. The first protocol evaluates the generalization of the methods under previously unseen illumination conditions and background scene. The second protocol evaluates the effect of input sensors (printers, displays) on the performance of the method. To study the effect of the input camera variation, the third protocol is used. And finally in the last protocol, generalization of the methods is evaluated across previously unseen environmental conditions, attacks, and input sensors. Sample images from OULU-NPU dataset are given in Fig. 7.

3.2 Detection and normalization of face and facial regions

For developing a fully automatic FPA detection system, the facial region should be detected first. The alignment and normalization of input images are essential for improving the classification accuracy. In this study, the general-purpose Dlib library [25] was used to align the input images and identify the facial regions. Dlib is a free library that provides effective solutions for aligning input images (Fig. 8a) with the help of pre-trained 5-point face mask as shown in Fig. 8b. By calculating the angles between these control points, the image was rotated, the head tilts were removed, and frontal faces were obtained (Fig. 8c). Finally, a pre-trained 68-point face mask was applied to frontal faces (Fig. 8d). In the study, these points

Fig. 6 Attack examples of the REPLAY-ATTACK dataset

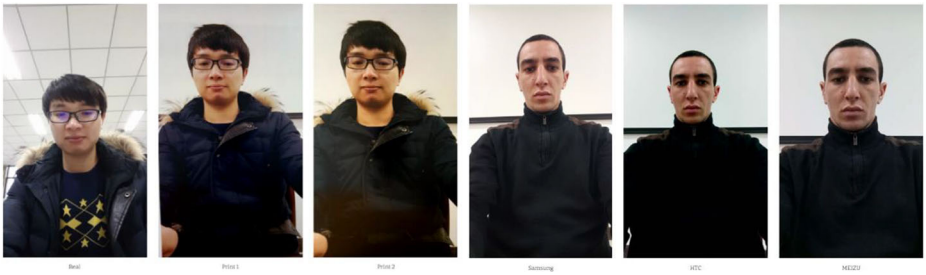


Fig. 7 Attack examples of the OULU-NPU dataset

used to detect the facial regions: wide face, cropped face, eye, nose, and mouth. The regions of which FPA detection performances were examined are shown in Fig. 9.

3.3 Feature extraction

3.3.1 Local binary patterns

LBP is a powerful method for describing texture information in digital images [31]. The LBP texture analysis operator is a gray-level independent texture extraction method. Its main goal is to label the pixel in the center of the 3 × 3 mask by thresholding the neighboring pixel values according to it. LBP codes are generated using eq. (1).

$$LBP_{P,R}(x_c) = \sum_{p=0}^{P-1} u(x_p - x_c) 2^p$$
$$u(y) = \begin{cases} 1, & if \quad y \geq 0 \\ 0, & if \quad y < 0 \end{cases} \tag{1}$$

In the equation x_c is the center pixel, x_p represents the neighbors of x_c, R is the distance of the neighbors from the center pixel, and P represents the number of neighbors.

Generally, the uniform LBP patterns (LBP^{U2}) are used in FPA detection systems [11, 27, 38, 42, 43]. Uniform patterns describe those with at most two bitwise 0/1 transitions between adjacent bits. For example, while the code "00111100" is uniform, the code "10110101" is not. According to this method, each uniform pattern represents a bin in the histogram, while all non-uniform patterns are collected in one bin. The number of uniform LBP patterns created this way is *2 + P(P-1)* [45]. For the input image $I(x,y)$, the $LBP_{8,1}^{U2}$ histogram for $P = 8$ neighbors at a distance of $R = 1$ pixel, is generated by eq. (2) below.

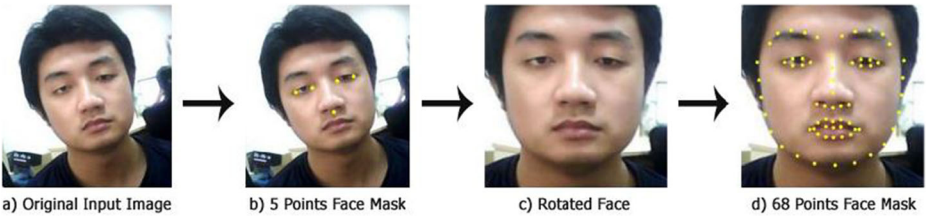


Fig. 8 Aligning the input images

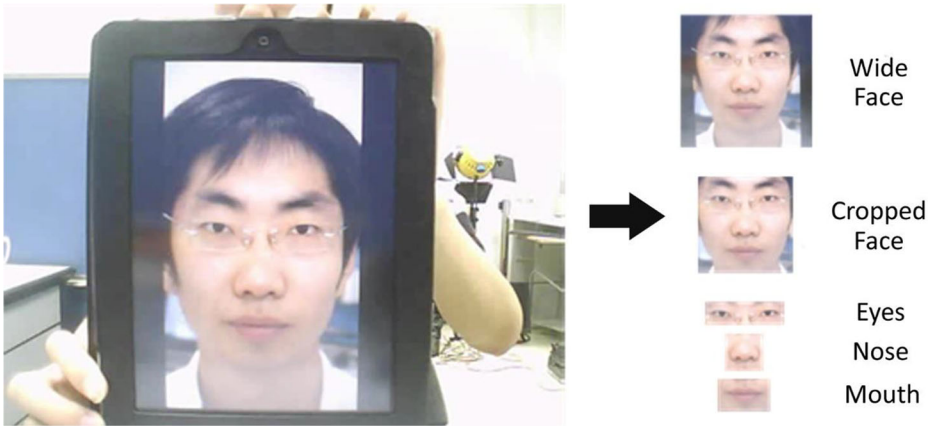


Fig. 9 Facial regions

$$H_i = \sum_{x_c \in I(x,y)} f\{LBP_{8,1}(x_c) = U(i)\}$$
$$i = 0, 1, \ldots, n-1 \qquad f(y) = \begin{cases} 1, & if\ y\ is\ true \\ 0, & if\ y\ is\ false \end{cases} \tag{2}$$

$U(i)$ is the array holding 58 uniform patterns produced in (8,1) neighborhood. This histogram carries information about micro-patterns such as edges, spots, flat areas on the whole image [20]. In the first stage of this study, uniform patterns were used for FPA detection. In the second stage, all patterns were used as the input set, and the FPA detection performance was examined.

3.3.2 Multi-block local binary pattern

The Multi-Block LBP (MB-LBP) is an improved model of the original LBP algorithm proposed for detailed examination of edges, spots, and flat areas on the image. In this model, the image is firstly divided into n blocks. Then, the original LBP operator is applied to all blocks, and regional histograms are obtained. Finally, the histograms of n regions are concatenated, and a feature vector is produced for the input image [6]. Figure 10 shows the feature extraction process with MB-LBP algorithm on the sub-blocks of a given input image.

