Journal of Network Intelligence ©2022 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **7**, Number **2**, May **2022**

Face anti-spoofing detection based on color texture structure analysis

Yi-Jia Zhang, Jia-Yin Chen

School of Information Science and Technology
Zhejiang Sci-Tech University
Hangzhou 310018, P. R. China
waiting@zstu.edu.cn

Zhe-Ming Lu

School of Aeronautics and Astronautics
Zhejiang University
No. 38, Zheda Road, Hangzhou 310027, P. R. China
Corresponding author: zheminglu@zju.edu.cn

Received October 2021; revised January 2022

Abstract. *Traditional face anti-spoofing detection schemes generally only focus on the study of color texture features, while ignoring the light intensity features of face images, which leads to the lack of partial face information. We propose a face anti-spoofing detection method based on the fusion of Co-occurrence of Adjacent LBPs (CoALBP), Local Phase Quantization (LPQ) and LDN-TOP (local directional number pattern from three orthogonal planes). The CoALBP descriptor can obtain rich texture information and spatial structure information of the image, and LPQ can just make up for the high resolution requirements of the CoALBP descriptor for the image, but the first two features only focus on the color texture signal, and ignore the transitional information of the face brightness. LDN can just encode the structural information of the face texture and the brightness intensity change of the light-dark transition, and can better extract the key point information and the brightness texture information of the face. Extending the LDN to three orthogonal planes, which enriches the appearance and movement information of the face image. We make up for this loophole in this research, fuse the color texture features and light intensity features of the human face, the fusion of these three features improves the accuracy of spoofing face classification, and achieved good experimental results. We verify our scheme on two public data sets and compared with other algorithms, which proves the superiority of our proposed scheme.*
Keywords: Face spoofing, color texture, presentation attack, feature fusion

---

1. **Introduction.** The current face recognition technology is becoming mature ,and it's widely and successfully applied to various authentication tasks. Following this are various face presentation attacks such as photo attacks, video attacks, and 3D masks attacks. These attacks pose a great threat to the security of face recognition systems. The face detection technology is an image detection technology to prevent the face recognition system from being attacked by this type of illegal face, thereby protecting the customer's information security and property security. At present, the face spoofing attacks that we generally discuss are mainly aimed at the re-photographed photos of legitimate users and video images containing the user's facial information. Because the acquisition of 3D face masks is more complicated, the production cost is high, and special materials and

production tools are required [1], the attack detection of 3D face masks is not discussed here. Attacking the face recognition system through illegal face spoofing is already a crisis that all face recognition systems must face. At present, some users have suffered property losses due to such illegal facial attacks. It can be seen that, effective face anti-spoofing detection algorithms have extremely high technical value in today's modern society [2]. Only when the security of the system is guaranteed can users be more assured to use the new generation of identity authentication systems.

In the early stage of the face anti-spoofing detection task, most of the countermeasures were to perform brightness analysis on grayscale images. Later, Boulkenafet [3] et al. proposed a novel countermeasure based on color texture, which mainly used a local texture descriptor LBP, it has demonstrated very good description and judgment ability in the scheme of texture analysis and face recognition, thus opening a new era of texture analysis in the face anti-spoofing detection scheme. However, color texture analysis has certain limitations, that is, although it captures the very important texture feature descriptor in face image detection, it completely ignores the light structure information of the face image. There is a feature descriptor LDN (local directional number pattern) in capturing texture structure and lighting information [4]. It is generally used for face detection and expression recognition, but due to its powerful description of face texture structure and brightness changes, it is also an excellent countermeasure used in face anti-spoofing detection.

In recent years, the existing software-based anti-spoofing liveness detection methods have made a lot of progress, which can be roughly divided into the following four types: (1) texture based methods; (2) micro-motion based methods; (3) vital signals frequency based methods; (4) other hardware based methods. In this section, we will briefly discuss these related methods.

Early texture-based methods focused on analyzing the brightness information of a face image to extract its gray-scale texture features, while ignoring the color texture information. Multi-scale local binary mode is used to obtain print quality defects, specular reflections and texture differences in shadows to detect whether the face in front of the camera is alive or a printed photo or video.The methods based on color texture are basically manual extraction and fusion of features. For example, Boulkenafet et al. [3,6] did not choose instead of conventional grayscale images, but innovatively extracted color texture features from HSV and YCbCr color spaces for the first time[7]. Compared with gray information, the combination of color information significantly improves the performance. Tiago et al. [5] used multi-scale LBP features, especially LBP-TOP features to extract the structure and dynamic features of the micro-texture of the face, and also extended the detection to the spatio-temporal domain, which effectively improved the detection performance. In addition, the development of Gaussian (DoG) [9] has also effectively assisted in improving the effect of face detection.Inheriting the ideas of the predecessors, Jukka et al. [10] focused on the study of the micro-texture difference between real face images and printed face images, using multi-scale LBP to extract such a feature space, the detection effect is good and computationally fast.

