Multimedia Tools and Applications (2021) 80:14693–14721
https://doi.org/10.1007/s11042-020-10338-1

Survey of non-intrusive face spoof detection methods

Pooja R. Patil[1] · **Subhash S. Kulkarni**[1]

Received: 11 May 2020 / Revised: 12 November 2020 / Accepted: 22 December 2020 /
Published online: 28 January 2021

Abstract

Biometrics are distinct physiological characteristics used to describe individuals. Compared to the traditional access control methods such as passwords and Person Identification Numbers (PIN) which can be forgotten and shared easily, biometrics are widely used in authentication systems. Even though the accuracy of face recognition systems is lower than that of the systems using fingerprint, iris, etc. as the acquisition devices of the latter evade the affine and photometric transformations, recognition systems with the face as a trait are widely used due to the contactless and non-intrusive nature of the acquisition device-camera. As the cameras are in-built in most of the handheld and portable devices such as mobile phones and laptops, the uncontrolled and/or unregulated immediacy of sharing the photographs via messaging services and uploading on social networks entices the attackers to create spoofs to deceive a face recognition system. Hence, it is necessary to incorporate a spoof detection algorithm in recognition systems before revealing the identity. This paper gives an overview of the steps involved in the face spoof detection process, the various databases available, the different measures to discern between live and spoof images, aligned with the perceived observance, the binary classifiers used, and the performance evaluation parameters revealed in the literature.

Keywords Biometrics · Face recognition systems · Security · Face liveness detection · Spoof attack and detection · Intrusive and non-intrusive methods · Binary classification

1 Introduction

Traditional access control methods that use passwords are still popular because of their static nature when compared with the recognition systems that use biometric data which can be subject to change either naturally or accidentally. But there is a high chance that the passwords are lent and shared and can also be forgotten easily which may take some time

✉ Subhash S. Kulkarni
sskul@pes.edu

Pooja R. Patil
ppoojapatil844@gmail.com

[1] PESIT - Bangalore South Campus, Bengaluru, Karnataka, India

to recover. To circumvent these limitations, the usage of biometric data in recognition systems is increasing. However, the use of biometric data raises security concerns which may stall the distribution of the data. There are various methods to generate spoofing attacks that emulate the specific physiological attributes of an individual so as to alleviate the performance of authentication systems. Hence liveness detection is very important in recognition systems to increase the level of security that can be implemented either by using hardware or software-based methods. Software-based liveness detection methods are widely used to avoid additional costs required in hardware-based methods. Among the other biometrics such as iris, fingerprint, palm print, etc., the face is the commonly exploited biometric trait due to it's non-intrusive and contactless nature. The face spoof detection process can be performed using one of the two categories: intrusive and non-intrusive, as stated in Table 1.

Face anti-spoofing methods using intrusive approach requires either user intervention or a closer contact of the user to the device capturing the biometric The nearer the device to the user, the more intrusive is the device. For example, a surveillance camera that can recognize people face remotely is less intrusive than a biometric system capturing the fingerprint or imaging sensors touching the user eye to scan the retina. While the non-intrusive methods using security cameras are a regular feature of many public spaces and their presence has become ubiquitous because of their non-intrusive abilities to detect, recognize, and identify individuals without requiring active participation or the knowledge of the subject.

Techniques using intrusive mode focus on vitality signs and base their decisions on human involuntary actions such as eye reflexes, lip, and head movements. The key, but a major limitation, of these methods, is based on the assumption that the user will experience such liveness indications within a given time frame and hence these methods fail if there is a delay in response. These methods mainly target static spoofing attacks and hence can be easily deceived by the replay (video) attacks or using eye-cut photos and hence pose a challenge either to detect or reject these replay attacks. An alternative for these approaches can be challenge-response methods, another intrusive approach, which explicitly asks the user to perform certain random action(s) to verify the liveness. Representatives of this type require cooperation from the user to capture this additional data. Other intrusive methods requiring users' response may fail when they reject to respond and methods requiring auxiliary devices such as sensors, cameras, etc. may become difficult for deployment. However, the software-based methods, which are purely based on image analysis without users' intervention, are widely used and accepted because of their non-intrusive in nature even though these methods lack generalization ability compared to some aforementioned approaches.

The face spoof detection process can be interpreted as a binary classification problem to classify the face images as either live or spoof. This initially begins with the exploration of the types of spoofing mediums and the available datasets which are generally preprocessed.

Table 1 Different spoof detection methods

Si. No.	Categories	Type	Challenges
1.	Intrusive methods	Based on human involuntary actions	These methods fail when video spoofs are presented
		Based on user interactions	These methods fail when users refuse to respond
2.	Non-intrusive methods	Hardware based methods	These methods become expensive due to additional auxiliary devices and are location dependent
		Software based methods	Widely accepted due to the non-intrusive nature

Face detection and size normalization are commonly used techniques in applications using face images. It is essential to explore, understand, and visualize the hidden patterns in the data one is working with to find out the differences between the live and spoof images. The different spoofing mediums and the datasets considered in the literature are discussed in Sections 2 and 3.

The next step involves the extraction of features selection. It is very challenging to select the features that effectively extract the underpinned patterns which can efficiently bring out the differences between the live and spoof images while being invariant to affine transformations and dynamic external factors such as changing illumination conditions. The step also demands features to be as minimum as possible to avoid the additional computational costs, training time, and memory usage. Section 4 gives a brief introduction of the features involved in the face spoof detection research.

The procedure continues with the classification step which requires the separation of the available dataset into training and test sets. The training set is divided into k sets and the model trained with k-1 sets is used to predict the results on the remaining set, for k-fold cross-validation, to improve the predictive performance of the classifier. Then a classifier is chosen to predict the exact class of the test set. The workflow of the spoof analysis is depicted in Fig. 1.

Section 5 describes some commonly used binary classifiers in face liveness detection. The organization of the rest of the paper is as follows: Section 6 gives the various evaluation metrics for comparative analysis of the suitability of the features used and the classifiers chosen. Section 7 gives a review of the experiments carried out in the spoof detection process by various authors followed by some conclusions drawn in Section 8.

2 Different types of spoofs

Several ways are possible to pose a security threat to the face recognition systems and hence the study of anti-spoofing methods against spoofing attacks has been active in recent years. Different types of spoofing attacks and different properties of the medium used to generate these spoofing attacks and external variations such as illumination and shadowing effects enhance the difficulty to find robust and efficient countermeasures.

2.1 Display or mobile spoofs

The image of a person's face can be presented to the biometric system through display devices such as laptops, mobile phones, monitors, etc.

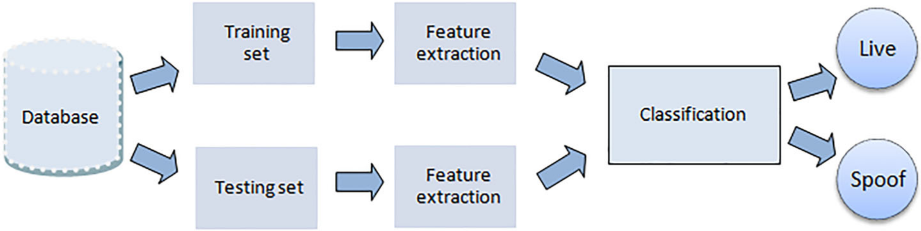

Fig. 1 Workflow of spoof analysis

2.2 Print spoofs

A face recognition system can be misled by printing the image of an authorized person's face using high-quality color printers.

2.3 Replay attacks or video spoofs

Videos of a genuine person are presented to masquerade a recognition system.

2.4 Face masks

Online 3D mask makers and 3D printing technology can be used to fraud a face recognition system.

2.5 Graphic spoofs

Face images printed on photographic sheets, posters or magazines are other kinds of threats to a biometric system.

2.6 Caricatures and line drawings

Caricatures and line drawings of human faces may be used to deceive a face recognition system.

These spoofs can be presented to the biometric system in a variety of ways: the use of a *flat printed photo* has greater potential to take place due to the availability of such pictures either photographed by an impostor or on social media, the *eye-cut photo* attack with eye regions of imposter cut off to exhibit blink behavior of the impostor, *warped photo* attacks, bending a printed photo in any direction to simulate facial motion. A *Replay attack* via video playback shows almost all behaviors similar to real faces, with many of the intrinsic features of valid user movements. This type of attack has physiological signs of life that are not presented in photos, such as eye blinking, facial expressions, and movements in the head and mouth, and it can be easily performed using tablets or large smartphones. 3D mask attacks are of two types: life-size wearable mask and paper-cut mask [64]. These attacks are addressed to anti-spoofing systems that analyze 3D face structures, is one of the most complex attacks to be detected. Mask manufacturing is much more difficult and expensive than the other types of attacks, requiring 3D scanning and printing special devices.

3 Databases available and their description

Depending on the attack types, the process of producing spoofing attacks is time-consuming and sometimes requires a lot of resources and certain manufacturing skills. Therefore, it is not difficult to imagine that collecting attack data for many subjects may become very demanding, and, for certain types of attacks such as 3D masks, too expensive. Perhaps this is one of the main reasons why the number of publicly available face spoofing databases and the number of subjects involved in their creation is limited. A brief description of commonly used face spoofing databases, differing in the data format, the number of clients and samples, protocol, types of attacks, as well as the quality of the recording devices is described below and some sample images are depicted in the Fig. 2. The different datasets used in