3.4 Dimensionality reduction

It is essential to minimize costs such as computational complexity, computation time, and storage in a classification process. Reducing the size of the feature vector

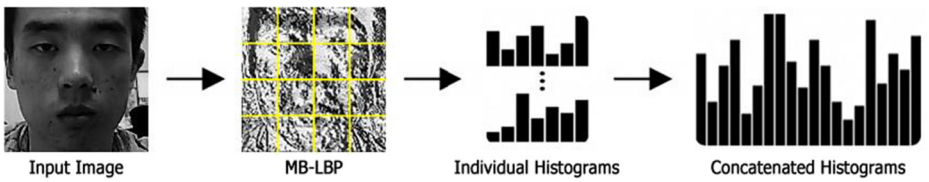


Fig. 10 Generating the MB-LBP feature vector

is one of the essential phases of this process. The dimensionality reduction aims to determine the subsets that will best represent the dataset and achieve the best classification accuracy.

3.4.1 Principal components analysis

PCA is a method of projecting the data in a multidimensional space to a lower-dimensional space in a way that maximizes the variance [3]. The main goal is to find linear combinations of variables, called principal components, which best represent the dataset. These principal components correspond to eigenvectors that maximize the variance of the data projected onto them [13]. The eigenvectors of the data covariance matrix (S) are obtained by the equation $W_{opt} = argmax_{\|W\| = 1} W^T S W$. When the equation is solved, the eigenvectors (W) of S corresponding to the largest d ($d \leq D$) eigenvalues are obtained. Then, dimensionality reduction is performed using $y_i = W^T x_i$ ($y_i \epsilon R^d$) [20]. Principal components with 95% eigenvalues were used in the study.

3.5 Classification

3.5.1 Support vector machines

SVM is a supervised classification algorithm based on statistical learning theory. The mathematical algorithms of SVM were initially designed for the linear classification of two-class data and then generalized for the classification of multi-class and non-linear data. SVM is based on finding the hyperplane that can best separate two classes from each other [22]. SVM model separating two classes is given in Fig. 11.

In Fig. 11a, H1 plane cannot separate the classes correctly. H2 plane successfully separated the classes, but the distance between the samples and the hyperplane are minimal. H3 plane, on the other hand, separated the class samples with maximum distance. The plane passing between the closest instances of two classes in Fig. 11b is the optimum solution for separating these classes. The samples that intersect the imaginary points equidistant from the plane are

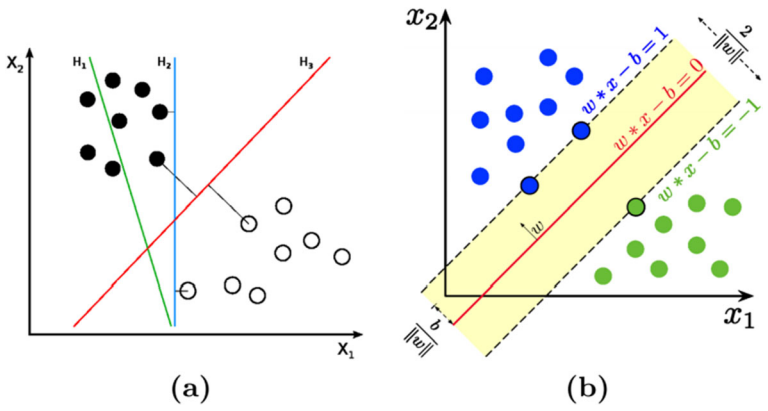


Fig. 11 SVM hyperplane detection

called support vectors. In the study, SVM with Radial Basis Function (RBF) kernel is used to classify the input image as real/fake.

4 Experimental results

In NUAA and CASIA datasets, all frames in which the facial regions detected were used in the experiments. Due to the large number of samples and/or test scenarios in REPLAY-ATTACK and OULU-NPU datasets, the images are collected by taking frames at 125 ms intervals from the videos to reduce processing time and complexity. The frames in which the facial regions detected correctly were used to evaluate the performance of the proposed method.

NUAA and CASIA datasets consist of only training and test sets. REPLAY-ATTACK and OULU-NPU dataset has a development set except the training and test sets. For this reason, 5-fold cross-validation was performed on the training sets of NUAA and CASIA datasets to produce the development sets. The training set was divided into five equal parts, four parts were used as the training set and one part as the development set. This process was repeated five times, and the average of the results was calculated.

The NUAA dataset contains 3459 training and 9067 test examples for a single attack type (printed photo). There are seven test scenarios in the CASIA dataset; three scenarios for different image qualities (low-quality, normal-quality, high-quality), three different scenarios for attack types (warped photo, cut photo, video replay), and a general test scenario where all data are used together. In the study, FPA detection results were obtained for all the seven test scenarios. On the other hand, REPLAY-ATTACK dataset includes six different attack types: high-definition attack, mobile attack, printed photo attack, digital+ printed photo attack, video replay attack and all attacks. The attacks were carried out both by holding the device in hand and positioning the device in a fixed place. Considering three scenarios according to the positioning type (hand, fixed, all), $6 \times 3 = 18$ different test scenarios were evaluated for REPLAY-ATTACK dataset. Finally, experiments were carried out for the 4 test protocols in the OULU-NPU dataset. The test scenarios and the numbers of real and fake images used in the experiments for these datasets are given in Table 1.

FPA detection systems are subject to two types of errors. These errors are denial of real accesses (false reject) and acceptance of attacks (false accept). The performance of these systems is usually measured by Half Total Error Rate (HTER) metric. HTER is half of the sum of False Acceptance Rate (FAR) and False Rejection Rate (FRR) and is calculated by eq. (3) below.

$$HTER(\tau) = \frac{FAR(\tau) + FRR(\tau)}{2} \tag{3}$$

Since FAR and FRR depend on the threshold value τ, increasing FAR causes FRR to decrease. Therefore, results are often represented by graphs showing the variation of FAR concerning FRR for different τ threshold values. Equal Error Rate (EER), another criterion used in FPA detection, is the value at the point where FAR and FRR are equal as shown in Fig. 12. The threshold τ corresponding to the EER is obtained from the development set and HTER is calculated from the test set using this threshold value.

Table 1 The number of samples used in the experiments (H_: attacks carried out by holding the input device in hand, F_: attacks carried out by positioning the input device on a fixed support, A_: the combination of these two sets)

Dataset	# of Train Sample		# of Development Sample		# of Test Sample		Attack Type
	Real	Fake	Real	Fake	Real	Fake	
NUAA	1743	1746	–		3356	5711	Printed Photo
CASIA	3140	11,015	–		5297	16,088	Low Quality
	3197	11,231			4830	16,085	Normal Quality
	4577	11,846			5782	17,291	High Quality
	10,914	12,859			15,909	19,165	Warped Photo
		9499				14,731	Cut Photo
		11,734				15,568	Video
		34,092				49,464	Overall
REPLAY- ATTACK	7185	4761	7197	4633	9596	6106	H_ Highdef
		4757		4651		6173	F_ Highdef
		9518		9284		12,279	A_Highdef
		4703		4706		6084	H_Mobile
		4717		4462		5731	F_Mobile
		9420		9168		11,815	A_Mobile
		7137		7104		9218	H_Printed_Photo
		7170		6927		9142	F_Printed_Photo
		14,307		14,031		18,360	A_Printed_Photo
		2395		2396		3114	H_Digital+Printed Photo
		2370		2342		3122	F_Digital+Printed Photo
		4765		4738		6236	A_Digital+Printed Photo
		4722		4631		6086	H_Video
		4674		4528		5884	F_Video
		9396		9159		11,970	A_Video
		11,859		11,735		15,304	H_Grandtest
		11,844		11,455		15,026	F_Grandtest
		23,703		23,190		30,330	A_Grandtest
OULU-NPU	3812	15,284	2743	11,067	1353	7833	Protocol I
	5782	11,799	4150	8445	5371	10,538	Protocol II
	4840	19,618	3590	14,618	928	3768	Protocol III
	3256	6622	2444	4916	255	600	Protocol IV