In addition to the spatio-temporal expansion of texture descriptors mentioned earlier, there are other motion-based methods. The difference between a real face and a spoofing face is that the real body has some important signals, such as heartbeat, blood flow, blinking, subtle movements of facial muscles, lip movements, etc. The face optical flow analysis method is to find the optical flow field of the previous frame by analyzing the change of the optical flow field in the face video or image sequence over time, and by analyzing the correlation between the optical flow field between adjacent frames The corresponding motion relationship with the optical flow field of the face in the current frame, so as to calculate the motion information of the face that changes over time. Bao et al. [11] proposed a face detection algorithm based on the optical flow field. Since the living human face has a three-dimensional spatial structure, the photo face is a two-dimensional planar structure. When the human face moves in front of the camera, the optical flow field of each area of the living human face changes more complicatedly, while the optical flow field of each area of the photo face changes more uniformly. Finally, by analyzing the change relationship of the optical flow field between different frames, the living human face and the photo face are distinguished. The algorithm has achieved good results in experiments, but because the optical flow field analysis method is extremely sensitive to the lighting conditions, it is difficult to use the optical flow analysis method due to the changes in the lighting conditions under actual application conditions, such as excessive or dark illumination. At the same time, the optical flow field analysis method cannot detect the forgery attack of the three-dimensional face mask.

Fourier transform is the basis of digital image processing, which analyzes the extracted image features by switching between the time domain and the frequency domain. LPQ is a fuzzy-insensitive texture feature description operator proposed by Heikkilä et al[8],it's a texture descriptor, but it is mainly used to improve the quality of the image, that is, to increase the resolution. It is generally used to process blurred images, which is very useful for the detection of blurred faces in face images.Bao et al. [11] judged true and false faces in the frequency domain, and gave two settings: a) Compared with the real face, the photo is a flat structure with fewer high-frequency components; b) There is no relative movement in the printed face photos, and there is no obvious change in a certain period of time. Based on these two points, there are certain shortcomings in the judgment of true and false faces. On the one hand, the high frequency components of the photo face increase with the improvement of the clarity of the printed photo, and on the other hand, the spatial information of the image is not considered. The Fourier spectrum analysis method is relatively simple, but the algorithm is not robust and is easily affected by image resolution and illumination. When the face image is clear, the Fourier spectrogram has more high-frequency components, and vice versa. With the widespread application of high-definition cameras, its shortcomings are highlighted.

In addition to VIS images captured by ordinary cameras, the researchers also studied clues in images captured by other hardware, such as near-infrared cameras, short-wave infrared cameras, thermal infrared cameras, multi-spectral cameras, light field cameras, and face anti-spoofing depth camera.

There are also methods based on image quality. Because of the surface material of the anti-spoofing face, there are specular reflections on many spoofing facial images, causing some printing attacks and concealment attacks to look blurry. The color diversity of spoofing faces largely depends on its production tools, and they lead to distortions in the colors of spoofing face images. Wen et al. [12] analyzed the causes of image distortion from the aspects of specular reflection, blur intensity, chromaticity moment, and color diversity to distinguish between spoofing faces and real faces. It analyzes the overall reflection difference of real human faces and obtains better generalization performance. The method based on image quality improves the generalization performance of face anti-spoofing, but it is difficult to accurately identify high-definition matte photos and display attacks.

Generally speaking, the image is stored in RGB format[13], but because the features we extract are not only related to color texture, but also related to brightness contrast changes, and Y in the YCbCr color space just represents the brightness channel, and Cb and Cr are colors. In this experiment, the YCbCr color space is used for the face anti-spoofing experiment.In this paper, we use three descriptors, i.e., CoALBP, LPQ and

LDN-TOP. They can well describe the face image in the two aspects of image lighting structure and color texture analysis. This is a face anti-spoofing detection scheme that combines brightness analysis and color texture analysis. In the current face anti-spoofing detection field, there are few such methods. We organize the remainder of this paper into four sections. Section 2 describes our method in detail. Section 3 introduces the databases used and experimental settings. Section 4 draws conclusions of this work.

All in all, the main contribution of this work is: in the face anti-spoofing detection task, it not only makes full use of the color and texture information of the face image, but also combines the texture movement information in the spatial structure of the face image under the change of external brightness. This combines color texture detection and brightness transition detection at the same time. This idea has not appeared in the previous face anti-spoofing live detection scheme, and the detection effect is excellent. At present, the EER on the Replay-Attack dataset has reached 0.07.

2. **Proposed Method.** This paper proposes a face anti-spoofing countermeasure based on the color texture structure analysis of CoALBP, LPQ and LDN. The overall face representation is extracted from the brightness and chroma images in different color spaces, and then different texture descriptors are studied. The performance effect in detecting various types of spoofing faces. The overall block diagram of the proposed face anti-spoofing detection scheme is shown in Fig .1.

