Engineering Applications of Artificial Intelligence 72 (2018) 368–381

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

# How far did we get in face spoofing detection?

Luiz Souza [a], Luciano Oliveira [a,*], Mauricio Pamplona [a], Joao Papa [b]

[a] IVISION Lab, Federal University of Bahia, Brazil
[b] RECOGNA Lab, São Paulo State University, Brazil

ARTICLE INFO

*Keywords:*
Face spoofing
Face recognition
Survey
Spoofing attack

ABSTRACT

The growing use of control access systems based on face recognition shed light over the need for even more accurate systems to detect face spoofing attacks. In this paper, an extensive analysis on face spoofing detection works published in the last decade is presented. The analyzed works are categorized by their fundamental parts, *i.e.,* descriptors and classifiers. This structured survey also brings a comparative performance analysis of the works considering the most important public data sets in the field. The methodology followed in this work is particularly relevant to observe temporal evolution of the field, trends in the existing approaches, to discuss still opened issues, and to propose new perspectives for the future of face spoofing detection.

## 1. Introduction

In the last decade, there has been an increasing interest in human automatic secure identification, being mainly based on unique personal biometric information (Jain et al., 2008). One of the main reasons for such focus concerns the high number of security breaches and transaction frauds in non-biometric systems, which are prone to be cracked due to inherent vulnerabilities (Meadowcroft, 2008), like stolen cards and shared passwords, just to name a few.

Biometrics may use physical or behavioral characteristics for identification purposes, and different alternatives have been explored over the years: fingerprint (Hasan and Abdul-Kareem, 2013; Marasco and Ross, 2015; Peralta et al., 2014), hand geometry (Al Eidan, 2013; Kah Ong Michael et al., 2012), palmprint (Tamrakar and Khanna, 2016), voice (Yadav and Mukhedkar, 2013; Choi et al., 2015), face (Zhao et al., 2003; Feng et al., 2016; Dora et al., 2017), and handwritten signature (Sanmorino and Yazid, 2012). Among those, face stands out for its acceptability and recognition cost, turning out to be one of the best option for a wide range of applications, from low-security uses (*e.g.,* social media and smartphone access control) to high-security applications (*e.g.,* border control and video surveillance in critical places).

This popularity, however, comes with a price: face recognition systems have become a major target of spoofing attacks. In such scenarios, an impostor attempts to be granted in an identification process by forging someone else's identity. As procedures to replicate human faces are very much standard nowadays (*e.g.,* photo and 3D printing), spoofing detection has become mandatory in any suitable face recognition system. Fig. 1 illustrates the complexity of this problem, and the following question can be raised: "Which half is real or fake?". It is sometimes a very challenging task, even for humans.

Several approaches for spoofing detection have been developed in the last decade. Recently, two main surveys on the subject present a comprehensive review (Galbally et al., 2014; Parveen et al., 2015): in Galbally et al. (2014), a survey on anti-spoofing methods focuses not only on face, but also on other biometric traits (*e.g.,* iris, voice, fingerprint); in Parveen et al. (2015), face anti-spoofing methods are discussed by considering the intrusiveness of each method, with few attention on comparative analysis and temporal evolution of the field. On the other hand, the proposed survey focuses only on face-oriented works, reviewing and analyzing the most relevant works on face spoofing detection in the literature towards depicting the advance of the detection methods in the last decade. An extensive set of face anti-spoofing methods is presented, also depicting the evolution of the existing works. In this sense, trends denoted throughout these years were pointed out, as well as open issues were remarked in order to provide new directions on research topics in the future. Next, the contributions of this survey are addressed and discussed in details with respect to the other existing surveys, with special attention to the gaps filled by the present work.