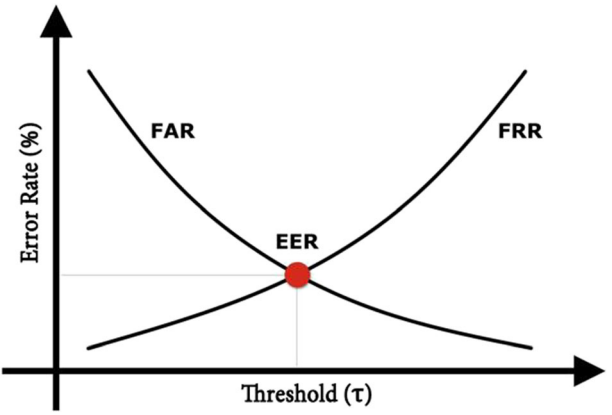


Fig. 12 Obtaining EER value from FAR and FRR graph

In the study, Area Under Curve (AUC) metric was also used to evaluate the system performance. AUC represents the area under the ROC probability curve. The ROC curve shows the false positive rate (FPR) change versus the true positive rate (TPR). These values are obtained by the following eqs. (4).

$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{TN + FP} \tag{4}$$

In the equation, TP, FP, TN, and FN represents the true positives, false positives, true negatives, and false negatives, respectively. The AUC criterion expresses how well the model can distinguish between classes. So, the higher the AUC, the better the model predicts.

In the study, regional LBP features were used for FPA detection from facial regions. Input images (wide face, cropped face, eye, nose, and mouth) were separated into various sub-regions, and regional features were extracted. In a previous study, it was seen that the features extracted using 8 × 8 regional decomposition process, did not increase the FPA detection performance but caused the vector size to increase significantly [21]. For this reason, feature vectors obtained from 1 × 1, 2 × 2, and 4 × 4 subregions of the input images were used in the study. After the images were divided into subregions, the $LBP_{8,1}^{U2}$ operator was applied. The histograms obtained from each subregion were concatenated, and the feature vector was obtained for the input image. Then these feature vectors are classified as real/fake using SVM. After that, the size of the feature vectors was reduced by PCA, and reclassification was performed.

The best HTER results obtained in FPA detection experiments on NUAA, CASIA, REPLAY-ATTACK and OULU-NPU datasets are given in Table 2. Due to the different test scenarios and the face parts used, the number of results obtained is quite large. For example, only REPLAY-ATTACK dataset contains $5 \times 3 \times 18 \times 2 = 540$ classification results for 5 different face regions (wide face, cropped face, eye, mouth, nose), 3 different regional parsing (1 × 1, 2 × 2, 4 × 4), 18 different test scenarios and 2 different feature sets ($LBP_{8,1}^{U2}$, $LBP_{8,1}^{U2}$ + PCA). Therefore, only the best results are shared in Table 2.

When the results in the table are examined, the performance of FPA detection is evaluated for five different face regions and 30 different test scenarios on 4 datasets. According to the results, in 24 of the 30 test scenarios, FPA detection of the wide face region is higher than the other regions (cropped face, eyes, nose, and mouth). In the remaining 6 test scenarios, the cropped face area is more successful. On the other hand, reducing the size of MB-$LBP_{8,1}^{U2}$ features with PCA ($LBP_{8,1}^{U2}$ + PCA) improves the FPA detection performance in 20 test scenarios.

For the only attack type in NUAA dataset, the MB-$LBP_{8,1}^{U2}$ features obtained from the cropped face region have 0.0353 HTER performance in FPA detection. In CASIA dataset, the highest performance of 0.0466 HTER was obtained with the MB-$LBP_{8,1}^{U2}$ features obtained from wide face region for video replay attack. The MB-$LBP_{8,1}^{U2}$ features obtained from wide face region are better in FPA detection in 4 of the 7 test scenarios defined in the dataset (low quality, normal quality, printed photo, video, and replay attacks). On the other hand, MB-$LBP_{8,1}^{U2}$ and MB-$LBP_{8,1}^{U2}$ + PCA features obtained from the cropped face region are more successful in high quality and cut photo attacks, respectively. When all attack types in this dataset are evaluated together, MB-$LBP_{8,1}^{U2}$ + PCA features produced from the cropped face region have the best performance with 0.1275 HTER. MB-$LBP_{8,1}^{U2}$ + PCA features have

Table 2 Best HTER values obtained for face regions and test cases in NUAA, CASIA and REPLAY-ATTACK datasets

Dataset	Attack Type	Face Regions									
		Wide Face		Cropped Face		Eye		Nose		Mouth	
		MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)
NUAA	Printed Photo	0.0793 (2×2)	0.1265 (1×1)	**0.0353 (2×2)**	0.1754 (1×1)	0.1978 (4×4)	0.1399 (1×1)	0.0952 (4×4)	0.2377 (4×4)	0.1135 (2×2)	0.1254 (2×2)
CASIA	Low Quality	**0.0700 (2×2)**	0.1663 (2×2)	0.1084 (2×2)	0.1281 (1×1)	0.2188 (4×4)	0.2326 (4×4)	0.1069 (2×2)	0.1020 (4×4)	0.1195 (4×4)	0.1281 (4×4)
	Normal Quality	**0.0904 (2×2)**	0.1991 (4×4)	0.1149 (2×2)	0.1441 (4×4)	0.2420 (2×2)	0.2233 (2×2)	0.1601 (2×2)	0.1574 (2×2)	0.1481 (4×4)	0.1572 (2×2)
	High Quality	0.2255 (4×4)	0.3366 (1×1)	**0.1872 (4×4)**	0.2924 (2×2)	0.3854 (2×2)	0.3767 (2×2)	0.3181 (1×1)	0.3030 (2×2)	0.2842 (4×4)	0.2783 (2×2)
	Warped Photo	**0.1116 (4×4)**	0.1496 (2×2)	0.1659 (4×4)	0.1341 (4×4)	0.3240 (4×4)	0.3128 (4×4)	0.2597 (2×2)	0.2365 (2×2)	0.2335 (2×2)	0.2440 (2×2)
	Cut Photo	0.1363 (4×4)	0.1238 (2×2)	0.1524 (4×4)	**0.1017 (4×4)**	0.1544 (4×4)	0.1592 (4×4)	0.2229 (1×1)	0.2138 (2×2)	0.1873 (2×2)	0.2268 (2×2)
	Video	**0.0466 (2×2)**	0.0784 (2×2)	0.1037 (4×4)	0.0824 (4×4)	0.2860 (1×1)	0.2875 (2×2)	0.2145 (2×2)	0.2135 (2×2)	0.2066 (4×4)	0.2084 (2×2)
	Overall	0.1360 (4×4)	0.1485 (2×2)	0.1589 (4×4)	**0.1275 (4×4)**	0.3164 (4×4)	0.3258 (4×4)	0.2565 (2×2)	0.2351 (2×2)	0.2452 (2×2)	0.2511 (2×2)
REPLAY-ATTACK	H_Highdef	0.0840 (2×2)	**0.0732 (2×2)**	0.1260 (2×2)	0.0972 (4×4)	0.1995 (2×2)	0.1706 (2×2)	0.2716 (1×1)	0.2761 (2×2)	0.2282 (2×2)	0.2346 (2×2)
	F_Highdef	0.1038 (2×2)	**0.0667 (2×2)**	0.1333 (1×1)	0.0854 (4×4)	0.1783 (2×2)	0.1651 (2×2)	0.2333 (2×2)	0.2433 (2×2)	0.2080 (2×2)	0.2010 (2×2)
	A_Highdef	0.0909 (2×2)	**0.0635 (2×2)**	0.1164 (2×2)	0.0963 (4×4)	0.1853 (2×2)	0.1718 (2×2)	0.2413 (1×1)	0.2632 (2×2)	0.2144 (2×2)	0.2302 (2×2)
	H_Mobile	0.0463 (2×2)	**0.0237 (4×4)**	0.0282 (2×2)	0.0351 (4×4)	0.0934 (2×2)	0.1329 (2×2)	0.1116 (2×2)	0.1164 (2×2)	0.1063 (2×2)	0.0991 (1×1)
	F_Mobile	**0.0076 (4×4)**	0.0094 (4×4)	0.0132 (2×2)	0.0168 (4×4)	0.0926 (2×2)	0.1331 (1×1)	0.0676 (2×2)	0.0681 (2×2)	0.1324 (2×2)	0.1209 (2×2)
	A_Mobile	0.0436	**0.0244**	0.0426	0.0384	0.1089	0.1473	0.1002	0.1041	0.1190	0.1198