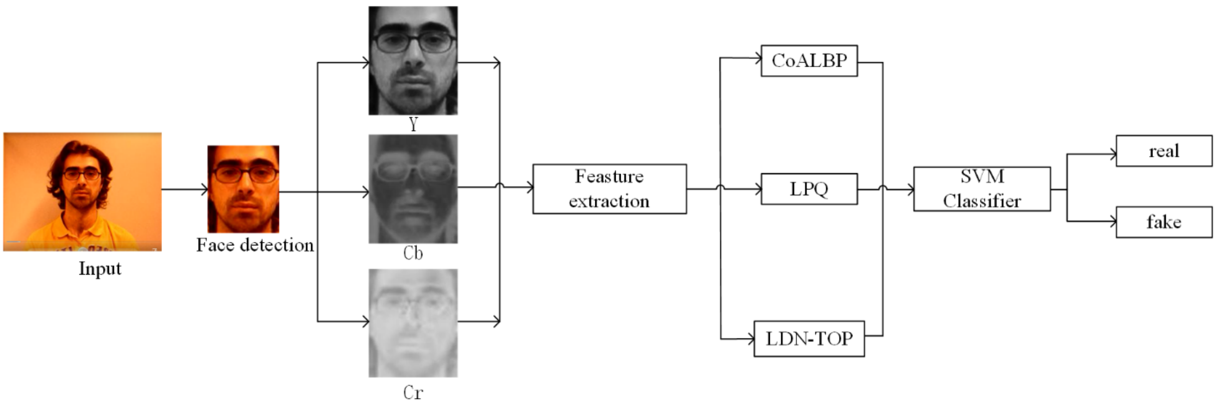

FIGURE 1. Framework of the proposed scheme

First, the input face video is cropped and normalized to a $64/times64$ pixel image. Second, the input face image information is extracted from the three channels of YCbCr to extract the CoALBP, LPQ, and LDN-TOP of the face information feature histogram. Then, the above feature data is normalized and stitched to finally obtain the overall representation of the face color texture histogram. Finally, the above training data is sent to the binary classifier, and then the model test is proceeded to obtain the detection result to determine whether the face in the tested video data is a real person or a spoofing face.

2.1. **Feature Extraction.** The descriptiveness and completeness of the extracted features are very important in the face anti-spoofing liveness detection algorithm, which largely determines the accuracy of detection and classification. This paper proposes a face anti-spoofing live detection method based on the mixed features of CoALBP, LPQ and LDN-TOP. They have shown very promising detection performance and effects in previous studies. The details of each of these features are described as follows.

2.1.1. *Co-occurrence of Adjacent Local Binary Patterns (CoALBP).* LBP feature descriptors are widely used in texture detection in the field of image detection. In order to obtain the spatial structure relationship information between pixels, the author in [7] proposed Co-occurrence of Adjacent Local Binary Patterns (CoALBP). CoALBP is an improved feature descriptor based on the LBP descriptor. It not only retains the powerful color and texture description performance of LBP, but also makes up for the spatial relationship information of image pixels missing from LBP, and can better resist the attack of light changes.

Compared with other descriptors (such as LBP), the motivation for using the CoALBP descriptor is: first, CoALBP completely contains the texture information contained in the LBP descriptor. Second, CoALBP also solves one of the LBP descriptors through the idea of co-occurrence big flaw-it is forcibly packed into a single histogram, which results in the lack of spatial relationship information between the descriptors, that is, the CoALBP descriptor can also obtain global image information and is robust to brightness transformation.

Let us first explain the mechanism of LBP. LBP mainly describes the size relationship between the central pixel and its neighboring pixels in a tiny image. Fig .2 shows the relationship between the micro image and the corresponding LBP.

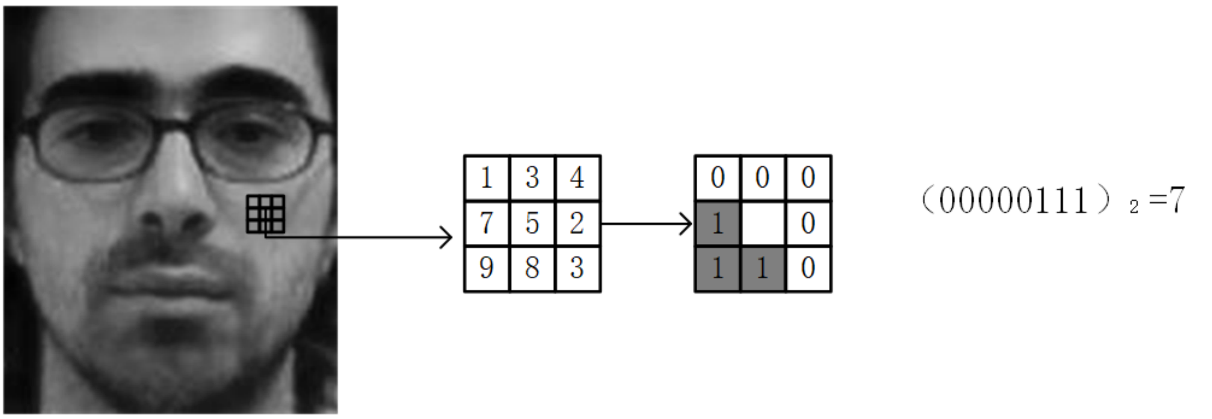

FIGURE 2. The micro pattern and its corresponding LBP feature descriptor

Generally $N_n = 8$ is adopted, that is, the displacement vector $\Delta s_i = 1$, and 8 points are taken around the center pixel, the obtained binary number is converted into a decimal number, and finally it is converted into a LBP histogram and shaping and connection are performed.

The idea of CoALBP feature descriptors comes from the combination of multiple LBPs that can show a richer spatial structure pattern of image pixels, so as to well solve the unity that a single LBP can only represent a simple image pattern.

The calculation process of the CoALBP feature descriptor is as follows:

First, with two simplified sparse configurations LBP(+) and LBP(×), a two-dimensional histogram is created in each direction between two adjacent LBP patterns. Convert each LBP into a vector ;

Then the correlation between four spatially adjacent modes is captured, and four directions are defined:

$$D = \{(0, \Delta B), (\Delta B, 0), (\Delta B, \Delta B), (-\Delta B, \Delta B)\} \tag{1}$$

Then, we calculate the histogram representation of the four $N_p \times N_p$ autocorrelation matrices $N_p \times N_p$ that co-occur in the space of adjacent LBPs:

$$H\left(a\right)=\sum_{r\in I}f\left(r\right)f\left(r+a\right) \tag{2}$$

The a in Eq. (2) is the displacement vector from the reference LBP to its adjacent LBP.