### 1.1. Contributions

To the best of our knowledge, there are only two surveys in the context of face spoofing detection (Galbally et al., 2014; Parveen et al.,

[*] Corresponding author.
*E-mail addresses:* luiz.otavio@ufba.br (L. Souza), lrebouca@ufba.br (L. Oliveira), mauricio@dcc.ufba.br (M. Pamplona), papa@fc.unesp.br (J. Papa).

https://doi.org/10.1016/j.engappai.2018.04.013
Received 17 June 2017; Received in revised form 9 April 2018; Accepted 13 April 2018
Available online 8 May 2018

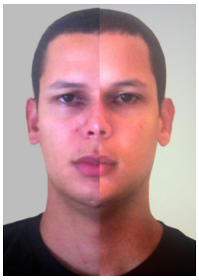

**Fig. 1.** Example of a half real (photo) and half fake face (photo of a photo). Which half is the real one? The answer is the one on the left.

2015). Although two face anti-spoofing competitions were organized (Chakka, 2011; Chingovska, 2013), and several data sets and methods have been published, the amount of gathered data and results were not still thorough and critically analyzed so far. Even these two existing surveys do not concentrate efforts to understanding the trends of this research field in terms of conception of the methods and results.

Galbally et al. (2014) published a survey based on a chronological evolution of multimodal anti-spoofing methods. Although a special attention was given to face anti-spoofing, other biometric traits were also presented and discussed. A proposed timeline takes into consideration fingerprint, iris, and face anti-spoofing detection competitions, being the latter one organized by one of the authors of the survey (Galbally et al., 2014). In regard to face-driven works, the authors provided an extensive and comprehensive description of different types of face attacks and public image data sets. The face anti-spoofing methods, categorized by Galbally et al., were according to three levels: sensor, features, and multi-modal fusion, but being only two levels employed to classify the analyzed works. Sixteen existing works compose the face study part, which was characterized by the level of the technique, type of attack, public image data set used, and a single error rate. At the end, a discussion was addressed showing that although competitive laboratory performances were achieved, some people were successfully able to hack the fingerprint recognition system of the Iphone 5s. In Galbally et al. (2014), also, some discussion about performance of face anti-spoofing methods resided in general considerations about cross-data set performance evaluation (in order to turn methods' evaluation more thoroughly accomplished), new relevant features acquired on facial blood flow, and new hardware that could be used along with cameras to improve face anti-spoofing detection. The remainder of the survey in Galbally et al. (2014) discusses philosophical aspects of performing an anti-spoofing detection approach within face recognition systems.

Parveen et al. (2015) followed a general architecture comprised of a sensor, pre-processing, feature extraction, and classification steps as a basis for a taxonomy of face anti-spoofing detection methods. The methods are categorized as non-intrusive or intrusive ones, addressed according to the stillness or motion detection presented in the detection process, respectively. Twenty-nine face anti-spoofing methods were studied, and the results of the existing works were individually analyzed over public image data sets. An experimental analysis was carried out by means of four error measures: *half total error rate* (HTER), *equal error rate* (EER), *area under curve* (AUC) and *accuracy* (ACC). At the end in Parveen et al. (2015), some pros and cons are highlighted with regard to implementation complexity, user collaboration and attack coverage.

Differently from Galbally et al. (2014), which spread out the discussion on various anti-spoofing methods using different traits, we present an extensive survey that is focused on the evolution of particularly face spoofing detection methods and existing benchmarks. Instead of following a more generic categorization as those proposed in Galbally et al. (2014) and Parveen et al. (2015), all gathered works here were organized in terms of their main component parts, *i.e.,* descriptors and classifiers (see Section 2). This taxonomy was devised to help the reader to better understand the processes behind each countermeasure, and to unveil technical trends concerning different types of attacks. Since all works comprise features and learning methods, this organization seems to be the best to depict a big picture of the state-of-the-art research related to face spoofing detection.