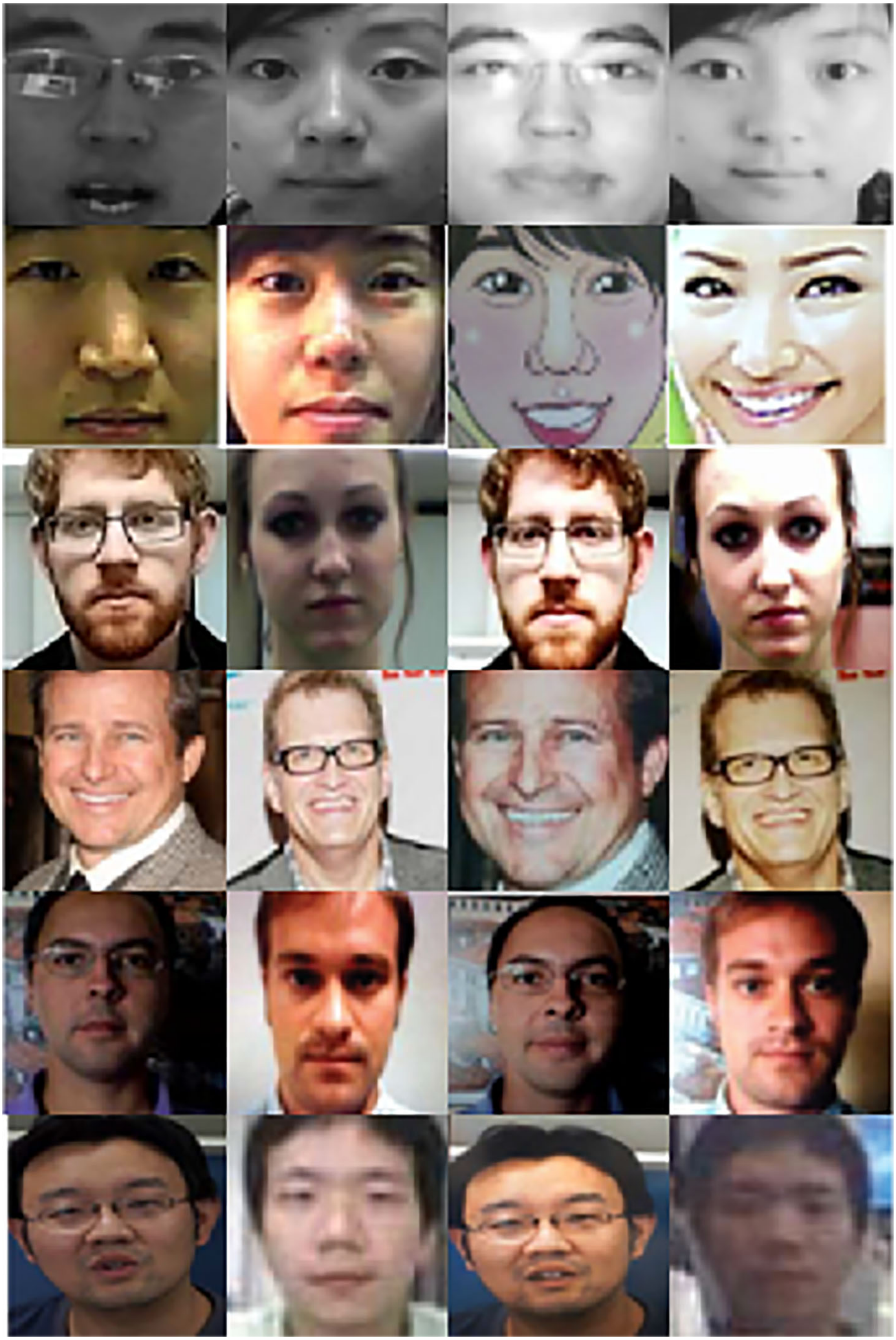

Fig. 2 First two coloumns represent live and last two columns represent spoof face images. Each row represent the sample images of NUAA, BERC Webcam and ATM, MSU MFSD, MSU USSA, Idiap Replay Attack and CASIA FAS datasets

Table 2 Commonly used datasets and their characteristics for face spoof analysis

Si. No.	Dataset	Spoof type	No. of subjects	Acquisition device for live face images	Acquisition device for spoof face images
1.	NUAA	Print spoofs	15	Webcam	Webcam
2.	MSU MFSD	Print spoof Replayed video	55	Nexus 5 MacBook	Canon 550D iPhone 5s
3.	MSU USSA	Print spoofs Replayed videos	1140	MacBook	Nexus 5 iPhone 6
4.	Replay Attack	Print spoof Replayed video	50	MacBook webcam	Canon PowerShot
5.	Print Attack	Print spoofs	50	MacBook webcam	MacBook webcam
6.	CASIA FASD	Print spoof Replayed video	50	Sony NEX - 5 USB camera	Sony NEX - 5 Webcam

spoof analysis with their spoof type are tabulated in the Table 2. Databases with different types of spoofs, discussed in Section 2, have been collected by the researchers and some of them with subjects' consent are publicly available or will be shared on request.

3.1 NUAA Photograph database

The NUAA imposter database is a publicly available database for face spoof detection.[1] The database is collected using generic webcams at different locations and under various illumination conditions with a gap of two weeks between successive intervals. A total of 15 subjects are invited to participate in the session. The series of live images are captured using webcams at a rate of 20 fps and 500 images are collected per subject with additional constraints imposed, such as restricted expressions, involuntary head movements, and eye blinks, to make live images look, as close as possible, like imposters. The spoof images are collected by first capturing face images using a Canon camera in such a way that the face covers 2/3rd of the area of the photograph and are then printed on (i) photographic papers of size 6.8 cm × 10.2 cm and 8.9 cm × 12.7 cm using a traditional photo printing method, and 70 GSM A4 paper using a color HP printer and finally these five categories of spoof images are recaptured using a webcam.

3.2 Zahid et. al's dataset

In [1], the authors used Print Attack, Replay Attack, NUAA Imposter Yale Recaptured datasets which are commonly used in the state of the art methods. As these datasets contain video clips, they extracted 20 images from each clip. The authors also created their dataset which comprises face images of 40 subjects with five different expressions per subject. The spoof images are collected by displaying five face images of each subject taken from social networks. As a whole, the dataset contains 200 live and 200 display spoof images.

[1] http://parnec.nuaa.edu.cn/_upload/tpl/02/db/731/template731/pages/xtan/NUAAImposterDB_download.html

3.3 Yale recaptured dataset

This dataset consists of 640 live face images and 1920 LCD spoof images of 10 subjects. A subset of the Yale face database B is used as genuine images that are captured using a geodesic lighting rig constructed with 64 xenon strobes controlled by a computer. The rig is used to vary the illumination and 64 images of a face under a specific pose are acquired in about 2 s and hence contain images with fewer facial expressions and head movements. For spoof images, the face images of Yale database B are displayed on three LCD monitors: LG Flatron 1900, a CTL 171Lx 1700 TFT, and a DELL Inspiron notebook and are recaptured using Kodak C813 of 8.2 megapixels (MP) and a Samsung Omnia i900 of 5 MP cameras from a distance of 50 cm from the screen [1].

3.4 BERC Webcam database

This database contains live images of 25 subjects collected using the conventional webcam and spoof images of four types: (i) The images of 20 subjects collected under three different illumination conditions: indoor without any external lightings, with strong frontal illumination, and with light source inclined to the camera axis and printed on photographic of size 10.2 cm × 15.2 cm and 29.7 cm × 21 cm using the conventional method (ii) The same set of images are printed on A4 sheets using a laser color printer. (iii) 60 face images each of size in the range 5–8 cm and 60 with their size ranging from 9–14 cm are collected from magazines. The distance here is measured from the line slightly above the eyebrows to the line above the chin. (iv) 60 different varieties of caricature images are collected from the web are printed in sizes of 5–8 cm and 9–14 cm using a laser color printer. All the images in this database are of size 640 × 480. The database contains a total of 1408 live and 6461 spoof images.

3.5 BERC ATM database

This database contains all the live and fake images collected under normal indoor lighting conditions and the resolution of these images is comparatively low due to the plastic cover on the in-built camera of the ATM. The live face images of 25 subjects and fake face images of 20 subjects are selected identical to that of the Webcam database. The dimension of both live and spoof images is 640 × 480. The dataset comprises 1797 live and 5802 face images [36].

3.6 J. Li. et. al's dataset

In [48], the authors created their dataset by inviting 4 subjects for the session. The live images are captured using the Logitech Quickcam Pro 4000 camera. The fake face images are collected by printing on photographic sheets of size 48 × 33 mm and 76 × 55 mm and A4 sheets using a color printer. Here, all the frames extracted are cut manually.

3.7 MSU Mobile Face Spoofing Database (MFSD)

In this database, the video clips of genuine faces are captured using MacBook Air 13 laptop camera with a resolution of 640 × 480 at a frame rate of 30 frames/sec (fps) with the duration of each clip maintained at least for 9 s and using Google Nexus 5 Android camera with a resolution of 720 × 480.[2] The average distance between the camera and the face is approximately 50 cm. The videos of the subjects are captured using Canan 550D Single Lens Reflex (SLR) and also using an iPhone 5s back camera with a resolution of 1900 × 1088 and replayed using an iPad air screen and iPhone 5s Android phone in front of the biometric system with an average distance of approximately 20 cm and 10 cm respectively between the display screen and the system. The images (5184 × 3456) captured using Canon 550D camera are printed on A3 paper using an HP color laserjet printer and are presented to the system at an average distance of ≈ 40 cm [74].

3.8 MSU Unconstrained Smartphone Spoof Attack (USSA) Database

In this database, the live face images of 1,000 subjects are collected from the web faces database [72] which comprises images of celebrities at different locations and under varying illumination conditions and resolutions.[3] In addition to 1,000 subjects, 50 subjects from Idiap, 50 from CASIA FASD, and 40 from MSU MFSD are also included, and thus the average resolution of the images of all 1140 subjects is 705 × 865. For spoof attacks, the images are collected using front and rear cameras of Google Nexus 5 and are presented using three display devices: MacBook, Nexus 5, and tablet screens to enable the researchers to study the effect of the quality of images on spoof detection. The public set-the set of images with the subjects' consent to make their data publicly available, of MSU USSA, has 6840 images of different quality captured using different spoof mediums. To create print attacks, all the images of 1140 subjects are printed on a matte of 8.5 × 11-inch paper using an HP color laser printer in such a way that the face covers most of the area of the paper and presented to the frontal and rear cameras of Nexus 5 in a manner to minimize the reflections of the indoor lighting. In all cases, the aspect ratio is maintained in order to avoid distortions. 2280 images are collected in this way [61].

3.9 IDIAP Replay Attack database

The live images of 50 subjects are collected indoor without external illumination in different illumination conditions and varying external lightings using a MacBook webcam.[4] Spoofing video attacks are produced using Canon PowerShot SX 150 cameras for each subject under the same lighting and background conditions used for capturing live spoofs [14].

3.10 IDIAP Print Attack Database

In Print-Attack Replay Database, the live face images are collected in the form of 200 video clips of 50 subjects under various lighting conditions collected using a webcam.[5] These images are collected printed on A4 sheets and are recaptured using the same webcam under the same environment set up for collecting live face images.