Table 2 (continued)

Dataset	Attack Type	Face Regions									
		Wide Face		Cropped Face		Eye		Nose		Mouth	
		MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)
		(2×2)	**(4×4)**	(2×2)	(4×4)	(2×2)	(2×2)	(2×2)	(2×2)	(2×2)	(2×2)
	H_Printed Photo	**0.0111**	0.0306	0.0451	0.0507	0.1401	0.1267	0.0721	0.0720	0.1749	0.1742
		(4×4)	(4×4)	(4×4)	(1×1)	(4×4)	(2×2)	(4×4)	(2×2)	(1×1)	(2×2)
	F_Printed Photo	**0.0236**	0.0257	0.0400	0.0508	0.1677	0.1604	0.0749	0.0661	0.1938	0.1766
		(4×4)	(2×2)	(4×4)	(1×1)	(1×1)	(2×2)	(1×1)	(2×2)	(1×1)	(4×4)
	A_Printed Photo	**0.0239**	0.0287	0.0411	0.0445	0.1518	0.1354	0.0756	0.0705	0.1878	0.1824
		(4×4)	(4×4)	(4×4)	(4×4)	(4×4)	(2×2)	(2×2)	(2×2)	(1×1)	(4×4)
	H_Digital+Printed Photo	0.0952	**0.0838**	0.1247	0.1073	0.2225	0.2199	0.2276	0.2326	0.2371	0.2390
		(1×1)	**(2×2)**	(2×2)	(2×2)	(2×2)	(2×2)	(4×4)	(2×2)	(2×2)	(2×2)
	F_Digital+Printed Photo	0.0983	**0.0514**	0.0852	0.0645	0.2406	0.2332	0.2053	0.1941	0.2306	0.2184
		(2×2)	**(2×2)**	(4×4)	(4×4)	(2×2)	(2×2)	(4×4)	(2×2)	(2×2)	(2×2)
	A_Digital+Printed Photo	0.1111	**0.0789**	0.1167	0.0890	0.2295	0.2366	0.2167	0.2206	0.2423	0.2382
		(2×2)	**(2×2)**	(4×4)	(4×4)	(2×2)	(2×2)	(4×4)	(1×1)	(2×2)	(2×2)
	H_Video	0.0681	0.0593	0.0796	**0.0521**	0.1284	0.1457	0.1835	0.1928	0.1689	0.1546
		(2×2)	(4×4)	(2×2)	**(4×4)**	(2×2)	(2×2)	(2×2)	(2×2)	(1×1)	(1×1)
	F_Video	0.0673	**0.0333**	0.0837	0.0401	0.1237	0.1453	0.1643	0.1721	0.1577	0.1641
		(4×4)	**(2×2)**	(4×4)	(4×4)	(2×2)	(1×1)	(2×2)	(2×2)	(2×2)	(2×2)
	A_Video	0.0738	0.0521	0.0771	**0.0490**	0.1262	0.1538	0.1721	0.1798	0.1627	0.1545
		(2×2)	(4×4)	(2×2)	**(4×4)**	(2×2)	(1×1)	(2×2)	(2×2)	(1×1)	(1×1)
	H_Grandtest	0.1006	**0.0761**	0.1049	0.0896	0.2081	0.1948	0.2336	0.2361	0.2373	0.2260
		(1×1)	**(2×2)**	(2×2)	(2×2)	(2×2)	(2×2)	(4×4)	(2×2)	(2×2)	(1×1)
	F_Grandtest	0.0937	**0.0414**	0.1001	0.0604	0.2175	0.2149	0.2013	0.1906	0.2129	0.2000
		(2×2)	**(2×2)**	(4×4)	(4×4)	(2×2)	(2×2)	(2×2)	(2×2)	(2×2)	(2×2)
	A_Grandtest	0.1012	**0.0709**	0.1029	0.0847	0.2119	0.2088	0.2244	0.2149	0.2248	0.2332
		(2×2)	**(2×2)**	(2×2)	(2×2)	(2×2)	(2×2)	(2×2)	(2×2)	(2×2)	(2×2)
OULU-NPU	Protocol I	0.1770 (4×4)	**0.1612** **(4×4)**	0.2475 (1×1)	0.2468 (4×4)	0.3000 (2×2)	0.3444 (2×2)	0.2204 (2×2)	0.2696 (2×2)	0.3012 (4×4)	0.3253 (4×4)

Table 2 (continued)

Dataset	Attack Type	Face Regions									
		Wide Face		Cropped Face		Eye		Nose		Mouth	
		MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)	MB-$LBP_{8,1}^{U2}$ (Blocks)	MB-$LBP_{8,1}^{U2}$+ PCA (Blocks)
	Protocol II	0.1364 (4×4)	**0.0842 (4×4)**	0.1515 (4×4)	0.1435 (4×4)	0.2973 (4×4)	0.3098 (4×4)	0.2104 (4×4)	0.1933 (4×4)	0.2593 (2×2)	0.2333 (2×2)
	Protocol III	0.2181 (4×4)	**0.1652 (4×4)**	0.2206 (4×4)	0.2146 (4×4)	0.3124 (4×4)	0.3207 (4×4)	0.2534 (4×4)	0.2281 (4×4)	0.2877 (2×2)	0.2624 (2×2)
	Protocol IV	0.2664 (4×4)	**0.2462 (4×4)**	0.3446 (1×1)	0.3493 (4×4)	0.3525 (4×4)	0.3659 (4×4)	0.2666 (1×1)	0.2983 (1×1)	0.3208 (1×1)	0.3137 (1×1)

The best result for the relevant attack scenario is shown in bold

better results for cut photo attack and the test scenario where all attacks are evaluated together in CASIA dataset.