Despite the other two surveys, our work resorts to a quantitative and analytical methodology (see Section 3) in order to support the analysis of trends of the existing face anti-spoofing approaches (see Section 4). A comparison of several methods was accomplished over the most currently used public data sets, taking into account the bias of the metrics used to assess face anti-spoofing performance (with several perfect results), differently from Galbally et al. (2014) and Parveen et al. (2015), where the results were individually analyzed. The goal is to numerically show how far spoofing detectors got considering only face. In order to fulfill such purposes, sixty-one face anti-spoofing methods were gathered (including the works that participated in the two competitions). Previous surveys did not include any in-depth assessment of existing face spoofing detection approaches (Galbally et al., 2014; Parveen et al., 2015), leaving unclear which ways should be followed and what need to be done in technical terms, considering only face spoofing detection. Differently from the philosophical and general discussion found in Galbally et al. (2014), concerning facts and challenges in the spoofing detection domain, the numerical-driven evaluation of the area allows suggesting other ways to evaluate the performance (avoiding supposedly perfect results), as well as new future research topics (*e.g.*, deep learning Fan et al., 2014, and collaborative clustering Cornujols et al., 2018) to be applied in face anti-spoofing methods (see Section 4).

### *1.2. Methodology*

This compilation of works is based on a literature search in the following data sets: Scopus,[1] IEEE Xplore,[2] Engineering Village,[3] and Google Scholar[4] On these sources, articles were consulted considering all publications with the following keywords: **face recognition, face spoofing detection, face liveness detection, countermeasure against spoofing attacks and face anti-spoofing detection methods**. The choice of the articles was made according to the following criteria: (i) they should follow the same protocol when evaluating the study; (ii) they should indicate the results using at least one of the metrics discussed in Section 3.2, (iii) they should be comparable to other studies using the same data set, and finally (iv) they must be peer-reviewed.

It is noteworthy that there were two competitions on face spoofing detection referred in Chakka (2011), Chingovska (2013). The results obtained by the competition teams were analyzed, and the names of the groups and universities were used as references to the methods used in the first face spoofing detection competition, such as: Ambient Intelligence Laboratory (AMILAB), Center for Biometrics and Security Research, Institute of Automation, Chinese Academy of Sciences (CASIA), Idiap Research Institute (IDIAP), Institute of Intelligent Systems and Numerical Applications in Engineering, *Universidad de Las Palmas de Gran Canaria* (SIANI), Institute of Computing, Campinas University (UNICAMP) and Machine Vision Group, University of Uolu (UOLU) (Chakka, 2011). As well as, the names CASIA, Fraunhofer Institute for Computer Graphics Research (IGD), joint team from IDIAP, UOLU, UNICAMP and CPqD Telecom & IT Solutions (MaskDown), the LNM Institute of Information Technology, Jaipur (LNMIIT), Tampere University of Technology (MUVIS), University of Cagliari (PRA Lab), *Universidad Autonoma de Madrid* (ATVS) and UNICAMP refer to the teams that participated in the second face spoofing detection competition (Chingovska, 2013). Throughout this text, these team names will be cited as the reference of the method in the competition (Chakka, 2011 or Chingovska, 2013).

[1] http://www.scopus.com/.

[2] http://ieeexplore.ieee.org/Xplore/home.jsp.

[3] http://www.engineeringvillage.com/.

[4] https://scholar.google.com.

## 2. Face spoofing detection

Face spoofing detection (Maatta et al., 2011, 2012; Schwartz et al., 2011; Bharadwaj et al., 2013; Tirunagari et al., 2015), face liveness detection (Yan et al., 2012; Peixoto et al., 2011; Yang et al., 2013; Wang et al., 2013; Tan et al., 2010), counter measure against facial spoofing attacks (Komulainen et al., 2013b; Pereira et al., 2012; Kose and Dugelay, 2013c, a, b), and face anti-spoofing (Chingovska et al., 2012; Erdogmus and Marcel, 2013; Galbally and Marcel, 2014) are all terms interchangeably used to denote methods to identify an impostor trying to masquerade him/herself as a genuine user in facial recognition systems.