[2] http://biometrics.cse.msu.edu/pubs/databases.html

[3] http://biometrics.cse.msu.edu/pubs/databases.html

[4] www.idiap.ch/dataset/replayattack

[5] www.idiap.ch/dataset/printattack

3.11 CASIA FASD

In this dataset, face images are acquired using Sony Nexus-5 and USB cameras of 50 subjects which are displayed on Ipad and recaptured using Sony Nexus-5 for generating spoofs [80].[6]

3.12 SFL

In this dataset, the live images of 23 subjects are captured using a smartphone under varying illumination conditions in both indoor and outdoor environments. The images displayed on this smartphone are (i) recaptured using another smartphone for generating mobile spoofs and (ii) printed on a photographic sheet of 12.5 × 17.7 cm which are finally captured using a smartphone [37].

The quality of the spoofing attacks for face mode is influenced by several factors. Firstly, the quality of the original image is used to generate a spoofing attack. For example, the original sample may be a mugshot image taken with the user's cooperation, or an image in adversary conditions taken from a distance or can be downloaded from the Internet. The quality of the recorded input may also vary and may depend on the circumstances under which the spoofing attack is performed, like the illumination conditions or the presence of supervision at the biometric system capturing device. Other factors, categorized by Common Criteria as important for attacks to any kind of information systems are technical expertise, knowledge about the capturing device, a window of opportunity, etc.

The attacker usually has direct influence neither on the quality of the original sample, which may likely be obtained in an opportunistic manner nor on the conditions at the side of the biometric system. However, he is fully responsible for the process of fabricating the attack, which includes the choice of the spoofing media, material, devices, and tools needed to perform the attack. These choices determine the type of the spoofing attack, as a broad description of its properties. The type of attack is the basic source of differences between the spoofing attacks, which often serve as cues to detect them. One of the properties of the spoofing attacks that is conditioned on their type is their dynamics. Another property is their dimensionality, i.e., the face spoofing attacks can be in 2D or 3D.

The basic types of attacks can further differ in several other aspects, which may or may not depend on the attacker's will. An example is an environment where the original sample is recorded, and it can be controlled or adversary. A fixed support or hand support can be used for holding the spoofing medium. On the contrary, in a close-up attack, the borders of the spoofing medium are integrally visible. This aspect is primarily influenced by the size of the original sample or the spoofing media used to display the attacks. The complexity, cost, and level of expertise to produce different types of spoofing attacks vary significantly. While producing a digital photo attack may require only access to the Internet and a consumer's mobile device, producing 3D masks may require expensive equipment, like a camera or 3D scanner and a 3D printer. The type of attack, as well as the properties related to its dynamics, dimensionality, or other factors, have an important impact on the choice of features used by the spoofing counter-measures.

[6] www.cbsr.ia.ac.cn/english/FaceAntiSpoofDatabases.asp

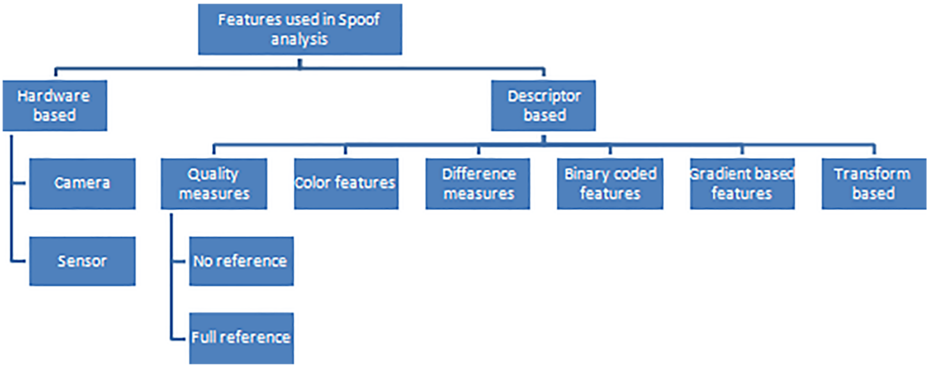

Fig. 3 Features used in spoof analysis

4 Features used in spoof analysis

The feature selection is the most prominent step in all ML applications and plays a major role in the accuracy of the proposed model and hence involves a careful observation of the data failing to which will have a negative impact on the predictor leads to misclassification. Figure 3 depicts different types of features and their characteristics which can contribute

Table 3 Characteristics of the features prevalent to distinguish between live and spoof images

	Feature type		Characteristics exploited to discern between live and spoof images
1.	Hardware based	Camera based	Multiple cameras are deployed to capture side views. Camera flashlight and focus is used to capture the DoF information near the local image parts.
		Sensor based	Various types of sensors are used to capture the involuntary actions of humans such as blood flow, odor etc.
2.	Descriptor based	Quality measures	Quality measures are extracted from pretrained models. Reflectance properties of the presenting, distortions introduced during multiple recapture, diffusion speed, energy ratios, SNR and it's variants and statistical moment descriptors are used.
		Color features	The decrease in diversity of colors of the spoofing medium due to the limitation on their resolutions are exploited using color moments and histograms in various color spaces.
		Difference measures	The difference between the original images and their corresponding blurred versions, sum of absolute pixel wise differences and their variants are used.
		Binary coded features	Each patch of an image is coded based on the intensity value of the center pixel to capture the micro-texture differences.
		Gradient based	The degradations in the edges of the images are exploited using line and edge detectors and histograms.
		Transform based	Fourier transform, wavelet and Gabor transforms are commonly used.

to Spoof analysis are summarized in Table 3 and a comprehensive study of each of these feature classes is described below.

4.1 Hardware based

The hardware-based spoof detection is employed for spoof analysis to capture the vitality signs of humans due to voluntary actions, which requires users' response and involuntary actions which can be captured using motion additional auxiliary devices such as sensors. These methods are highly accurate but the complexity of implementation, cost, and intrusiveness make these systems unacceptable.

4.1.1 Camera based

Focus and flashlight of the cameras or multiple cameras are used to focus the region of interest and to enhance the distinction between the spoof and live images. In [50], the flashlight of the camera is used to enhance the distinction between focused on the hair region of live and spoof images, [77] focuses on the nose and mouthparts of the face image to capture the depth information using Fourier transform.

4.1.2 Sensor based

Rhythmic changes in breathing is captured in [30, 63] uses thermal information, [48] uses motion analysis using Fourier spectra, [8] employs optical flow, [67] and [60] exploited the count of eye blinks, [59] and [59] uses odor, temperature, blood flow and blood pressure of the human body.

4.2 Descriptor based

The descriptive based Spoof analysis methods involve the effective representation of images and hence reduces the computation cost and time and are widely used because of the ease of implementation, inexpensiveness, and non-intrusive nature. The different descriptors used for Spoof analysis are as follows:

4.2.1 Quality measures

The quality measures for Spoof analysis are computed either using predictive models trained on thousands of images or by the comparative analysis of the quality of live and spoof images to capture degradations of the spoof images introduced during multiple captures. The difference in the reflective nature of the spoofing medium is also exploited for classification. In [11], the authors use no reference-quality measures obtained from pre-trained models such as (i) Natural Image Quality Evaluator (NIQE) which uses a model trained based on multivariate distribution to predict the quality, (ii) Blind Image Integrity Notator (BLINDS-2) which makes predictions using the model trained on normalized DCT (Discrete Cosine Transform) coefficients [44, 47] and so on. Galbally et al. [21] uses discriminant analysis and some correlation measures. For full reference methods, Gaussian blurred images are used.

In [74], the authors proposed four measures which are described as follows: (i) specular reflection features: The specular reflection component is extracted from the input face image and the fixed percentage along with the first two moment descriptors of the specular component by excluding the monochromatic pixels with high-intensity [22], (ii) Blurriness measures: The blurring effect caused when the spoofing attacks presented to the biometric system from a very short distance to conceal the boundaries is devised using two types of features: one based on the difference between the original image and its blurred version [15] and the other based on the average width of edges in the image [57] along with color moments and diversity features. The length of this feature is 101. All four features are concatenated together to form the Image Distortion Analysis (IDA) feature vector for liveness detection. The overlapping of the digital grids of the capturing and recapturing devices, called Moire patterns, can be observed as distinct peaks in the Fourier domain in the mid and high-frequency regions which are detected using a correlation-based peak detector exploited as features in [23]. In [37], diffused images are obtained using pixel-wise Total Variation (TV) flow with a stable Additive Operator Splitting (AOS) scheme, and the pixel-wise diffusion speed is calculated as the absolute difference between the original and diffused image in logarithmic space. This difference image is encoded locally to form a feature vector, called Local Speed Pattern (LSP), for spoof detection. In [37], random reflection characteristics of live images due to 3D shape is exploited using the diffusion process. Slower diffusion is observed for spoof images is due to the uniform distribution of illumination energies compared to live images and hence the difference between original and diffused images i.e., the diffusion, is used as a clue to discriminate between live and spoof images. In [42], the authors used Variance Energy Ratio (VER) which captures the changes in successive frames due to the vitality signs in the eye region for static analysis. The value greater than a certain threshold indicates liveness. The authors exploited the discontinuity between the foreground and the background using edge detection followed by line detection and Border Energy Ratio (BER) calculation. A higher value indicates liveness. In [24], the features used in this work are as follows:

luminance This measure is different for different surfaces and proportional to contrast in the images and is randomly distributed throughout for live images.

SSIM The quality of grayscale images is exploited using this measure and has the range (-1, 1).

Energy The sum of the pixels of the log of the Fourier transform of a single channel is calculated.

Entropy It is calculated at each pixel of an RGB image and then converted to grayscale.

Mean The mean of all the pixel values of an image in RGB and YCbCr color space is computed.

Skewness The symmetricity of the face is exploited using this measure and is zero for higher symmetry.

These measures are concatenated to form a single feature vector. Bhogal et al. [11] uses no-reference quality measure predictions of the pre-trained models based on statistical distributions for spoof analysis. In [54], specular reflection ratio which depends on the refractive index and intensity distribution of Hue channel and statistical measures of the GLCM (Gray Level Co-occurrence Matrix) of original and low pass filtered images are exploited as features. Yeh and Chang [79] uses distortion dependent pixel similarity deviation of a mean subtracted and contrast normalized (using standard deviation) image at selected gradient positions.