In REPLAY-ATTACK dataset, which has more samples, attack types, and test scenarios, the highest performance was obtained as 0.0076 HTER with the MB-$LBP_{8,1}^{U2}$ features generated from wide face region in mobile attack performed over a fixed source (F_Mobile). In all mobile attacks (fixed + hand), 0.0244 HTER success was achieved with MB-$LBP_{8,1}^{U2}$ + PCA features. MB-$LBP_{8,1}^{U2}$ + PCA features generated from the cropped face region increased the performance of FPA detection in only 2 of the 18 test scenarios in this dataset (H_Video and A_Video). While MB-$LBP_{8,1}^{U2}$ features produced from the wide face region had better performance in 4 of the remaining 16 test scenarios, MB-$LBP_{8,1}^{U2}$ + PCA features produced from the wide face region in 12 test scenarios perform better FPA detection. When all attacks in the REPLAY-ATTACK dataset are evaluated together, 0.0709 HTER is obtained. This value decreases to 0.0414 HTER only for the fixed source attacks.

According to the results on OULU-NPU dataset, the highest performance of 0.0842 HTER was obtained with the MB-$LBP_{8,1}^{U2}$ + PCA features obtained from wide face region for Protocol II which evaluates the effect of input sensors (printers, displays) on the performance of the method. According to the results, the proposed method is robust to the different input sensors used to create the attacks. In protocol 4, where all factors (different lighting, background scene, input sensor and camera variation) are evaluated together, MB-$LBP_{8,1}^{U2}$ + PCA features produced from wide face region gives the best FPA detection performance of 0.2462 HTER. The MB-$LBP_{8,1}^{U2}$ + PCA features obtained from wide face region are better in FPA detection in all the test protocols defined in this dataset.

When the FPA detection performances of eye, nose, and mouth regions were examined, the best results are, 0.0952 HTER (nose region) for NUAA dataset, 0.1020 HTER (nose region) for low-quality attack in CASIA dataset, 0.0676 HTER (nose region) for F_Mobile attack in REPLAY-ATTACK dataset and 0.1933 HTER (nose region) for Protocol II in OULU-NPU dataset. In CASIA dataset, the mouth region succeeds in 4 test scenarios, the nose region in 2 test scenarios and the eye region in 1 test scenario (cut photo attack). When all attacks were evaluated together, MB-$LBP_{8,1}^{U2}$ + PCA attributes extracted from the nose region showed 0.2351 HTER performance.

In REPLAY-ATTACK dataset, the eye region is more successful in detecting FPA in 10 of the 18 test scenarios. This creates a prediction about the detection performance of attacks which are carried out only in the eye region due to mask usage in current and future pandemic conditions. In the remaining 8 test scenarios, it is understood that the performance of the nose area is high. For this reason, it is presumed that evaluating the eye and nose regions together will improve the FPA detection performance. The mouth region did not show any superiority over the eye and nose regions on the test scenarios in this dataset. When all attack types are included, FPA detection can be performed with the MB-$LBP_{8,1}^{U2}$ + PCA features obtained from the eye region with a HTER success of 0.2088.

In the OULU-NPU dataset, the nose region shows better FPA detection performance in all of the 4 protocols. In the most challenging test scenario, Protocol 4, MB-$LBP_{8,1}^{U2}$ features produced from nose region have the highest FPA detection performance of 0.2666 HTER.

The regional FPA detection results showed that the performances of nose and eye regions are generally better than that of mouth region. The performance of the nose region is due to the smaller variation compared to the eye and mouth regions when aligning the face. Since the eye region has a more dynamic structure, texture differences in real/fake images in this region are

more evident. The mouth area is generally unsuccessful in FPA detection. This is because the mouth region may not contain attack patterns.

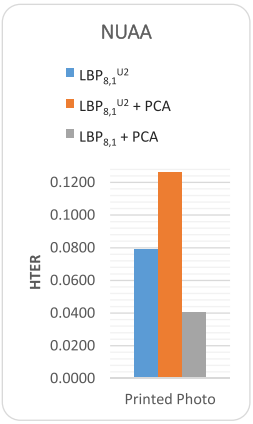
As one can see from the whole table, the information around the face region (wide face) makes positive contributions to the performance of FPA detection system. Therefore, subsequent experiments were performed only on the wide facial region. On the other hand, uniform LBP features are generally used in texture recognition studies. In this study, all the MB-$LBP_{8,1}$ feature extracted from the wide face region were used to examine the effect of all LBP features on the FPA detection performance.

Figure 13 shows the performance of MB-$LBP_{8,1}^{U2}$, MB-$LBP_{8,1}^{U2}$ + PCA, and MB-$LBP_{8,1}$ + PCA features extracted from the wide face region, according to test scenarios in NUAA, CASIA, and REPLAY-ATTACK and OULU-NPU datasets. Using all the LBP patterns for printed photo which is the only attack type in the NUAA dataset, increased the performance by 48.8% (Fig. 13a). It can be seen from the figure that the real/fake detection in the NUAA dataset is performed with 0.0406 HTER. The AUC value for this test scenario is 0.9594, and the classification accuracy is 95.88%.

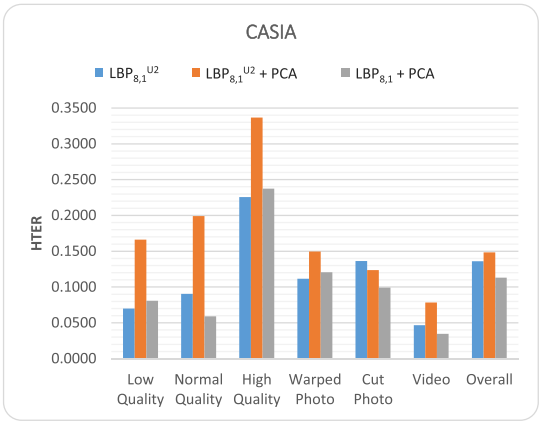
The results produced for the test scenarios of CASIA dataset are given in Fig. 13b. It can be seen from the figure that the performance of the FPA detection system decreases as the picture quality increases. One of the discriminating factors between the real accesses and attacks is the high frequency content of images. In spoofing attacks this content is likely to be attenuated. However, increasing the device quality strengthened the high frequency content of attacks so the ability to distinguish them from real accesses is diminished. As can be also seen from the graph, the use of all LBP patterns increases the performance of FPA detection in 4 of 7 test scenarios (normal quality, cut photo, video replay, and overall). It is essential to improve 16.9% in the test scenario where all attack types are evaluated together. In the experiment performed on all images in the CASIA dataset, real/fake detection can be made with 0.1130 HTER. The obtained AUC value and classification accuracy were 0.8870 and 89.91%, respectively.