Finally, the obtained histograms are sorted and connected to form the final a feature vectors.

2.1.2. *Local Phase Quantization (LPQ).* Local Phase Quantization (LPQ) is a variant of LBP. The image is given by the Fourier phase spectrum obtained by quantization. For descriptions that are not sensitive to blur and linear illumination changes, the LPQ descriptor improves the image quality by deblurring.

The motivation for using LPQ descriptors is: first, LPQ has image blur insensitivity and strong robustness, that is, when the image is affected by blur, the LPQ descriptors can perform better than LBP descriptors in texture analysis. Second, the CoALBP descriptor requires high resolution of the image, and the texture analysis performed when the image is blurred is not good, so the combination of the two greatly enhances the effect of color texture analysis.

In the spatial domain, the blurred image is the convolutional representation of the image intensity and the point spread function (PSF). In the frequency domain, the process can be represented by a function:

$$G=O\times H \tag{3}$$

Among them,G,O,H represent the blurred image, the original image and the point spread function PSF respectively. From this idea, the calculation process of the LPQ operator is inspired: First, for the input image $f\left(x\right)$, the SIFI algorithm is used in the 3×3 area N_x to obtain the SIFI feature value of the spatial frequency u at the pixel x :

$$F\left(u,x\right)=\sum_{y\in N_x}f\left(x-y\right)e^{-2j\pi u^T y} \tag{4}$$

where the spatial frequency is:

$$u_1{=}[\alpha,0]^T,u_2{=}[0,\alpha]^T,u_3{=}[\alpha,\alpha]^T,u_4{=}[\alpha,-\alpha]^T \tag{5}$$

where α is the PSF of $f\left(x\right)$,that is, the scalar of $h\left(x\right)$ under the first zero crossing.
Second, $Q\left(F\left(u,x\right)\right)$ is quantified as follows:

$$Q\left(F\left(u,x\right)\right)=\mathrm{sgn}\left(\mathrm{Re}\left\{F\left(u,x\right)\right\}+2\mathrm{sgn}\left(\mathrm{Im}\left\{F\left(u,x\right)\right\}\right)\right) \tag{6}$$

Again, for the 8 coefficients in the 3×3 area, each $Q\left(F\left(u,x\right)\right)$ gets the integer value of 2^8=256 binary quantization coefficients, and they are connected in series to get the LPQ texture value of pixel x :

$$\rho_{LPQ}=\sum_{1}^{256}Q\left(F\left(u,x_i\right)\right)\cdot 2^{2(i-1)} \tag{7}$$

Then the decorrelation processing is carried out based on the whitening transformation to obtain the statistically independent LPQ value m_i ; Finally, the histogram is normalized as follows

$$m_i'=\frac{m_i}{\sum\limits_{i=1}^{256}m_i} \tag{8}$$

and the histogram of the LPQ feature operator $m'_1, m'_2, ..., m'_{256}$ is obtained.

2.1.3. *Local Directional Number Pattern from Three Orthogonal Planes (LDN-TOP).* Reference [4] proposed a face descriptor named Local Directional Number Pattern (LDN) for enhancing the ability of face recognition. It encodes the structure of the local neighborhood by analyzing its direction information, so as to realize face recognition. The structural information of texture and the encoding of intensity changes are generally used for face recognition and expression recognition, but the reference [4] proves that LDN performs equally well in face spoofing detection.Zhou et al.[14] extends the LDN feature into three orthogonal planes, and the results obtained further prove that the brightness change feature is extremely meaningful for face anti-spoofing detection.

The motivation for using the LDN descriptor is: first, LDN encodes each pixel structure in the surrounding area by analyzing the direction information, so as to generate corresponding descriptors for different texture structures of the face image, so it can distinguish between bright and dark changes. Second, the derivative of partial Gaussian is used to create an asymmetric compass mask. The size of the mask is variable, so that we can merge facial features of different resolutions, and the mask can resist a single light change and noise. At the same time, the edge response detection effect is good, and these two are exactly the weaknesses of the basic descriptor of color texture analysis LBP. The fusion of LDN descriptors just compensates for the defects of texture analysis in resisting light changes and noise attacks. Finally, LBP only selects sparse points for calculation, but LDN makes full use of the pixel information of the entire field, and it is only six bits in length, the description code is less but the information contained is richer.

In our experiments, it is constructed in the center pixel of the image sequence. The LDN operator is calculated on the three orthogonal planes (namely XY, XT and TY) that intersect in the middle, thus enriching the spatial structure information and motion information of the face texture image.

Input the face image I , and finally calculate the corresponding feature vectors H_i^{XY}, H_i^{XT} and H_i^{TY} on the three planes of XY, XT and TY respectively.