4.2.2 Color based features

In [74] uses color moment features: The degradation involved in chromatic reproduction property of display devices and printers which are commonly used to create spoofing attacks have been used and skewness along with the first two spatial moment descriptors of each channel in HSV color space are proposed. In addition to these chromatic features, the percentage of pixels in the histogram bins which contain minimum and the maximum count has been exploited as the other two features. Thus, there are five features per each channel of a given input image, and the difference between the diversity of colors of live and that of spoof images exploited by computing the histogram bin counts of the top 100 frequently appearing colors and the total number of distinct colors in quantized (32- bit) input image. In [70], color, edge, and Gabor descriptors and histograms in RGB and HSV color spaces of few frames for static analysis. Similarity scores have been computed for all the features (one per each feature) and are combined using the Dynamic Score Combination (DSC) method [69] with the majority voting rule. For, video analysis, the authors used the eye blink detector in [60] with some additional heuristics and some constraints on the number of blinks via a logistic function. Motion measurement is also performed to combat the spoofs of higher quality which, however, may fail when fixed photo attacks are used. In [31], authors extracted statistical textural features [35] from illumination maps in three color spaces HSV, YCbCr and LAB along with quality measures [21] to differentiate live faces from mobile spaces.

4.2.3 Difference measures

In [62], Bruno et al. focus on the problem of a variable which is very often in the operational scenario of a face recognition system: illumination. The blurring effect due to the brightness of the LCD screen has been, captured using DoG filter with two Gaussian masks having different standard deviations: $\sigma_1 = 0.5$ and $\sigma_2 = 1$ is exploited to distinguish between live face images from LCD spoofs in which high-frequency information i.e. edges of the image becomes difficult to detect. In [68], the Difference of Gaussian (DoG) filtered image transformed into frequency domain in the columnized form (i. e. R^p; $p = m \times n$ pixels) is used as a feature vector. Here, the DoG filter acts as a band pass filter where the very high frequencies are rejected to remove noise. The inability of the DoG filter to detect edges of live images with partial occlusions and shadows on one side of the image has been witnessed in [62] and the authors also used Contrast Limited Adaptive Histogram Equalization (CLAHE) that operates on small patches of the image and enhances the contrast of each patch and hence increases the ability of the DoG filter to detect edges even when the parts of the face are shadowed. The accuracy of the classification results depends on the choice of the following parameters: (i) number of tiles (ii) contrast enhancement limit and (iii) distribution parameter. The experimental observations show that higher contrast limit increase the discrimination ability and Rayleigh distribution preserves the originality as opposed to exponential distribution which introduces distortions in the brighter regions of the image.

4.2.4 Binary coded features

In [1], Locally Uniform Comparison Image Descriptor (LUCID) [81] is employed as a feature to discriminate between live and spoof images. The authors highlight the suitability of their proposed method to mobile applications due to the high speed as the feature descriptor does not require any floating-point operations and involve simple permutations of the intensity values of the image patch. The insensitivity of LUCID to the photometric transformations under varying illumination conditions and noise introduced by the image sensor during the capture and recapture of images motivated the authors to use the descriptor. The computation of LUCID is as follows: For grayscale images, all the values in the patch p of size n × n will be arranged in ascending order. This is repeated for all the patches in the image and finally, all the vectors are concatenated. For an image of size M × M, the size of the feature vector will be $1 \times M^2$. For RGB images, p will be an n × n patch with c color channels. The dimension for RGB image will be $1 \times 3M^2$. In [2], three different features are used to distinguish between live and spoof images: (i) LUCID [81] is computed on patch size of 24 × 24 of RGB image blurred using averaging kernel, (ii) CENsus TRasnform hiSTogram (CENTRIST) [75] on 3 × 3 image patches and (iii) Patterns of Oriented Edge Magnitudes (POEM) [71] built on pixel block of size 8 × 8 and cells of size 7 × 7 with unsigned representation having 3 bins. In [61], Local Binary Patterns ($LBP_{(8,1)}$) is computed on image patches of size 32 × 32 with an overlap of 16 pixels. These features from each patch are concatenated along with color moments (mean, variance and skewness) [74] to form a feature vector of length 4248. In [5, 7, 10, 17, 36, 51, 55], LBP and its variations are used in the literature to capture the texture information of spoof images. In [2], the proposed liveness detection system uses either single or multiple descriptors based on the Security Level (SL) of the biometric system which is controlled by the user. When SL is set to low, the system uses LUCID [81] as a descriptor to train the SVM classifier. When SL is set to medium, CENTRIST, which preserves the global information is used in addition to LUCID to capture both local and global information. Decision scores obtained by the classifiers trained separately using LUCID and CENTRIST are fused using logical AND rule for spoof detection. When SL is set to high, the system employs POEMs, in conjunction with LUCID and CENTRIST, which enhances both the local and global information by the distribution of edge directions. Again the AND based fusion method is used to combine the decision scores of three SVM classifiers trained individually using the above-mentioned features.

4.2.5 Gradient based

In [18], each extracted frame is normalized to the size 64 × 64 and transformed into HSV and YCbCr color spaces. Thus the input frame is split into six color channels: H, S, V, Y, Cb, and Cr. The gradient face image is computed for each of these channels using color gradient operator [18] and the histograms, of length 64, of each channel, are concatenated to form a feature vector. In [78], the authors divide the whole image into 3 × 3 patches and consider only the nose (central patch) and cheek (right bottom corner) parts of the image and reduce the size of the parts to 1/3rd of their original size. The authors highlight the difference in the gradient magnitude of the original and Gaussian blurred versions of both the live and spoof image parts. The classification is based on the fact that the sum of all the pixels of the difference image will be larger for live image parts due to the prominent gradients of the live images than those of recaptured face images. In [20], the proposed metrics of Image Quality Assessment (IQA) try to estimate the human perception of looking at an image. The quality metrics include some full-reference measures which require an undistorted version of the test image as a reference and hence the grayscale image is distorted by a 3 × 3 low pass Gaussian kernel with standard deviation $\sigma = 0.5$ produce a distorted version. As the spoof images captured using multiple cameras of different resolution introduce distortions, the quality of these recaptured images will be low when compared to live images which pass through a single camera, and hence the performance of the edge and corner detectors degrades when used for spoof images. These attributes are considered and nine pixel difference based measures, three correlation-based measures and two edge-based measures have been used to generate feature vector.

4.3 Transform based

In [23], several zero mean band pass filtered versions of the input images are transformed into the Fourier domain. The log magnitude of each of the images is fed to a correlation-based peak detector. The absence of peaks in all the filtered images indicates liveness. In [36], the detected face image is transformed into the frequency domain using two dimensional Discrete Fourier Transform (2D DFT). This produces a complex image and hence is analyzed using either real and imaginary parts or phase and magnitude responses and in most cases, the magnitude response is used for analysis in image processing as it preserves most of the structural information of the image in the spatial domain. The larger dynamic range of the Fourier image make it difficult to be displayed and thus necessitates the use of logarithmic transformation and in most implementations, the Fourier image is displayed in such a way that the image mean i.e. the DC or zero frequency value is at the center and the frequency increases with the distance from the center. This resultant image of size N × N is divided into $N/2$ concentric rings and finally, the feature vector is constructed by concatenating the average values of all the rings. The length of this feature vector is $N/2$. The commonly used feature LBP is employed to emphasize the loss of micro-texture of the spoof images compared to the live images. The histogram of the uniform LBP ($LBP_{(8,1)}$) coded image is used for classification. In [41], 3D shape information is used to capture the differences in the low-frequency region. Li et al. [49] uses energy and the standard deviation of the Fourier transformed image is used along with LBP for shape and micro-texture analysis. In [48], the difference between the surface normal and hence the intensity variations concerning the Lambertian model [9] is exploited as a feature to distinguish between live and fake images using High-Frequency Descriptor (HFD). The method fails when high-quality spoofs are used. The authors created their database with 4 subjects to test the performance of their proposed method.

Spoof detection methods based on motion analysis use properties of the human motion patterns in front of the system, to distinguish them from those in the presence of spoofing attacks. Some of these methods base their approach on the assumption that the movement patterns of a 2D object due to either handheld presentation attacks or replay attacks are different from 3D objects—live people. Some authors used the relative motion difference between the face parts closer to the camera (nose) and away from the camera (ears). Some methods exploit the correlation between the face parts and the background as a distinction factor and are suitable mainly for scenic spoofing attacks. These methods may fail when a significant amount of motion is introduced by an attacker or if the presumed motion patterns are absent during the acquisition period, similar to the techniques based on liveness cues, and hence can address only the static attacks and can be easily deceived by the replay attacks. In addition, some protocols as in [46] and [4] can resist replay attacks.

Anti-spoofing methods analyze the visual appearances and stand behind a strong argumentation about the differences in the visual properties of real accesses and spoofing attacks by proposing suitable measures, explained in several works. It is interesting to note that majority of visual appearance-based methods work even with only a single image at the input. They are usually applied either on the detected face, face parts, or the full input image. Recently, the analysis of the visual appearance has been extended into a temporal domain, i.e., the frames will be extracted from the video and the extracted features from each frame will be fused to classify the images. Also to increase the robustness of the countermeasures, some researchers have used multiple biometric traits like fingerprint along with face, to make it difficult for a fraudulent user to deceive the biometric systems.

All the methods described above have been reported with different success rates, which cannot be easily compared because they are obtained on different types of attacks and on various databases most of which are not released publicly. As one of their advantages, they are non-intrusive and user-friendly and do not depend on the user's behavior. These methods are expected to successfully detect any of the static or dynamic attacks, but their success may be questioned if the spoofing attacks are printed or displayed using high-resolution media, thus lack some of the artifacts these methods can rely on. Their generalization properties, when applied to different acquisition conditions or types of attacks they are not trained for are uncertain since the visual appearance of the images often depends on numerous factors viz: illumination conditions, acquisition devices, and their resolution or display media, etc. A way to mitigate this problem is to fuse several different anti-spoofing methods [45] to build a more generalized liveness detection scheme effective against a multitude of attack types using appropriate fusion schemes.

Following the trend in computer vision, the research community, working on spoof detection, has started experimenting via deep learning (DL) to automatically extract and learn the features directly from the data. This is in contrast to the machine learning (ML) methods where the features are crafted by closer inspection of the data inspired by some characteristics that are common in either all of the live or spoofing attacks. Although it can be argued that countermeasures engineered this way are suitable only for the type of artifacts they are designed for, some recent works using DL reported lower performance than ML-based methods.