The results in REPLAY-ATTACK dataset shows that, the use of all LBP patterns increases the performance of FPA detection, especially in types of attacks where the device is hold in hand (Fig. 13c). These are H_Highdef, H_Digital+Printed Photo, H_Video, and H_Grandtest. Additionally, performance has been increased in the A_Video (video replay attacks from both fixed and hand-held source) attack. FPA detection performance was improved by 10.38% in the H_Grandtest test scenario, which encompasses all sorts of attacks in which the device is hold in hand. As the attacks are generally made from hand-held devices, this improvement is quite significant. The HTER value obtained in this scenario is 0.0682, the AUC value is 0.9318, and the classification accuracy is 92.73%. In addition, in the A_Grandtest test scenario, which was created by evaluating all attack types in the REPLAY-ATTACK dataset together, all LBP patterns provided a 1.5% improvement in the performance of FPA detection. The results are 0.0698 HTER, 0.9285 AUC value, and 92.19% classification accuracy.

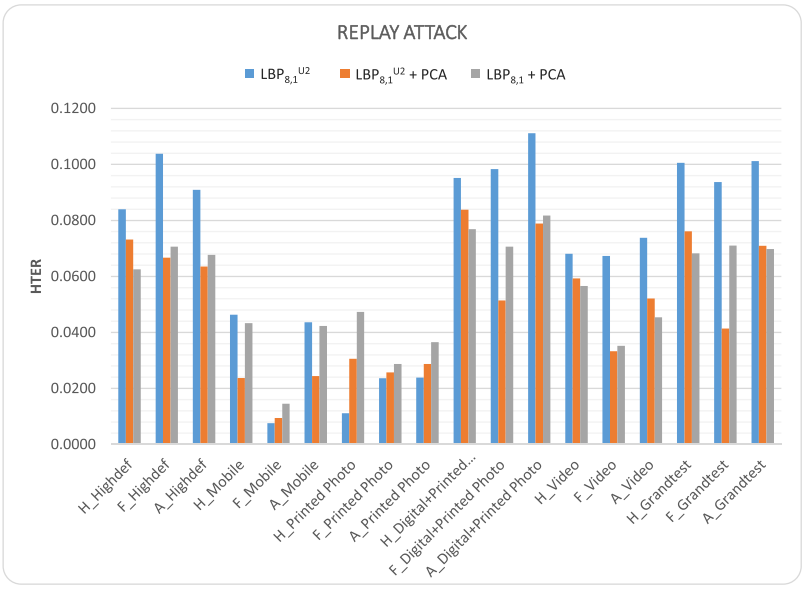
The FPA detection performance of the proposed method on OULU-NPU dataset is shown in Fig. 13d. The use of all LBP patterns increases the PFA detection performance for all test protocols. OULU-NPU has much more complex and challenging test scenarios than other datasets. Despite this, an adequate level of FPA detection performance has been achieved in Protocol I and II with the proposed method. For Protocol I and II the use of all LBP patterns improves the FPA detection performance by %41.3 and %18.2, respectively. The results for Protocol I are 0.0946 HTER, 0.9053 AUC value, and 92.24% classification accuracy. For


NUAA
LBP8,1U2
LBP8,1U2 + PCA
LBP8,1 + PCA
HTER
0.1200
0.1000
0.0800
0.0600
0.0400
0.0200
0.0000
Printed Photo

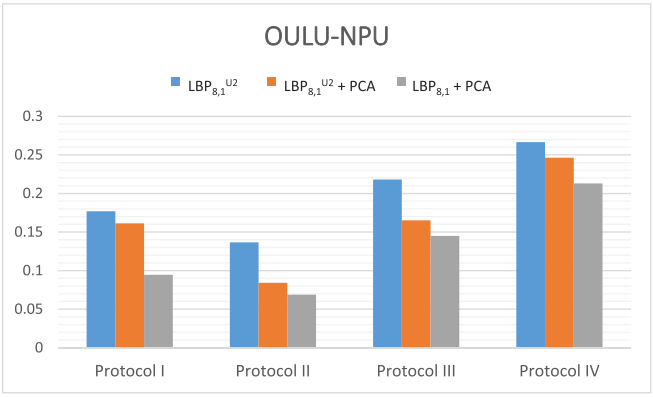
(a)


CASIA
LBP8,1U2
LBP8,1U2 + PCA
LBP8,1 + PCA
HTER
0.3500
0.3000
0.2500
0.2000
0.1500
0.1000
0.0500
0.0000
Low Quality
Normal Quality
High Quality
Warped Photo
Cut Photo
Video
Overall

(b)


REPLAY ATTACK
LBP8,1U2
LBP8,1U2 + PCA
LBP8,1 + PCA
HTER
0.1200
0.1000
0.0800
0.0600
0.0400
0.0200
0.0000
H_Highdef
F_Highdef
A_Highdef
H_Mobile
F_Mobile
A_Mobile
H_Printed Photo
F_Printed Photo
A_Printed Photo
H_Digital+Printed...
F_Digital+Printed Photo
A_Digital+Printed Photo
H_Video
F_Video
A_Video
H_Grandtest
F_Grandtest
A_Grandtest

(c)


OULU-NPU
LBP8,1U2
LBP8,1U2 + PCA
LBP8,1 + PCA
0.3
0.25
0.2
0.15
0.1
0.05
0
Protocol I
Protocol II
Protocol III
Protocol IV

(d)

Fig. 13 FPA detection performances of MB-LBP8,1 U2, MB-LBP8,1 U2 + PCA and MB-LBP8,1 + PCA attributes extracted from the wide face region, according to test scenarios **a**) NUAA **b**) CASIA **c**) REPLAY-ATTACK d) OULU datasets

Protocol II 0.0688 HTER, 0.9312 AUC value, and 93.93% classification accuracy are achieved. As the third and fourth protocol are more complex, the FPA detection performance is slightly lower. However, when evaluated in general, it is an important result in terms of the generality of the proposed method that the use of all LBP patterns in all protocols increases the performance. All these results reveal that the patterns other than the uniform LBP patterns contain information in the FPA detection problem.

The comparison of the proposed method with the studies using LBP features in the literature is given in Table 3. Since the studies in the literature perform FPA detection on the entire face image, the comparison is made according to the results obtained with the wide face region. Therefore, it was not possible to compare the results related to the performance of the facial regions in FPA detection. On the other hand, when the previous studies are examined, generally the results produced on the whole dataset are given. Various test scenarios as in the study were not studied. In this respect, the analysis of FPA detection performance according to 30 different test scenarios and five different face regions in this study will shed light on relevant studies on this subject.

As shown in Table 3 the FPA detection performances on NUAA and CASIA datasets are better than the other studies. In previous studies, EER results were mostly reported on NUAA and CASIA datasets. In this study, HTER results were calculated by generating development sets from the training sets of these datasets using 5-fold cross-validation method. For NUAA dataset 0.17% EER and 4.06% HTER performance were obtained. The results for the CASIA dataset were 0.22% EER and 11.30% HTER. The HTER result obtained from the REPLAY-ATTACK dataset is 6.98%, which is better than some studies in the literature.