### 2.1. Types of face spoofing attacks

Face spoofing systems usually consider the following types of spoofing attacks:

- The use of a **flat printed photo** (see Fig. 2(b)) is the most common one, with great potential to take place, since most people have facial pictures available on the Internet (*e.g.,* social media) or could be photographed by an impostor without collaboration or permission.
- In the **eye-cut photo** attack, eye regions of a printed photo are cut off to exhibit blink behavior of the impostor (see Fig. 2(c)).
- **Warped photo** attacks consist in bending a printed photo in any direction to simulate facial motion (see Fig. 2(d)).
- An attack via **video playback** shows almost all behaviors similar to real faces, with many of the intrinsic features of valid user movements (see Fig. 2(e)). 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.
- **Mask** attacks are of two types: life-size wearable mask (see Fig. 2(f)) and paper-cut mask (see Fig. 2(g)). These attacks are addressed to anti-spoofing systems that analyze 3D face structures, being 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.

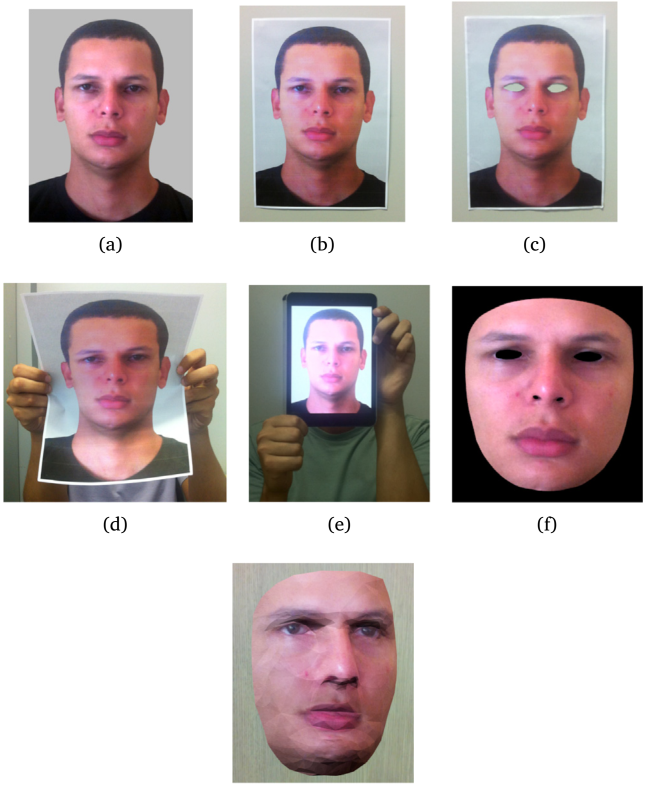

**Fig. 2.** Types of face spoofing attack: (a) genuine user; (b) flat printed photo; (c) eye-cut photo; (d) warped photo, (e) video playback; (f) life-size wearable mask and (g) paper-cut mask.

### 2.2. Taxonomy of face spoofing methods

Face recognition systems based on 2D and 3D images can be exposed to spoofing attacks, which can be verified by different approaches. In order to summarize them, we organized all gathered works in terms of descriptors and classifiers. Descriptors were categorized as texture, motion, frequency, color, shape or reflectance; while classifiers are organized as discriminant, regression, distance metric or heuristic. Table 1 presents the summary of the proposed taxonomy, and the descriptors and classifiers are respectively discussed and analyzed in Sections 2.3 and 2.4.

### 2.3. Descriptors

#### 2.3.1. Texture

Texture features are extracted from face images under the assumption that printed faces produce certain texture patterns that do not exist in real ones. Texture is probably the strongest evidence of spoofing, since more than 69% of the works (see Table 1) use texture alone or combine it with other descriptors in their countermeasures.