5 Classifiers for spoof analysis

The unlimited number of possible ways to masquerade an identity management system by the attackers necessitates the knowledge of highly sophisticated statistical and analytical algorithms to unmask the disguised patterns in the data and to overcome the security breaches and threats. The quest for such statistical techniques led to the development of the predictive models—the machine learning or classification algorithms. The purpose of these algorithms is to obtain good predictive results which can be either the class labels or continuous real values. The only way to make the qualitative analysis of such algorithms is to compare the obtained predictions with the known labels, which is the procedure for supervised learning [56].

To generalize the algorithms to the unseen samples, the available data is divided into training and test sets with each instance in the training set containing its features and the corresponding class label, and the test set consisting only the features and the classification task is to build a model that can predict the class labels of these. The predictive performance of the model is proportional to the number of samples [28].

Most of the available supervised learning algorithms and their variants follow the same basic procedure to obtain a predictive model which is as follows:

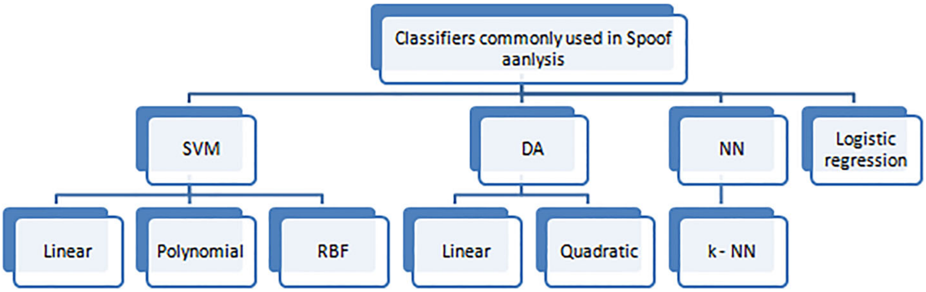

Fig. 4 Classifiers used commonly in spoof analysis

Data preparation The input to the classification algorithms will be in the form of a matrix with each row representing an instance (for example a person) and column, a feature. That is, the dimension of the matrix depends on the number of unique instances (rows) and the number of features or attributes extracted from each sample or instance (columns).

Choice of an algorithm The choice of a classification algorithm is a trial and error process because the accuracy of prediction of all the machine learning (ML) algorithms solely depends on the historical data and hence a single algorithm cannot provide the best results for every problem. However, it is essential to be aware of the key characteristics of the algorithms and the trade-offs to make a start such as training speed, predictive performance, memory usage, and transparency.

Fitting a model The fitting functions applicable for binary classification are:

- Support Vector Machines (SVMs)
- Discriminant analysis
- Naive Bayes
- Nearest neighbor

Choosing a Validation Method The accuracy of the obtained fitted model can be examined based on the resubstitution error[7] and cross-validation error.

Examining and tuning the fit until satisfied After validation, the models can be tuned by checking the performance of the available algorithms and using their various fitting parameters, for example, by using the different kernels for SVM viz: linear, polynomial, etc., different distributions for Naive Bayes classifier such as Gaussian, multinominal, etc., and so on, for better accuracy.

Using the resulting model for predictions Finally, the model can be used to make predictions on the new data to evaluate the performance of the classifier [12].

The process of face spoof detection can be visualized as a binary classification problem where the classifier takes the features of the sample face images with their categorical class as the input to classify the new sample as belonging to one of the two classes: live or spoof. Some classifiers which serve this purpose are described below and are depicted in Fig. 4.

[7] The difference between the known response of the training data and the predictions made by the classifier on the training data. A higher value indicates lower classification accuracy and a lower value does not guarantee good predictive results for unseen data.

5.1 Support Vector Machine (SVM)

An SVM finds the hyper plane (an M-1 dimensional space for M dimensional feature space), i.e., the plane that can separate data points of one class to lie on one side and those of the other on the other side of the hyper plane. The best hyper plane is the one with the maximum perpendicular distance between the two nearest points on both sides of the hyper plane. The points are referred to as support vectors [16]. When the training set is not linearly separable in the domain of existence, SVM transforms the data points in the original feature space (X, Y), representing the corresponding responses, to a higher-dimensional space, with the help of a function called the kernel function and the basic kernels of SVM are linear, polynomial, radial basis function (RBF) and sigmoid.

5.2 Linear Discriminant Analysis (LDA)

LDA projects the data points in the existing space into a new space to maximize the inter-class separability and minimize the intra class variance under the assumption that the data has Gaussian (normal) distribution. The various discriminant rules of separability are Bayes discriminant Rule, maximum likelihood, and Fisher's linear discriminant rule. With the assumption that each class has a different normal distribution, LDA first computes the mean of the data points of each class, the covariance between the data points by subtracting the mean of each class from the data points from that class and finally computes the covariance matrix of the result [26]. The trained model assigns the new data to the class that produces the lowest misclassification costs [19].

5.3 Naive Bayes classifier

The naive Bayes classifier is a probabilistic classifier with the strong assumption that the features are highly uncorrelated from one another within each class, but in practice, found to work well when this independence assumption is not valid. The classifier estimates underlying parameters of the probability distribution of each class using the training data with the assumption that the features are conditionally independent. The posterior probability of the test data belonging to each class is computed and assigned to the class with the maximum posterior probability [65].

5.4 Nearest neighbor

This classifier categorizes the test points according to their distance from the data points of the training set. Given the training set features and a distance metric viz: Euclidean, city block, Mahalanobis, etc., to be used, the model assigns the test data to the class with the minimum distance. A k-nearest neighbor classifier considers k-data points for analysis.

5.5 Logistic regression

In this classifier, the predictive class labels are considered as binary dependent variables, and the relation to these with the features is modeled using a probability distribution for making predictions [26].

The performance of classification algorithms depends only on the historical data that the model learns and hence cannot predict the behavior and/or suitability of a classification algorithm to a specific problem. However, some results, based on the analysis of many datasets in a study that involved up to 7000 observations, 80 predictors, and 50 classes, to compare the classification algorithms used commonly in spoof analysis are listed in the Table 2 and some classifiers which serve this purpose are depicted in the Fig. 4 with their characteristics tabulated in Table 4.

6 Performance evaluation measures

Performance evaluation metrics define the accuracy of the recognition systems and are necessary to make quantitative assessments of the biometric system. The typical measures and graphics which effectively describe the overall performance of the biometric system used in face liveness detection are as follows:

6.1 False Positive Rate (FPR)

False Positive Rate (FPR) gives the percentage of the fake test samples incorrectly classified as belonging to the positive class. It is also referred to as False Acceptance Rate (FAR), False Match Rate (FMR), and False Genuine Rate (FGR). In applications that demand high security, FPR is expected to be very low. It is mostly expressed in percentage.

6.2 False Negative Rate (FNR)

False Negative Rate (FNR) means what percentage of the positive test samples incorrectly classified as belonging to the negative class. It can be noticed that both FPR and FNR counter each other, i.e. the requirement of low FPR may unavoidably increase FNR and vice versa. It is also termed as False Rejection Rate (FRR), False Non-Match Rate (FNMR), and False Non-Genuine Rate (FNGR) It is very often represented in percentage and this value should be as low as possible.

6.3 Half Total Error Rate (HTER)

HTER is the average of FPR and FNR.

Table 4 Characteristics of different classification algorithms

Classifier	Memory usage (in MB)	Prediction speed (in seconds)	Interpretability
SVM	4 for linear, 100 for others	1 for linear, 0.01 for others	Easy for linear, hard for other kernel types
Discriminant analysis	1 for linear, 100 for quadratic	100	Easy
Naive Bayes	1 for simple distributions, 4 for kernel distributions & high dimensional data	1 for simple distributions, 0.01 for kernel distributions & high dimensional data	Easy
Nearest neighbor	4	0.01 for cubic, 1 for others	Hard

6.4 Equal Error Rate (EER)

EER is the point when FPR matches FNR.

6.5 Receiver Operating Characteristic (ROC) curve

A ROC curve is a plot commonly used classification algorithms for exhibiting the performance of a classifier under different criteria. The x-axis of ROC is the FPR and the y-axis is the (True Positive Rate) TPR Any point on the curve implies the trade-off between the achieved TPR and the accepted FAR. ROC curves summarize the entire performance of a classifier and allow the comparison of different classifiers under similar conditions. The analysis of the curve assists in selecting a possibly optimal model and to discard suboptimal models from either the cost context or the class distribution independently before specifying. The best predictive model yields a point in the upper left corner, at the coordinate (0,1) of the ROC space, representing zero false negatives and false positives, and hence the (0,1) point is also called a perfect classification point. The Area Under the Curve (AUC) is also a commonly used parameter and has the range (0, 1) and 1 being the best. The relation between various quality measures are depicted in Fig. 5 below.

Since most of the considered data sets are not balanced (i.e., the number of impostors and genuine images is different), ACC may lead to biased performance analysis. All other

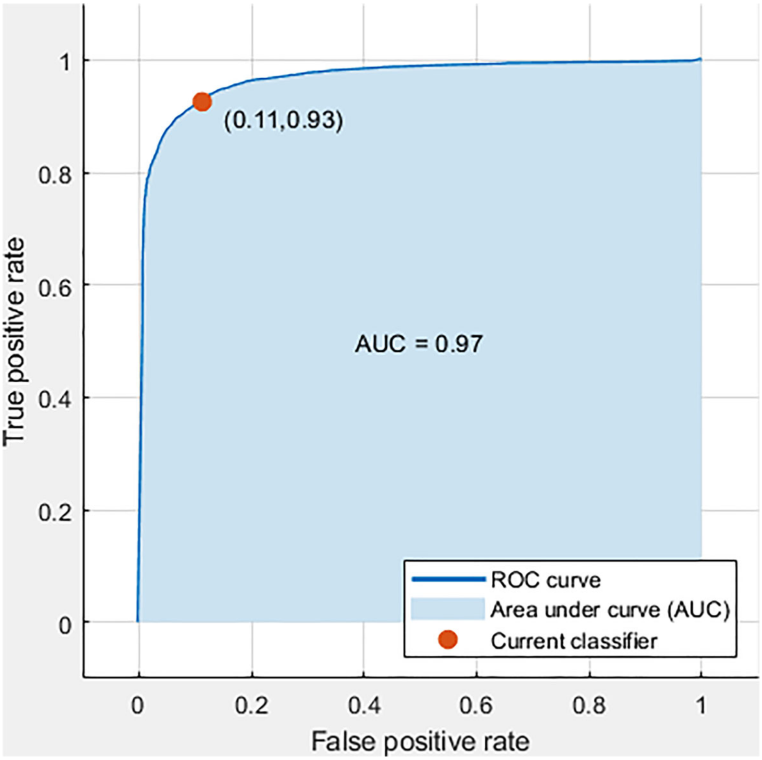

Fig. 5 Relation between the various quality measures used in spoof analysis

metrics are based on a separate evaluation of FAR and FRR, so they are more reliable for comparative analysis.