Table 3 Comparison of the proposed method with other methods presented in the literature

Method	NUAA		CASIA		REPLAY-ATTACK	
	EER (%)	HTER (%)	EER (%)	HTER (%)	EER (%)	HTER (%)
Jukka Määttä et al., 2011 [27]	2.9	–	–	–	–	–
Chingovska et al., 2012 [12]	–	19.03	–	18.17	–	15.16
J. Määttä et al., 2012 [28]	1.1	–	–	–	–	–
Yang et al., 2013 [41]	1.9	–	11.8	–	–	–
Komulainen et al., 2013 [26]	–	–	–	–	5.11	–
Tirunagari et al., 2015 [38]	–	–	–	21.75	–	3.75
Boulkenafet et al., 2015 [7]	–	–	6.2	–	0.4	2.9
Tian & Xiang, 2016 [37]	–	–	–	18.06	–	0.0
Kim et al., 2017 [24]	–	–	4.89	–	–	5.5
De Souza et al., 2017 [16]	1.8	1.7	–	–	–	–
W. Zhang & Xiang, 2020 [43]	–	–	5.56	–	0.0	0.0
Shu et al., 2021 [33]	–	–	1.11	–	0.0	0.0
Proposed Method	**0.17**	4.06	**0.22**	**11.30**	9.28	6.98

Successful results obtained with the proposed method compared to other studies are shown in bold

5 Discussion

In this study it is aimed i) to examine the FPA detection performance of facial regions (wide face, cropped face, eye, nose, and mouth), ii) to determine how much information LBP features carry about FPA detection (the uniform patterns and all patterns) and iii) to create a simple, interpretable, and effective face spoofing detection system with low computational complexity.

The printed photo (printed, warped, cut, digital) and video replay (mobile, tablet) attacks, which are frequently used in the real-world scenarios are emphasized in the study. Production of high-quality 3D masks is quite expensive and complex. Also, there is no high-quality 3D mask attack videos in the CASIA, REPLAY-ATTACK and OULU-NPU datasets which are the frequently used datasets in the literature. But in the CASIA dataset, warped photo and cut photo attacks can be included in the class of low-quality 3D mask attacks.

In the study, the performances of facial regions (wide face, cropped face, eye, nose, mouth) in FPA detection were investigated. In the first stage of the study, FPA detection was performed with MB-$LBP_{8,1}^{U2}$ patterns obtained from 5 different face regions for 4 datasets and a total of 30 different test scenarios. The results show that the wide face region is more successful in detecting FPA than other facial regions in 24 test scenarios. In the remaining 6 test scenarios, the cropped face region was successful. This indicates that the entire facial region provides essential information for FPA detection. On the other hand, it is understood that the background information contained in the input images positively affects the performance.

When the FPA detection performances of the eye, nose, and mouth regions are evaluated, different regions perform better according to the datasets and test scenarios. In 10 test scenarios on the REPLAY-ATTACK dataset and 1 test scenario on the CASIA dataset, the eye region was more successful than the nose and mouth region. This situation creates a prediction about the detection performance of attacks to be made only from the eye region due to the use of masks in current and future pandemic conditions. On the other hand, the nose region gives the best FPA detection performance for the 15 test scenarios (1 NUAA, 2 CASIA, 8 REPLAY-ATTACK, 4 OULU-NPU) on all datasets. From these results it can be concluded that the hybrid use of eye+nose regions can increase FPA detection performance.

In the second stage of the study, the FPA detection performance of the MB-$LBP_{8,1}$ + PCA features obtained from the wide face region was examined. The results have showed that these features increase the FPA detection performance in 22 test scenarios. The obtained results have revealed that all the LBP patterns carry significant information in the FPA detection problem.

All videos in the CASIA dataset were taken under the same lighting conditions. In the REPLAY-ATTACK dataset, the attack videos were taken under two different lighting conditions. In the first environment, there is a fixed background and fluorescent lighting, while there is a non-uniform background in daylight in the second environment. The OULU dataset is designed to evaluate the generalization of FPA detection methods. Especially in the first protocol, it is tested how the methods behave in the previously unseen illumination conditions and background scene. Also, LBP texture descriptor extracts local features from local areas, so it is less affected by various lighting conditions. According to our experimental results, the proposed method obtained good results under different lighting conditions.

In future studies, the effects of these patterns on facial regions and the hybrid use of eye+nose regions on FPA detection can be examined. In addition, FPA detection performances of the regions can be examined with deep learning-based approaches.

Data Availability Not applicable.

Code availability Not applicable.

Author contributions Not applicable.

Funding Not applicable.

Declarations

Conflict of interest The authors have no conficts of interest to declare that are relevant to the content of this article.

References

1. Agarwal A, Singh R, Vatsa M (2016) Face anti-spoofing using Haralick features. 2016 IEEE 8th international conference on biometrics theory, applications and systems, BTAS 2016. https://doi.org/10.1109/BTAS.2016.7791171
2. Alotaibi A, Mahmood A (2017) Deep face liveness detection based on nonlinear diffusion using convolution neural network. SIViP 11:713–720. https://doi.org/10.1007/s11760-016-1014-2
3. Alpaydin E (2010) Introduction to machine learning third edition. Introduction to Machine Learning. https://doi.org/10.1007/978-1-62703-748-8_7
4. Anjos A, Chakka MM, Marcel S (2014) Motion-based counter-measures to photo attacks in face recognition. IET Biometrics 3:147–158. https://doi.org/10.1049/iet-bmt.2012.0071
5. Arashloo SR, Kittler J (2018) An anomaly detection approach to face spoofing detection: A new formulation and evaluation protocol. IEEE International Joint Conference on Biometrics, IJCB 2017 2018-Janua: 80–89. https://doi.org/10.1109/BTAS.2017.8272685
6. Bekhouche SE, Ouafi A (2015) AUTOMATIC AGE ESTIMATION AND GENDER CLASSIFICATION IN THE WILD Laboratory of LAGE, University of Ouargla, Algeria LAMIH Laboratory, UMR CNRS 8201 UVHC, University of Valenciennes, France Center for Machine Vision Research, University of Oulu, Finla
7. Boulkenafet Z, Komulainen J, Hadid A (2015) Face anti-spoofing based on color texture analysis. Proceedings - International Conference on Image Processing, ICIP 2015-Decem:2636–2640. https://doi.org/10.1109/ICIP.2015.7351280
8. Boulkenafet Z, Komulainen J, Hadid A (2016) Face spoofing detection using colour texture analysis. IEEE Trans Inf Forensics Secur 11(8):1818–1830. https://doi.org/10.1109/TIFS.2016.2555286
9. Boulkenafet Z, Komulainen J, Hadid A (2017) Face antispoofing using speeded-up robust features and fisher vector encoding. IEEE Signal Process Lett 24:141–145. https://doi.org/10.1109/LSP.2016.2630740
10. Boulkenafet Z, Komulainen J, Li L et al (2017) OULU-NPU: a Mobile face presentation attack database with real-world variations. In: proceedings - 12th IEEE international conference on automatic face and gesture recognition
11. Boulkenafet Z, Komulainen J, Hadid A (2018) On the generalization of color texture-based face anti-spoofing. Image Vis Comput 77:1–9
12. Chingovska I, Anjos A, Marcel S (2012) on the effectiveness of local binary patterns in face anti-spoofing. In: Proceedings of the International Conference of the Biometrics Special Interest Group, BIOSIG 2012
13. Croux C, Filzmoser P, Fritz H (2013) Robust sparse principal component analysis. Technometrics 55:202–214. https://doi.org/10.1080/00401706.2012.727746