For each plane, we first convolve the 3×3 area I_i of a certain central pixel (x, y) in the face image I with the Kirshc operator mask M^i in 8 directions to obtain the i-th 8 edge response values of the direction:

$$e_i(x, y) = I_i * M^i, 0 \leq i \leq 7 \tag{9}$$

The aim of the partial Gaussian derivatives is to create an asymmetric compass mask, which is used to calculate the edge response of a smooth face, the mask is robust to noise and illumination changes, and can generate powerful edge response. A variable σ is defined here as the width of the Gaussian bell. By changing σ, the Gaussian mask can be changed, and then the size of the mask can be changed to obtain richer facial features. As shown in the following formulas (10) and (11), the Gaussian mask is defined as:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2}\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{10}$$

and the mask is defined as

$$M_\sigma(x, y) = G'_\sigma(x + k, y) * G_\sigma(x, y) \tag{11}$$

Second, the number of directions is calculated corresponding to the maximum and minimum edge response values:

$$I_{x,y} = \arg\max_i \{e_i(x, y) \,|\, 0 \leq i \leq 7\} \tag{12}$$

$$J_{x,y} = \arg\max_{j} \left\{ e_j \left(x,y \right) | 0 \leq j \leq 7 \right\} \tag{13}$$

Then, the LDN feature value of pixel (x, y) is calculated by $I_{x,y}$, $J_{x,y}$:

$$LDN\left(x,y\right) = 8 \times I_{x,y} + J_{x,y} \tag{14}$$

Thus, the LDN operators $H_i^{XY}, H_i^{XT}, H_i^{TY}$ of the pixel (x, y) on the XY, XT and TY planes are obtained in turn.

Finally, the LDN operators of the three planes of XY, XT and TY are combined into

$$f_i = \left[H_i^{XY}, H_i^{XT}, H_i^{TY} \right] \tag{15}$$

and the final feature vector is expressed as

$$LDN{-}TOP_{[\sigma_1,...,\sigma_n]} = \left[f_1, ..., f_i, ..., f_I \right] \tag{16}$$

3. **Database and Experimental Setup.** This section begins with the introduction of datasets, followed by the evaluation of our proposed algorithm, and we compare the method with a single algorithm. All the experiments are conducted on a desktop equipped with Core-i7 and 8-GB RAM, and the implementation and the experimentation of the algorithms were carried out using MATLAB R2018b version.

3.1. **Datasets.** We choose two public datasets for evaluation, i.e., Replay-Attack, MSU MFSD, to demonstrate the effectiveness of our scheme. More details can be found below.

Replay-Attack: This dataset contains 1200 videos. The videos collected the real and deceptive face information of 50 volunteers. The holding methods of the attack images are divided into fixed mode and handheld mode, and they are divided into the training set with 360 videos, the verification set with 360 videos, and the test set with 480 videos. Each object has 24 videos, including 4 legitimate requests and 20 spoofing attacks. Each video is longer than 9s. All the videos are in 3 different types. Scene and recording are under 2 different lighting conditions.

MSU MFSD: This dataset has a total of 280 real and deceptive face videos. The video recording time is over 9 seconds, and the average recording time is 12 seconds, and 25 pictures per second. It uses a mobile phone to collect images. The deception attack is printed in a paper of the same size, and a picture containing all the background information can cover the entire phone camera.

3.2. **Evaluation.** In this experiment, we strictly follow the official test standards of the two databases Replay-Attack and MSU MFSD, and make a reasonable comparison with the methods proposed in the above-mentioned sections. Among them, the Replay-Attack database provides a separate development set to adjust model parameters; while MSU MFSD lacks a predefined test set, so we use Cross-Validation to train and adjust model parameters, and use equal error rate (EER) and Half error rate (HTER) evaluates the performance of face spoofing detection, FRR is the probability of predicting a real face as a fake face, FAR is the probability of predicting a fake face as a real face. EER is half of the sum of FRR and FAR when FRR equals to FAR. HTER is defined in Eq. (17). HTER means half of the sum of FAR and FRR. The specific meaning is half of the sum of the proportions of true and false faces that are each judged wrong. It is the compromise error rate under different thresholds. Usually the threshold is EER.

$$HTER = \frac{FRR + FAR}{2} \tag{17}$$

The pre-processing of the face image includes image frame sampling and face alignment, also the position of the eyes of the face image should also be aligned and normalized to 64×64 pixels, which is also helpful to reduce the impact of face normalization.

For the CoALBP descriptor, the features were a concatenation of three histograms computed using the LBP(+) operator with radius and the corresponding directions defined by the distances $B = [2, 4, 8]$. For the LDN descriptor, we set the width of the Gaussian bell $\sigma = [1.0, 1.5, 2.0]$, and divide each plane into 16×16 blocks. The binary classification was performed using a linear support vector machine (SVM) classifier (using LIBLINEAR [15]).

3.3. **Discussion.** We design several groups of experiments on Replay-Attack and MSU MFSD. Above all,we perform some single tests of color texture descriptors such as CoALBP, LPQ, and LDN-TOP On the YCbCr color space, which proves that all three descriptors contribute to face anti-spoofing detection. Next,according to the characteristics of the three descriptors in terms of color texture description, improvement of image clarity, and image light intensity description, it can be seen that these descriptors are complementary, and through a certain feature fusion method, the superior performance of the combination of these three descriptors is obtained.At last, we compare the performance of our method with the state-of-the-art algorithms.