7 Existing modalities in spoof analysis: observations and discussions

The existing non-intrusive ML-based spoof detection methods are organized according to the classification algorithms employed along with some of the works exploiting fusion-based methods at the end with the overview of each type presented as observed in the literature. The first few paragraphs giving the details of the works based on SVM, then DA followed by NN classifiers and it's variants.

7.1 Non-intrusive feature based classification (ML) algorithms

Among most of the works, SVM and its variants are commonly used classifiers. The performance of the SVM classifier with the linear kernel is reported: in [1] on Print Attack, Replay Attack, NUAA Imposter and Yale Recaptured datasets. The authors also created their dataset which is described in Section 3. Experiments are carried on each dataset with 40% of the images as the training set and the remaining 60% as the test set. All the reported recordings are conducted on the Asus K52F Intel dual-core laptop with 3GB RAM using MATLAB, in [18], Variant Color Roberts Cross Operator (VCRCO) has been employed as a feature in HSV and YCbCr color spaces. The experiment is conducted on the CASIA Face Anti-Spoofing Database (FASD) and Replay Attack database. The authors reported that the proposed color gradient feature vector in HSV color space performed better than YCbCr color space and a considerable decrease in error rate (EER and HTER) has been observed with the fusion-based approach. The experiments are conducted using C++, in [37], random reflection characteristics of live images due to 3D shape is exploited using the diffusion process. The authors employed their proposed method on NUAA, SFL, and Replay Attack datasets, and evaluation of the experiments is done using accuracy and HTER. In [27], the authors used (i) (Scale Invariant Feature Transform) SIFT [14] (ii) uniform LBP [53] and (iii) the Gabor wavelet features at five scales and 8 orientations along with statistical moments for spoof analysis [55]. This framework is applied to Replay Attach, CASIA FASD, and NUAA datasets for analysis. Three SVM classifiers with linear and polynomial kernels, one for each feature, are used and the final decision of liveness is based on score level fusion of the scores from each classifier. The performance of the proposed approach is measured using AUC.

Among the three kernels used (linear, polynomial, and RBF), RBF is the most widely used kernel, and the works which reported acceptable results with this are as follows: In [74], IDA feature discussed in the Section 4 is used as a feature to distinguish between genuine and spoof images. Two classifiers are trained: one for the Replay Attack database and the other for the Print Attack database. In the testing stage, the input feature is fed to both the classifiers, and their outputs are fused using min and sum rules [29] to attain the final result and it has been identified that the min rule outperforms the sum rule. The performance of the proposed method has been tested using training and testing sets from the same database-intra dataset protocol. To evaluate the generalization ability of the proposed features, the authors also used the cross-database protocol in which training is carried out on one database and tested on other databases. The authors compared the performance of their proposed method with the two state-of-the-art methods: LBP [55] based and DoG-LBP based, using Replay Attack, CASIA FASD, and their own MSU MFSD databases. Here an ensemble classification method-different SVMs for different types of spoofs is used when training on a dataset with multiple types of spoofs and the results are analyzed using HTER and ROC performance evaluation metrics. Promising and better results have been observed for the proposed IDA based method compared to the other two methods mentioned above in the cross-database scenarios. The proposed approach is also tested for images with three different Inter-Pupillary Distance (IPD): 44, 60, and 70 pixels. The best performance has been noticed for images with an IPD of 70 pixels. The average time to extract the IDA feature for each frame is observed to be 0.12 s. The experiments are implemented using MATLAB.

In [61], the improved version of LBP, Multi-scale LBP (MLBP) [25], SIFT [53] and LUCID [81] are analyzed and LBP and color moments (mean, variance and skewness) are exploited as features for face liveness detection. The MSU USSA, CASIA FASD, Idiap Replay Attack, MSU MFSD, and RAFS [61] face spoof databases are considered for experimental analysis and the influence of different parameters such as image quality, size of the database, and color channels have been reported and the experiments are carried out using intra and cross-database protocols via LibSVM [13]. The performance has been evaluated using HTER and ROC characteristics have been analyzed. To reduce the processing time and memory space, the bezel detector has been used to reject some of the images with black stripes of the spoofing mediums before implementing the spoof detection algorithm and the results analyzed. The experiments are conducted on Intel 3 GHz quad-core CPU with 8 GB RAM on Windows 7 platform using MATLAB.

In [36], the authors exploited Local Binary Pattern (LBP) [58] and 2D Discrete Fourier Transform (DFT) using two databases: Webcam and ATM databases comprising live and spoofs printed on photographic sheets, A4 papers, magazines, and caricatures. Lastly, feature vector described by concatenating the decision scores obtained by SVM classifier [73] trained with frequency-based features and those trained with the LBP base feature vector is fed into an SVM classifier with RBF kernel again with parameters optimized using Genetic algorithm [34]. Here, all four types of spoofs are included in the training and testing sets such that their type is unknown. The performance of frequency-based, LBP based, and fusion-based methods are described using a ROC curve and EER has been reported for fusion-based methods. The performance of each type of mask using EER has been listed and analyzed. The overall performance of this method degrades when high-resolution printers which can capture the micro-texture are used and the frequency-based approach is inappropriate when 3D masks are presented.

In [2], the proposed liveness detection system uses either single or multiple descriptors based on the SL of the biometric system which is controlled by the user. LUCID [81], CENTRIST and (POEMs) [71], are used. The experiments are conducted on Print Attack and NUAA Imposter databases and Asus laptop with 3GB RAM and Intel dual-core CPU using MATLAB and the performance o classifier has been evaluated using FAR, FRR, EER, and HTER. In [54], specular reflection ratio and statistical measures of the GLCM of original and low pass filtered images are for the classification of the images of the NUAA dataset, and the recognition accuracy is listed along with previously proposed methods for comparative analysis. In [79], the authors used distortion dependent pixel similarity deviation of a mean subtracted and contrast normalized as features and evaluated their work on CASIA, Replay Attack, and UVAD datasets [63] and quantitative measurements are performed using HTER and Relative ER (RER). The authors also make a comparison using inter dataset protocols. In [18], the gradient of all the channels of an image is computed in HSV and YCbCr color spaces and are concatenated to form a feature for face spoof analysis on CASIA and

Replay Attack datasets using intra dataset protocol. The performance is evaluated using HTER and EER.

Some of the works are reported promising results with DA classifiers which are described in brief below In [20], the proposed nine pixel difference based measures, three correlation-based measures, and two edge-based measures have been used to generate feature vector and presented to a Linear Discriminant Analysis (LDA) classifier to assign one of the two classes: live or spoof. The authors report the same HTER as that obtained in [14] with the Replay Attack dataset but the speed and simplicity of the IQ (Image Quality) measures make it more suitable for real-time applications. Also, the results of IQA based method is compared with DoG filter based liveness detection under seven different attacking scenarios and better performance is observed that the proposed method outperforms DoG based approach for acquisition sensors of higher resolution. The accuracy of the classifier decreases when high-quality spoofing attacks are presented. The experiments are implemented using MATLAB. In [31], authors extracted statistical textural features and evaluated their framework on their database of 25 subjects and measured the performance using accuracy, sensitivity, and specificity. Here, Quadratic Discriminant Analysis (QDA) is used for classification.

The k-NN based liveness detection is employed in [78]. The authors consider only the nose (central patch) and cheek (right bottom corner) parts of the image. The authors highlight the difference in the gradient magnitude of the original and Gaussian blurred versions of both the live and spoof image parts. The classification is based on the fact that the sum of all the pixels of the difference image will be larger for live image parts due to the prominent gradients of the live images than those of recaptured face images. Here, the k-NN algorithm is used for classification. The authors have created their dataset for evaluating their framework.

Recent works use CNN which can accept even raw images (with background) and make acceptable predictions. These models can also be used as classifiers. In [43], spectral energy density as a function of time is used as a feature for a unique representation of images preprocessed for illumination correction, and a CNN is used for classification. The proposed method is evaluated on CASIA, Replay Attack datasets, and HTER is used for comparative analysis with other works for both inter and intra datasets. In [3], the authors used modified nonlinear or anisotropic diffusion is used for preserving edges of the images, and a five-layer Convolutional Neural Network (CNN) is used to detect the edges of the shapes for spoof detection. The authors used the Replay Attack dataset to analyze the performance of their approach and used HTER for quantitative measurement. In [32] Eigenvalues of the covariance matrix are used as features for spoof analysis. In [33], constrained CNN architecture is used to decompose the spoof face into a spoof noise and a live face, and then utilizing the spoof noise for classification. In [6], randomly extracted LBP texture features with depth maps are used.