Different texture descriptors can be used to detect facial spoofing, but local binary patterns (LBP) (Ojala et al., 1996) is the very first choice, as it can be observed in Table 1. Indeed, nearly half of the surveyed works explore LBP or any of its variations. LBP is a grayscale, illumination-invariant, texture-coding technique that labels every pixel by comparing it with its neighbors, concatenating the result into a binary number. The number of neighbors, neighborhood radius, and coding strategy are all parameters of the method. The final computed labels are then organized in histograms to describe the texture, which can be performed for the entire image or even image paths. Different LBP configurations can be found in spoofing detection, such as the original LBP (Komulainen et al., 2013b; Chingovska et al., 2012; Erdogmus and Marcel, 2013; Kim et al., 2012b; Boulkenafet et al., 2015; de Souza et al., 2017, IDIAP Chakka, 2011, MaskDown Chingovska, 2013), multi-scale LBP (Maatta et al., 2011, 2012; Yang et al., 2013; Kose and Dugelay, 2013c, a, 2014; Erdogmus and Marcel, 2014; Yang et al., 2015; Kim et al., 2012a, UOULU Chakka, 2011, CASIA Chingovska, 2013, LNMIIT Chingovska, 2013, Muvis Chingovska, 2013), LBP variance (LBPV) (Kose and Dugelay, 2012), and LBP from three orthogonal planes (LBP-TOP) (Pereira et al., 2012; Asim et al., 2017, MaskDown Chingovska, 2013). LBP-TOP can be considered a hybrid texture-motion descriptor, since it combines both spatial and temporal information. Other texture-coding techniques were also explored for spoofing detection, including local phase quantization (LPQ) (Yang et al., 2013; Boulkenafet et al., 2016), which uses invariant blurring properties when extracting features from images. This descriptor is a phase information of the locally calculated Fourier spectrum for each position of the pixel in the image. Different approaches include local graph structure (LGS) (Housam et al., 2014b) and improved LGS (ILGS) (Housam et al., 2014a)$x'$, which were used to extract texture features by comparing a target pixel and its neighboring pixels. Other methods have also used texture descriptors in face spoofing detection competitions to analyze face spoofing attacks (AMILAB Chakka, 2011, PRA Lab Chingovska, 2013). Multiscale local phase quantization on three orthogonal planes (MLPQ-TOP) is an extension of LPQ for time-varying texture analysis, which explores the