14. Curtis S (2013) iPhone 5s fingerprint sensor "hacked" within days of launch. In: The Telegraph. http://www.telegraph.co.uk/technology/apple/iphone/10327635/iPhone-5s-fingerprint-sensor-hacked-within-days-of-launch.html. Accessed 5 Jan 2021
15. De Luis-García R, Alberola-López C, Aghzout O, Ruiz-Alzola J (2003) Biometric identification systems. Signal Process 83:2539–2557. https://doi.org/10.1016/j.sigpro.2003.08.001
16. De Souza GB, Da Silva Santos DF, Pires RG et al (2017) Deep texture features for robust face spoofing detection. IEEE Trans Circuits Syst II: Express Briefs 64:1397–1401. https://doi.org/10.1109/TCSII.2017.2764460
17. Erdogmus N, Marcel S (2014) Spoofing face recognition with 3D masks. IEEE Trans Inf Forensics Secur 9:1084–1097. https://doi.org/10.1109/TIFS.2014.2322255
18. Galbally J, Marcel S, Fierrez J (2014) Biometric antispoofing methods: a survey in face recognition. IEEE Access 2:1530–1552. https://doi.org/10.1109/ACCESS.2014.2381273
19. Galbally J, Marcel S, Fierrez J (2014) Image quality assessment for fake biometric detection: application to Iris, fingerprint, and face recognition. IEEE Trans Image Process 23:710–724. https://doi.org/10.1109/TIP.2013.2292332
20. Günay A, Nabiyev V (2017) Yüz Bölgelerinin Yaş Tahmini Başarımlarının Yaş Gruplarına Göre Değerlendirilmesi. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi 9:1–10
21. Günay Yılmaz A, Turhal U, Nabiyev VV (2020) Effect of Feature Selection With Meta-Heuristic 5:48–59
22. Kavzoğlu T, Çölkesen İ (2010) Destek Vektör Makineleri ile Uydu Görüntülerinin Sınıflandırılmasında Kernel Fonksiyonlarının Etkilerinin İncelenmesi. 73–82
23. Khurshid A, Tamayo SC, Fernandes E et al (2019) A robust and real-time face anti-spoofing method based on texture feature analysis. In: International Conference on Human-Computer Interaction. Springer, pp. 484–496
24. Kim I, Ahn J, Kim D (2017) Face spoofing detection with highlight removal effect and distortions. 2016 IEEE international conference on systems, man, and cybernetics, SMC 2016 - conference proceedings 4299–4304. https://doi.org/10.1109/SMC.2016.7844907
25. King DE (2009) Dlib-ml: a machine learning toolkit. J Mach Learn Res 10:1755–1758
26. Komulainen J, Hadid A, Pietikainen M et al (2013) Complementary countermeasures for detecting scenic face spoofing attacks. Proceedings - 2013 international conference on biometrics, ICB 2013. https://doi.org/10.1109/ICB.2013.6612968
27. Määttä J, Hadid A, Pietikäinen M (2011) Face spoofing detection from single images using micro-texture analysis. In: 2011 international joint conference on biometrics, IJCB 2011
28. Määttä J, Hadid A, Pietikäinen M (2012) Face spoofing detection from single images using texture and local shape analysis. IET Biometrics 1:3–10. https://doi.org/10.1049/iet-bmt.2011.0009
29. Marcel S, Nixon MS, Li SZ (2014) Handbook of biometric anti-spoofing-trusted biometrics under spoofing attacks
30. Ming Z, Visani M, Luqman MM, Burie J-C (2020) A survey on anti-spoofing methods for face recognition with RGB cameras of generic consumer devices
31. Ojala T, Pietikäinen M, Mäenpää T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24:971–987. https://doi.org/10.1109/TPAMI.2002.1017623
32. Raja R, Sinha TS, Patra RK, Tiwari S (2018) Physiological trait-based biometrical authentication of human-face using LGXP and ANN techniques. Int J Inf Comput Secur 10:303–320. https://doi.org/10.1504/IJICS.2018.091468
33. Shu X, Tang H, Huang S (2021) Face spoofing detection based on chromatic ED-LBP texture feature. Multimedia Systems 27:161–176. https://doi.org/10.1007/s00530-020-00719-9
34. Souza L, Oliveira L, Pamplona M, Papa J (2018) How far did we get in face spoofing detection? Eng Appl Artif Intell 72:368–381. https://doi.org/10.1016/j.engappai.2018.04.013
35. Sthevanie F, Ramadhani KN (2018) Spoofing detection on facial images recognition using LBP and GLCM combination. J Phys Conf Ser 971:012014. https://doi.org/10.1088/1742-6596/971/1/012014
36. Tan X, Li Y, Liu J, Jiang L (2010) Face liveness detection from a single image with sparse low rank bilinear discriminative model. In: lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). Pp 504–517
37. Tian Y, Xiang S (2016) Detection of video-based face spoofing using LBP and multiscale DCT. In: International Workshop on Digital Watermarking. Springer, pp. 16–28
38. Tirunagari S, Poh N, Windridge D, Iorliam A, Suki N, Ho ATS (2015) Detection of face spoofing using visual dynamics. IEEE Trans Inf Forensics Secur 10:762–777. https://doi.org/10.1109/TIFS.2015.2406533
39. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3:71–86. https://doi.org/10.1162/jocn.1991.3.1.71
40. Wen D, Han H, Jain AK (2015) Face spoof detection with image distortion analysis. IEEE Trans Inf Forensics Secur 10(4):746–761. https://doi.org/10.1109/TIFS.2015.2400395

41. Yang J, Lei Z, Liao S, Li SZ (2013) Face liveness detection with component dependent descriptor. Proceedings - 2013 international conference on biometrics, ICB 2013. https://doi.org/10.1109/ICB.2013.6612955
42. Yang J, Lei Z, Yi D, Li SZ (2015) Person-specific face antispoofing with subject domain adaptation. IEEE Trans Inf Forensics Secur 10:797–809
43. Zhang W, Xiang S (2020) Face anti-spoofing detection based on DWT-LBP-DCT features. Signal Process Image Commun 89:115990. https://doi.org/10.1016/j.image.2020.115990
44. Zhang Z, Yan J, Liu S et al (2012) A face antispoofing database with diverse attacks. In: proceedings - 2012 5th IAPR international conference on biometrics, ICB 2012. Pp 2–7
45. Zhao Q (2021) Research on the application of local binary patterns based on color distance in image classification. Multimed Tools Appl 80:27279–27298
46. Zhao X, Lin Y, Heikkila J (2018) Dynamic texture recognition using volume local binary count patterns with an application to 2D face spoofing detection. IEEE Trans Multimedia 20:552–566. https://doi.org/10.1109/TMM.2017.2750415