3.3.1. *Fusion of Complementary Color Texture Representations.* The features extracted in this experiment are all extracted and feature fused in the YCbCr color space. Table.1 shows the hidden complementary performance of CoALBP, LPQ and LDN-TOP face descriptors. We can see that when only the color texture descriptor is the CoALBP feature, the EER and HTER on the Replay-Attack and MSU MFSD data sets are relatively high, and the spoofing detection effect of the face is relatively general, especially the EER on the MSU MFSD reaches As of 4.21, when the LPQ feature that improves the image resolution is integrated, EER and HTER have significantly decreased,the reason is that the face color texture descriptors and the brightness dynamic texture descriptors have excellent description performance for the face.Previous studies have shown that color texture and brightness dynamic texture are usually used separately for face anti-spoofing detection research. Our experiment innovatively connects their excellent descriptors through their generated histograms to achieve the fusion of color texture features and brightness dynamic texture features. These three face texture representations that combine color texture features and bright dynamic texture features can learn from shortcomings and complement each other, the performance improvement of face anti-spoofing detection is very significant.

Table 1. The performance on the development and test set of the Replay-Attack and MSU databases in terms of EER(%) and HTER(%), respectively

Descriptor	Replay-Attack		MSU MFSD	
Evaluation Index	EER	HTER	EER	HTER
CoALBP	1.34	1.41	4.21	4.07
LPQ	1.93	1.70	3.79	3.82
LDN-TOP$_{[1.0,1.5,2.0]}$	0.87	1.04	2.04	2.02
CoALBP+LPQ	1.70	1.92	1.98	2.87
CoALBP+LPQ+LDN-TOP$_{[1.0,1.5,2.0]}$	0.07	1.30	1.75	1.71

3.3.2. *Comparison with the State of the Art Schemes.* In the following experiment, our method is compared with the most advanced face anti-spoofing scheme proposed in the literature.Fig .3,Fig .4,Table. 2 and Table. 3 show that some current comparisons of the most advanced technologies on the Replay-Attack and MSU MFSD databases. The following results show that our method based on color texture and brightness dynamic texture is better than other most advanced face anti-spoofing schemes. Both EER and HTER are low, that is, the detection accuracy is well and the generalization ability is strong,and the experiment demonstrates well competitive performance on the Replay-Attack Database.

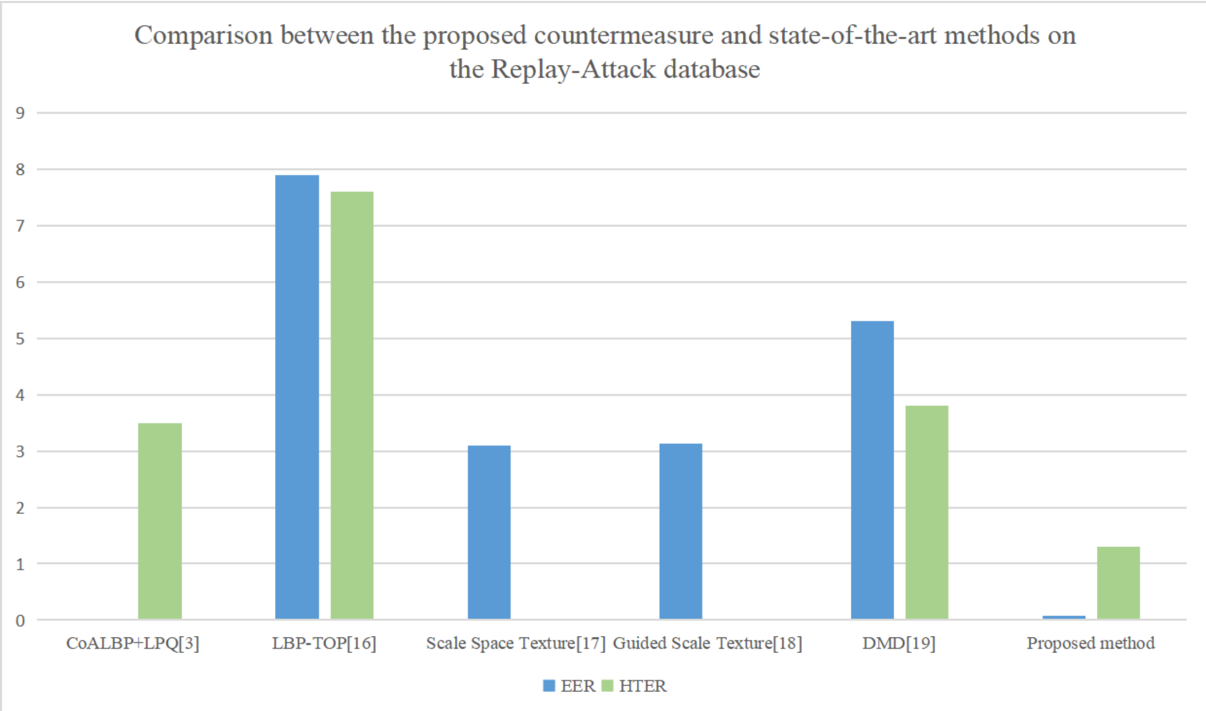

Figure 3. Comparison between the proposed countermeasure and state-of-the-art methods on the Replay-Attack database

Table 2. Comparison between the proposed countermeasure and state-of-the-art methods on the Replay-Attack database

Descriptor	Replay-Attack	
Evaluation Index	EER	HTER
CoALBP+LPQ[3]	0.0	3.5
LBP-TOP[16]	7.9	7.6
Scale Space Texture[17]	3.1	-
Guided Scale Texture[18]	3.13	-
DMD[19]	5.3	3.8
Proposed method	0.07	1.30

4. **Conclusion.** In this paper, we proposed a novel and effective face anti-spoofing method based on dynamic color texture construction. This paper uses the strong image texture information described by the original descriptor LBP retained in CoALBP, and also makes