7.2 Intrusive combinational modalities for spoof analysis

In [38], the detection of spoofing attacks has been performed based on the decision of two independent methods: one analyzing 3D properties of the head and one the eye-blinking of the user. In [70], the authors develop a fusion scheme at a frame level and apply it to a set of visual appearance cues. In [66], this fusion is done at the feature level. In [76], bring the intuition that the fusion can have a bigger impact if done with complementary counter-measures, that address different spoofing attack cues. In a particular case, a method based on motion analysis is fused with a method based on visual appearance. To measure the level of independence of two anti-spoofing systems, and thus to get an estimation of

Table 5 Overview of different modalities used in spoof analysis

Si. No.	Kernel/Classifier type		Database	Reference	Remarks
1.	SVM	Linear	Print Attack,Replay Attack, NUAA, Yale, Their own	[64]	LUCID effectively captures the differences in reflective properties in the uniform face regions such as cheeks and performed better for LCD spoof
			CASIA, Replay Attack	[7]	Roberts cross operator is used which performs better in HSV than YCbCr, decrease in error rate is observed with fusion based approach.
			NUAA, SFL, Replay Attack	[14]	Randomly distributed illumination energies in live face images diffuse faster compared to spoof face images and the approach performed for NUAA and Local Speed Patterns database compared to SFL dataset with print and LCD spoofs.
		Polynomial	Replay Attack, CASIA, NUAA	[53]	SIFT, uniform LBP, Gabor wavelets and statistical moments are used used to train SVMs with linear and polynomial kernels and polynomial kernel performed better.
		RBF	Replay Attack,Print Attack	[48]	IDA uses specular reflection, blurriness and color moments and the performance degrades when high quality images are used and the background information can decrease the performance.
			MSU USSA, CASIA, Replay Attack, MSU MFSD & RAFS	[74]	LBP and color moments are used along with bezel detector to reject the spoof images with borders and hence to increase the performance efficiency and reduce the computational costs incurred.
			BERC Webcam & ATM	[1]	LBP captures the micro texture differences and DFT captures both low frequency and high frequency detail. Performance degrades when high resolution printers used.
			Print Attack, NUAA	[35]	Takes security information from user and uses three features: LUCID, CENTRIST and POEM, when lower security is encountered
		*	NUAA	[57]	Specular reflection and statistical measures of GLCM are used.

Table 5 (continued)

Si. No.	Kernel/Classifier type		Database	Reference	Remarks
			Replay Attack, UVAD	[23]	Distortion dependent pixel similarity deviation is used.
			CASIA, Replay Attack	[7]	Gradient in HSV and YCbCr color spaces are used.
2.	DA	Linear	Replay Attack	[17]	Uses difference and correlation based edge measures, obtained same HTER in but outperformed DoG based approach.
		Quadratic	Their own	[57]	Statistical texture features are used.
3.	k-NN			[10]	Gradient magnitude is used to capture the difference in edges of live and spoof images.
4.	CNN		CASIA, Replay Attack	[29]	Evolution of spectral energy density is used. Live images contain most of their energy at low frequencies.
				[25]	Anisotropic diffusion is used.
				[13]	Eigen values of covariance matrix are exploited as features.

* indicates that the type of the classifier is not specified in the stated papers.

their complementarity and effectiveness of their fusion, [38] propose employing a statistical analysis based on [40]. Komulainen et al. [39] showed that score-level fusion of several simple anti-spoofing methods that do not involve complex inefficient classifiers may be favorable concerning a single one requiring time and memory. The summary of the above reviewed works is summarized in Table 5.

Most of the works in the literature presented their results using different performance metrics such as ACC, AUC, HTER, and EER and reported acceptable results for each of the datasets they considered for their work overlooking the performance of the model for the unknown attacks. Far from showing that face spoofing detection is a solved problem, this fact indicates the lack of a challenging data set that allows a thorough analysis of the proposed methods. We believe that a large data set with the worst scenarios is more likely to promote breakthroughs. However, there are works that investigate the effect of unknown spoofing attacks [52] In addition to a large number of images and/or videos, multiple types of attacks should be covered, be diverse in terms of ethnicity, age and gender, and present real-world scenarios with different environments, acquisition devices, lighting conditions, and human behaviors.

Comparing different works is a difficult task since most of the time we do not have access to the source codes, and reproducing codes and experimental results are very complicated. However, the determination of the best method based on the reported results is not an easy task. It is possible to make mistakes even when comparing works that use the very same data set. Besides a commonly available data set, it is of underlying importance to follow the same methodology and to have the same metrics when comparing different countermeasures.

8 Conclusions and scope of research

This survey introduces the various steps involved in the face spoof detection process, the limitations associated with various intrusive and non-intrusive methods which emphasize either a single frame or multi-frame based approaches on various types of spoofing attacks: print spoofs, display and replay attacks, etc. A comprehensive study on different features used to discern live and spoof images along with their characterization is presented. The various classification algorithms which can serve the purpose of binary classification for spoof analysis are discussed. Finally, the survey ends with a detailed description of different modalities employed in spoof analysis over a few decades.

Although different spoof detection methods present promising results on the available datasets, the datasets used do not contain all the worst-case scenarios and the existing methods fail to correctly classify the test data when they encounter new cases. Also, the performance accuracy depends not only on the features used but also on the number of training samples and restrictions on the parameters of the classification models. Hence there is a need to: (i) construct a large dataset with worst-case scenarios and (ii) construct robust features that can exploit the characteristics of different spoofing attacks and are affine invariant and robust to illumination variations. Lastly, the spoof detection problem can also be viewed as a multiclass classification problem rather than binary classification so that spoofing attacks can be classified according to their characteristics. Though challenges in economic spoof detection persist, the work presented is based on different papers available in the literature. This survey has been organized and presented to provide a proper perception on spoof detection ranging from types, modalities, countermeasures, classification algorithms, and performance measures along with merits and limitations observed and mainly to explore the possibility of the spoof detection system with the database created using commonly available resources for acquisition, not following any presumed database testing protocols to make it deployable in real-time applications with minimum user intervention and computational cost and since most of the available techniques are tailored to work with known attacks to learn the decision framework making it impossible to predict how they perform for unknown attacks the performance achieved by state-of-the-art methods for unknown attacks is far from their application to real-life cases. Looking into this survey, there is an exclusive requirement at the resources and acquisition side. However, this survey is a prelude to find directions for further work addressing (i) commonly available acquisition devices and natural environments (ii) exploring very effective features to address these circumstances. To begin with, we are restricted to spoof detection addressing print spoofs.

References

1. Akhtar Z, Michelon C, Foresti GL (2014) Liveness detection for biometric authentication in mobile applications. In: 2014 International Carnahan conference on security technology (ICCST). IEEE, pp 1–6
2. Akhtar Z, Micheloni C, Piciarelli C, Foresti GL (2014) Mobio_livdet: mobile biometric liveness detection. In: 2014 11th IEEE international conference on advanced video and signal based surveillance (AVSS). IEEE, pp 187–192
3. Alotaibi A, Mahmood A (2016) Enhancing computer vision to detect face spoofing attack utilizing a single frame from a replay video attack using deep learning. In: 2016 International conference on optoelectronics and image processing (ICOIP). IEEE, pp 1–5
4. Amin R, Islam SH, Biswas G, Khan MK, Leng L, Kumar N (2016) Design of an anonymity-preserving three-factor authenticated key exchange protocol for wireless sensor networks. Comput Netw 101:42–62

5. Angadi SA, Kagawade VC (2018) Detection of face spoofing using multiple texture descriptors. In: 2018 International conference on computational techniques, electronics and mechanical systems (CTEMS). IEEE, pp 151–156
6. Atoum Y, Liu Y, Jourabloo A, Liu X (2017) Face anti-spoofing using patch and depth-based cnns. In: 2017 IEEE international joint conference on biometrics (IJCB). IEEE, pp 319–328
7. Bai J, Ng T-T, Gao X, Shi Y-Q (2010) Is physics-based liveness detection truly possible with a single image? In: Proceedings of 2010 IEEE international symposium on circuits and systems. IEEE, pp 3425–3428
8. Bao W, Li H, Li N, Jiang W (2009) A liveness detection method for face recognition based on optical flow field. In: 2009 International conference on image analysis and signal processing. IEEE, pp 233–236
9. Basri R, Jacobs DW (2003) Lambertian reflectance and linear subspaces. IEEE Trans Pattern Anal Mach Intell 25(2):218–233
10. Benlamoudi A, Samai D, Ouafi A, Bekhouche SE, Taleb-Ahmed A, Hadid A (2015) Face spoofing detection using local binary patterns and fisher score. In: 2015 3rd International conference on control, engineering & information technology (CEIT). IEEE, pp 1–5
11. Bhogal APS, Söllinger D, Trung P, Uhl A (2017) Non-reference image quality assessment for biometric presentation attack detection. In: 2017 5th International workshop on biometrics and forensics (IWBF). IEEE, pp 1–6
12. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
13. Chang C-C, Lin C-J (2011) Libsvm: a library for support vector machines. ACM Trans Intell Syst Technol (TIST) 2(3):1–27
14. Chingovska I, Anjos A, Marcel S (2012) On the effectiveness of local binary patterns in face anti-spoofing. In: 2012 BIOSIG-proceedings of the international conference of biometrics special interest group (BIOSIG). IEEE, pp 1–7
15. Crete F, Dolmiere T, Ladret P, Nicolas M (2007) The blur effect: perception and estimation with a new no-reference perceptual blur metric. In: Human vision and electronic imaging XII, vol 6492. International Society for Optics and Photonics, p 64920I
16. Cristianini N, Shawe-Taylor J et al (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press, Cambridge
17. Dhawanpatil T, Joglekar B (2017) Face spoofing detection using multiscale local binary pattern approach. In: 2017 International conference on computing, communication, control and automation (ICCUBEA). IEEE, pp 1–5
18. Dong J, Tian C, Xu Y (2017) Face liveness detection using color gradient features. In: 2017 International conference on security, pattern analysis, and cybernetics (SPAC). IEEE, pp 377–382
19. Fisher RA (1936) The use of multiple measurements in taxonomic problems. Ann Eugenics 7(2):179–188
20. Galbally J, Marcel S (2014) Face anti-spoofing based on general image quality assessment. In: 2014 22nd International conference on pattern recognition. IEEE, pp 1173–1178
21. Galbally J, Marcel S, Fierrez J (2013) Image quality assessment for fake biometric detection: application to iris, fingerprint, and face recognition. IEEE Trans Image Process 23(2):710–724
22. Gao X, Ng T-T, Qiu B, Chang S-F (2010) Single-view recaptured image detection based on physics-based features. In: 2010 IEEE international conference on multimedia and expo. IEEE, pp 1469–1474
23. Garcia DC, de Queiroz RL (2015) Face-spoofing 2d-detection based on moiré-pattern analysis. IEEE Trans Pattern Anal Mach Intell 10(4):778–786
24. Garud D, Agrwal S (2016) Face liveness detection. In: 2016 International conference on automatic control and dynamic optimization techniques (ICACDOT). IEEE, pp 789–792
25. Han H, Klare BF, Bonnen K, Jain AK (2012) Matching composite sketches to face photos: a component-based approach. IEEE Trans Inf Forensics Secur 8(1):191–204
26. Härdle W, Simar L (2007) Applied multivariate statistical analysis, vol 22007. Springer, Berlin
27. Hassan MA, Mustafa MN, Wahba A (2017) Automatic liveness detection for facial images. In: 2017 12th International conference on computer engineering and systems (ICCES). IEEE, pp 215–220
28. Hsu C-W, Chang C-C et al (2003) A practical guide to support vector classification. Tech. rep., Department of Computer Science, National Taiwan University
29. Jain A, Nandakumar K, Ross A (2005) Score normalization in multimodal biometric systems. Pattern Recognit 38(12):2270–2285
30. Jan V, Drahanskỳ M, Dvor R, Yanushkevich SN et al (2012) Thermal face recognition: a fusion approach. In: 2012 Third international conference on emerging security technologies. IEEE, pp 39–42
31. Jayan TJ, Aneesh R (2018) Image quality measures based face spoofing detection algorithm for online social media. In: 2018 International CET conference on control, communication, and computing (IC4). IEEE, pp 245–249

32. Jiang C, Chen S, Zhang B, Chen Y, Bo Y, Feng Z (2018) Effectiveness analysis of the covariance matrix for spoofing detection application. In: 2018 Ubiquitous positioning, indoor navigation and location-based services (UPINLBS). IEEE, pp 1–5
33. Jourabloo A, Liu Y, Liu X (2018) Face de-spoofing: anti-spoofing via noise modeling. In: Proceedings of the European conference on computer vision (ECCV), pp 290–306
34. Jung HG, Kim J (2010) Constructing a pedestrian recognition system with a public open database, without the necessity of re-training: an experimental study. Pattern Anal Appl 13(2):223–233
35. Kim JK, Park HW (1999) Statistical textural features for detection of microcalcifications in digitized mammograms. IEEE Trans Med Imaging 18(3):231–238
36. Kim G, Eum S, Suhr JK, Kim DI, Park KR, Kim J (2012) Face liveness detection based on texture and frequency analyses. In: 2012 5th IAPR international conference on biometrics (ICB). IEEE, pp 67–72
37. Kim W, Suh S, Han J-J (2015) Face liveness detection from a single image via diffusion speed model. IEEE Trans Image Process 24(8):2456–2465
38. Kollreider K, Fronthaler H, Bigun J (2008) Verifying liveness by multiple experts in face biometrics. In: 2008 IEEE Computer Society conference on computer vision and pattern recognition workshops. IEEE, pp 1–6
39. Komulainen J, Hadid A, Pietikäinen M, Anjos A, Marcel S (2013) Complementary countermeasures for detecting scenic face spoofing attacks. In: 2013 International conference on biometrics (ICB). IEEE, pp 1–7
40. Kuncheva LI, Whitaker CJ (2003) Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. Mach Learn 51(2):181–207
41. Lagorio A, Tistarelli M, Cadoni M, Fookes C, Sridharan S (2013) Liveness detection based on 3d face shape analysis. In: 2013 International workshop on biometrics and forensics (IWBF). IEEE, pp 1–4
42. Lai C-L, Chen J-H, Hsu J-Y, Chu C-H (2013) Spoofing face detection based on spatial and temporal features analysis. In: 2013 IEEE 2nd global conference on consumer electronics (GCCE). IEEE, pp 301–302
43. Lakshminarayana NN, Narayan N, Napp N, Setlur S, Govindaraju V (2017) A discriminative spatio-temporal mapping of face for liveness detection. In: 2017 IEEE international conference on identity, security and behavior analysis (ISBA). IEEE, pp 1–7
44. Leng L, Zhang J, Khan MK, Chen X, Alghathbar K (2010) Dynamic weighted discrimination power analysis: a novel approach for face and palmprint recognition in dct domain. Int J Phys Sci 5(17):2543–2554
45. Leng L, Li M, Teoh ABJ (2013) Conjugate 2dpalmhash code for secure palm-print-vein verification. In: 2013 6th International congress on image and signal processing (CISP), vol 3. IEEE, pp 1705–1710
46. Leng L, Teoh ABJ, Li M, Khan MK (2014) A remote cancelable palmprint authentication protocol based on multi-directional two-dimensional palmphasor-fusion. Secur Commun Netw 7(11):1860–1871
47. Leng L, Li M, Kim C, Bi X (2017) Dual-source discrimination power analysis for multi-instance contactless palmprint recognition. Multimed Tools Appl 76(1):333–354
48. Li J, Wang Y, Tan T, Jain AK (2004) Live face detection based on the analysis of fourier spectra. In: Biometric technology for human identification, vol 5404. International Society for Optics and Photonics, pp 296–303
49. Li Y, Po L-M, Xu X, Feng L, Yuan F (2016) Face liveness detection and recognition using shearlet based feature descriptors. In: 2016 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, pp 874–877
50. Liu W (2014) Face liveness detection using analysis of fourier spectra based on hair. In: 2014 International conference on wavelet analysis and pattern recognition. IEEE, pp 75–80
51. Liu X, Lu R, Liu W (2017) Face liveness detection based on enhanced local binary patterns. In: 2017 Chinese automation congress (CAC). IEEE, pp 6301–6305
52. Liu Y, Stehouwer J, Jourabloo A, Liu X (2019) Deep tree learning for zero-shot face anti-spoofing. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4680–4689
53. Lowe DG (1999) Object recognition from local scale-invariant features. In: Proceedings of the seventh IEEE international conference on computer vision, vol 2. IEEE, pp 1150–1157
54. Luan X, Wang H, Ou W, Liu L (2017) Face liveness detection with recaptured feature extraction. In: 2017 International conference on security, pattern analysis, and cybernetics (SPAC). IEEE, pp 429–432
55. 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). IEEE, pp 1–7
56. Marsland S (2015) Machine learning: An algorithmic perspective. CRC Press
57. Marziliano P, Dufaux F, Winkler S, Ebrahimi T (2002) A no-reference perceptual blur metric. In: Proceedings. International conference on image processing, vol 3. IEEE, pp III–III

58. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987
59. Okereafor K, Onime C, Osuagwu O (2017) Enhancing biometric liveness detection using trait randomization technique. In: 2017 UKSim-AMSS 19th international conference on computer modelling & simulation (UKSim). IEEE, pp 28–33
60. Pan G, Sun L, Wu Z, Lao S (2007) Eyeblink-based anti-spoofing in face recognition from a generic webcamera. In: 2007 IEEE 11th international conference on computer vision. IEEE, pp 1–8
61. Patel K, Han H, Jain AK (2016) Secure face unlock: spoof detection on smartphones. IEEE Trans Pattern Anal Mach Intell 11(10):2268–2283
62. Peixoto B, Michelassi C, Rocha A (2011) Face liveness detection under bad illumination conditions. In: 2011 18th IEEE international conference on image processing. IEEE, pp 3557–3560
63. Pinto A, Schwartz WR, Pedrini H, de Rezende Rocha A (2015) Using visual rhythms for detecting video-based facial spoof attacks. IEEE Trans Inf Forensics Secur 10(5):1025–1038
64. Ramachandra R, Busch C (2017) Presentation attack detection methods for face recognition systems: a comprehensive survey. ACM Comput Surv (CSUR) 50(1):1–37
65. Schütze H, Manning CD, Raghavan P (2008) Introduction to information retrieval, vol 39. Cambridge University Press, Cambridge
66. Schwartz WR, Rocha A, Pedrini H (2011) Face spoofing detection through partial least squares and low-level descriptors. In: 2011 International joint conference on biometrics (IJCB). IEEE, pp 1–8
67. Sun L, Pan G, Wu Z, Lao S (2007) Blinking-based live face detection using conditional random fields. In: International conference on biometrics. Springer, pp 252–260
68. Tan X, Li Y, Liu J, Jiang L (2010) Face liveness detection from a single image with sparse low rank bilinear discriminative model. In: European conference on computer vision. Springer, pp 504–517
69. Tronci R, Giacinto G, Roli F (2009) Dynamic score combination: a supervised and unsupervised score combination method. In: International workshop on machine learning and data mining in pattern recognition. Springer, pp 163–177
70. Tronci R, Muntoni D, Fadda G, Pili M, Sirena N, Murgia G, Ristori M, Ricerche S, Roli F (2011) Fusion of multiple clues for photo-attack detection in face recognition systems. In: 2011 International joint conference on biometrics (IJCB). IEEE, pp 1–6
71. Vu N-S, Caplier A (2010) Face recognition with patterns of oriented edge magnitudes. In: European conference on computer vision. Springer, pp 313–326
72. Wang D, Hoi SC, He Y, Zhu J, Mei T, Luo J (2013) Retrieval-based face annotation by weak label regularized local coordinate coding. IEEE Trans Pattern Anal Mach Intell 36(3):550–563
73. Waske B, Benediktsson JA (2007) Fusion of support vector machines for classification of multisensor data. IEEE Trans Geosci Remote Sens 45(12):3858–3866
74. Wen D, Han H, Jain AK (2015) Face spoof detection with image distortion analysis. IEEE Trans Inf Forensics Secur 10(4):746–761
75. Wu J, Rehg JM (2010) Centrist: a visual descriptor for scene categorization. IEEE Trans Pattern Anal Mach Intell 33(8):1489–1501
76. Yan J, Zhang Z, Lei Z, Yi D, Li SZ (2012) Face liveness detection by exploring multiple scenic clues. In: 2012 12th International conference on control automation robotics & vision (ICARCV). IEEE, pp 188–193
77. Yang L (2014) Face liveness detection by focusing on frontal faces and image backgrounds. In: 2014 International conference on wavelet analysis and pattern recognition. IEEE, pp 93–97
78. Yeh C-H, Chang H-H (2017) Face liveness detection with feature discrimination between sharpness and blurriness. In: 2017 Fifteenth IAPR international conference on machine vision applications (MVA). IEEE, pp 398–401
79. Yeh C-H, Chang H-H (2018) Face liveness detection based on perceptual image quality assessment features with multi-scale analysis. In: 2018 IEEE Winter conference on applications of computer vision (WACV). IEEE, pp 49–56
80. Zhang Z, Yan J, Liu S, Lei Z, Yi D, Li SZ (2012) A face antispoofing database with diverse attacks. In: 2012 5th IAPR international conference on Biometrics (ICB). IEEE, pp 26–31
81. Ziegler A, Christiansen E, Kriegman D, Belongie SJ (2012) Locally uniform comparison image descriptor. In: Advances in neural information processing systems, pp 1–9