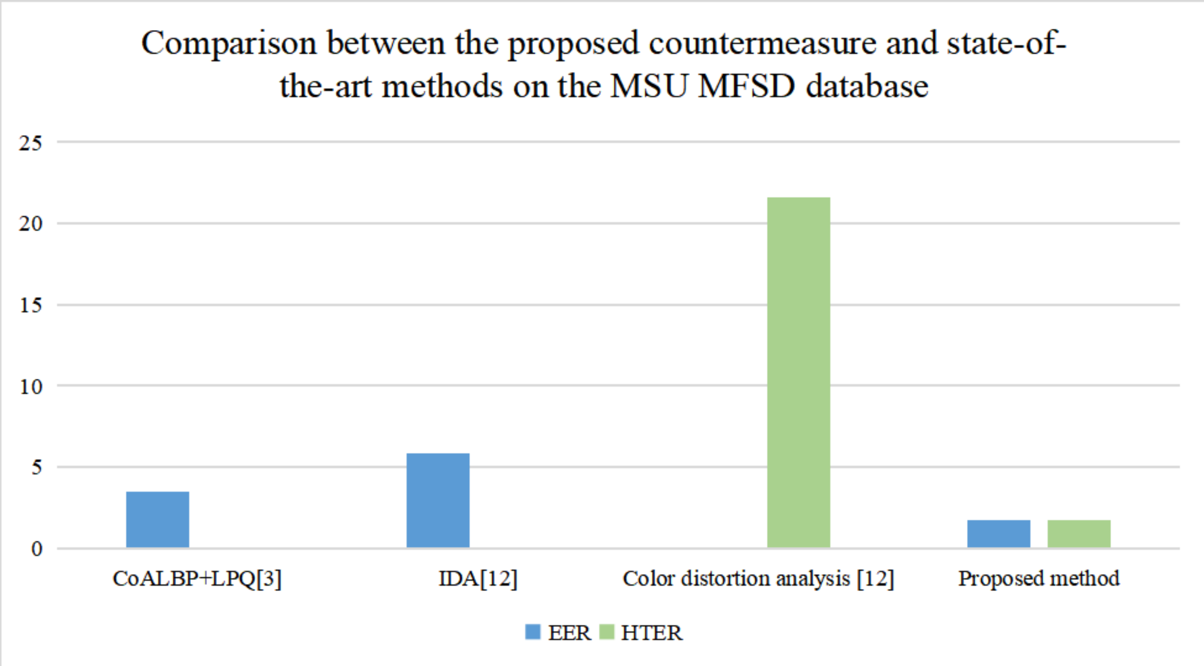

FIGURE 4. Comparison between the proposed countermeasure and state-of-the-art methods on the MSU MFSD database

up for the rich image structure spatial information missing in LBP, and then uses the deblurring and denoising characteristics of LPQ to obtain a clearer image. The texture feature can better deal with the problem of insufficient clarity of the input image, and further enhance the processing ability of image texture information. LDN can encode the structural information of the face texture and the brightness intensity change of the light-dark transition, and better extract the key point information and brightness texture information of the face. The fusion of the three features improves the accuracy of spoofing face classification, and the algorithm has achieved good detection results on multiple authoritative data sets. Our method achieved the excellent performance on Replay-Attack and MSU MFSD with EER and HTER. The generalization ability of the datasets is strong, and it has good practical application scenarios in the future.

The mainstream face anti-spoofing detection basically only performs the feature extraction of color texture.This experiment is different from the traditional method,the brightness feature and the color texture feature are combined to make up for the missing parts of the two types of different features,better realize a more comprehensive extraction of facial features. Future work focuses on combining other much better features [20-22]

TABLE 3. Comparison between the proposed countermeasure and state-of-the-art methods on the MSU MFSD database

Descriptor	MSU MFSD	
Evaluation Index	EER	HTER
CoALBP+LPQ[3]	3.5	-
IDA[12]	5.82	-
Color distortion analysis [12]	-	21.6
Proposed method	1.75	1.71

to obtain much better performance,and we will incorporate ideas such as image comparison learning, and further combine advanced IoT technologies[23] and deep learning technologies[24] to improve the integrity of the algorithm as much as possible, and improve the accuracy of algorithm detection, so as to better cope with the upcoming era of comprehensive information.

Acknowledgment. This work is supported in part by the National Key Research and Development Program of China under Grant No.2020AAA0140004 and Science Foundation of Zhejiang Sci-Tech University(ZSTU) under Grant No.19032458-Y. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

[1] Z. Sun,L. Sun and Q. Li, Investigation in spatial-temporal domain for face spoof detection, *IEEE International Conference on Acoustics, Speech and Signal Processing*,pp. 1538-1542, 2018.

[2] T. Edmunds, and A. Caplier, Motion-based countermeasure against photo and video spoofing attacks in face recognition, *Journal of Visual Communication and Image Representation*, Vol.2018, No. 50, pp. 314-332, 2018.

[3] Z. Boulkenafet, J. Komulainen, and A. Hadid, Face spoofing detection using colour texture analysis, *IEEE Transactions on Information Forensics and Security*, Vol. 11, No. 8, pp. 1818-1830, 2016.

[4] A. R. Rivera, J. R. Castillo, and O. O. Chae, Local directional number pattern for face analysis: face and expression recognition, *IEEE Transactions on Image Processing*, Vol. 22, No. 5, pp. 1740-1752, 2013.

[5] T. de F. Pereira, J. Komulainen, A. Anjos, J. M. De Martino, A. Hadid, M. Pietikäinen, and S. Marcel, Face liveness detection using dynamic texture, *EURASIP Journal on Image and Video Processing*, Vol.2014, No.2, pp. 1-15, 2014.

[6] Z. Boulkenafet, J. Komulainen, and A. Hadid, Face anti-spoofing based on color texture analysis, *IEEE International Conference on Image Processing*,pp. 2636-2640, 2015.

[7] R. Nosaka, Y. Ohkawa, and K. Fukui, Feature extraction based on co-occurrence of adjacent local binary patterns, *Pacific Rim Conference on Advances in Image and Video Technology*, Springer-Verlag, pp.82-91,2012.

[8] J. Heikkilä, E. Rahtu, and V. Ojansivu, Local phase quantization for blur insensitive texture description, *Local Binary Patterns: New Variants and Applications. Studies in Computational Intelligence*, Vol. 506, pp. 49-84, 2014.

[9] W. E. Polakowski, D. A. Cournoyer, S. K. Rogers, M. P. DeSimio, D. W. Ruck, J. W. Hoffmeister, and R. A. Raines, Computer-aided breast cancer detection and diagnosis of masses using difference of Gaussians and derivative-based feature saliency, *IEEE Transactions on Medical Imaging*,Vol. 16, No. 6, pp. 811-819, 1997.

[10] J. Määttä, A. Hadid, and M. Pietikäinen, Face spoofing detection from single images using micro-texture analysis, *International Joint Conference on Biometrics*, pp.1-7, 2011.

[11] W. Bao, H. Li, N. Li, and W. Jiang, A liveness detection method for face recognition based on optical flow field, *International Conference on Image Analysis and Signal Processing*, pp. 233-236, 2009.

[12] D. Wen, H. Han, and A. K. Jain, Face spoof detection with image distortion analysis, *IEEE Transactions on Information Forensics and Security*, Vol. 10, No. 4, pp. 746-761, 2015.

[13] Chi-Ming Chen, Chien-Chih Lin, Yu-Sheng Liu, Jin-Xiang Liao, Ming-Tsung Yeh, Yi-Nung Chung and Chao-Hsing Hsu, Applying Image Processing Technology to Automatically Detect and Adjust Paper Benchmark for Printing Machine, *Journal of Information Hiding and Multimedia Signal Processing*,Vol. 12, No. 2, pp. 56-64, June 2021.

[14] J. Zhou, K. Shu, P. Liu, J. Xiang and S. Xiong, Face Anti-Spoofing Based on Dynamic Color Texture Analysis Using Local Directional Number Pattern, *International Conference on Pattern Recognition*,pp. 4221-4228, 2021.

[15] R. E. Fan, K. W. Chang, C. J. Hsieh, X. R. Wang, and C. J. Lin, LIBLINEAR: A library for large linear classification, *Journal of Machine Learning Research*, Vol. 9, No. 61, pp. 1871-1874, 2008.

[16] T. de F. Pereira, A. Anjos, J. M. De Martino, and S. Marcel, Can face anti-spoofing countermeasures work in a real world scenario? *International Conference on Biometrics (ICB)*,pp. 1-8, 2013.

[17] Z. Boulkenafet, J. Komulainen, X. Feng, and A. Hadid, Scale space texture analysis for face anti-spoofing, *International Conference on Biometrics*, pp. 1-6, 2016.

[18] F. Peng, L. Qin, and M. Long, Face presentation attack detection using guided scale texture, *Multimedia Tools and Applications*, Vol. 77, No. 7,pp. 8883-8909, 2018.

[19] S. Tirunagari, N. Poh, D. Windridge, A. Iorliam, N. Suki, and A. T. S. Ho, Detection of face spoofing using visual dynamics, *IEEE Transactions on Information Forensics and Security*,Vol. 10, No. 4,pp. 762-777, 2015.

[20] G. D. Simanjuntak, K. N. Ramadhani, and A. Arifianto, Face spoofing detection using color distortion features and principal component analysis, *International Conference on Information and Communication Technology*, pp. 1-5, 2019.

[21] F. Utaminingrum, A. D. Purwanto, M. R. R. Masruri, K. Ogata, and I. K. Somawirata, Eye movement and blink detection for selecting menu on-screen display using probability analysis based on facial landmark, *International Journal of Innovative Computing, Information and Control*, Vol.17, No. 4, pp.1287-1303, 2021.

[22] L. B. Zhang, F Peng, L Qin, and M. Long, Face spoofing detection based on color texture markov feature and support vector machine recursive feature elimination, *Journal of Visual Communication and Image Representation*,Vol.2018, No. 51, pp.56-69, 2018.

[23] Ke Wang, Chien-Ming Chen, Zuodong Liang, Mohammad Mehedi Hassan,Giuseppe M.L. Sarné, Lidia Fotia, Giancarlo Fortino, A trusted consensus fusion scheme for decentralized collaborated learning in massive IoT domain, *Information Fusion*,Vol.72, pp.100-109, 2021.

[24] Eric Ke Wang, Xun Zhang, Fan Wang, Tsu-Yang Wu, and Chien-Ming Chen, Multilayer Dense Attention Model for Image Caption, *IEEE ACCESS*,Vol.7, pp. 66358-66368, 2019.