**Table 1**
Summary of the gathered works on face spoofing detection.

| Descriptors | Related works |
|---|---|
| Texture | **LBP and variations** (Komulainen et al., 2013b; Chingovska et al., 2012; Erdogmus and Marcel, 2013; Kim et al., 2012b; Boulkenafet et al., 2015; de Souza et al., 2017, IDIAP Chakka, 2011, MaskDown Chingovska, 2013, Maatta et al., 2011, 2012; Yang et al., 2013; Kose and Dugelay, 2013c, a, 2014; Erdogmus and Marcel, 2014; Yang et al., 2015; Kim et al., 2012a, UOULU Chakka, 2011, CASIA Chingovska, 2013, LNMIIT Chingovska, 2013, Muvis Chingovska, 2013; Kose and Dugelay, 2012; Pereira et al., 2012; Asim et al., 2017, MaskDown Chingovska, 2013), **LPQ** (Yang et al., 2013; Boulkenafet et al., 2016), **LGS** (Housam et al., 2014b), **ILGS** (Housam et al., 2014a), **Texture** (AMILAB Chakka, 2011, PRA Lab Chingovska, 2013), **MLPQ-TOP** (Arashloo et al., 2015), **MBSIF-TOP** (Arashloo et al., 2015), **LDP-TOP** (Phan et al., 2017, 2016), **HOG** (Maatta et al., 2012; Schwartz et al., 2011; Yang et al., 2013, 2015; Komulainen et al., 2013a), **Gabor Wavelets** (Maatta et al., 2012, Muvis Chingovska, 2013), **GLCM** (Schwartz et al., 2011; Kim et al., 2012a; Pinto et al., 2012, 2015b, UNICAMP Chakka, 2011; Chingovska, 2013, MaskDown Chingovska, 2013), **DoG** (Peixoto et al., 2011; Zhang et al., 2012, UNICAMP Chakka, 2011), **HSC** (Feng et al., 2016; Schwartz et al., 2011, UNICAMP Chakka, 2011), **AOS** (Alotaibi and Mahmood, 2017), **DNN** (Menotti et al., 2015) |
| Motion | **CRF** (Pan et al., 2007), **OFL** (Feng et al., 2016; Kollreider et al., 2008, 2009), **HOOF** (Bharadwaj et al., 2013), **HMOF** (CASIA Chingovska, 2013), **RASL** (Yan et al., 2012, CASIA Chakka, 2011), **Motion Correlation** (Komulainen et al., 2013b, CASIA Chingovska, 2013), **GMM** (Yan et al., 2012; Pinto et al., 2015a, CASIA Chakka, 2011, LNMIIT Chingovska, 2013), **DMD** (Tirunagari et al., 2015), **Motion** (SIANI Chakka, 2011, IGD Chingovska, 2013) |
| Frequency | **2D-DFT** (Kim et al., 2012b; Phan et al., 2017; Pinto et al., 2012, 2015b, a; Garcia and de Queiroz, 2015, UNICAMP Chingovska, 2013), **1D-FFT** (CASIA Chingovska, 2013), **2D-FFT** (LNMIIT Chingovska, 2013), **Haar Wavelets** (Yan et al., 2012, CASIA Chakka, 2011) |
| Color | **CF** (Schwartz et al., 2011, UNICAMP Chakka, 2011), **IDA** (Kim et al., 2012a; Wen et al., 2015), **IQA** (Galbally and Marcel, 2014), **IQM** (ATVS Chingovska, 2013), **Color** (Boulkenafet et al., 2015, 2016; Lakshminarayana et al., 2017, AMILAB Chakka, 2011, PRA Lab Chingovska, 2013) |
| Shape | **CLM** (Wang et al., 2013) |
| Reflectance | **Variational Retinex** (Tan et al., 2010; Kose and Dugelay, 2013b, 2014) |
| Classifiers | Related works |
| Discriminant | **SVM** (Maatta et al., 2012; Kose and Dugelay, 2013c, a, b, 2014; Boulkenafet et al., 2016; Komulainen et al., 2013a, CASIA Chingovska, 2013; Maatta et al., 2011; Bharadwaj et al., 2013; Pereira et al., 2012; Chingovska et al., 2012; Erdogmus and Marcel, 2013; Kim et al., 2012b; Boulkenafet et al., 2015; Pinto et al., 2012; Wen et al., 2015; Tirunagari et al., 2015; Phan et al., 2017, 2016, (LNMIIT Chingovska, 2013; Yang et al., 2013; Wang et al., 2013; Komulainen et al., 2013b; Kim et al., 2012a; Zhang et al., 2012; Pinto et al., 2015a, UOULU Chakka, 2011, AMILAB Chakka, 2011, PRA Lab Chingovska, 2013), **LDA** (Bharadwaj et al., 2013; Erdogmus and Marcel, 2013; Galbally and Marcel, 2014; Erdogmus and Marcel, 2014, MaskDown Chingovska, 2013, ATVS Chingovska, 2013, **MLP** Komulainen et al., 2013b, **NN** Feng et al., 2016, **CNN** de Souza et al., 2017; Asim et al., 2017; Alotaibi and Mahmood, 2017; Menotti et al., 2015; Lakshminarayana et al., 2017, **BN** (SIANI Chakka, 2011), **Adaboost** (IGD Chingovska, 2013) |
| Regression | **LLR** (Komulainen et al., 2013b, MaskDown Chingovska, 2013), **LR** (Yan et al., 2012, CASIA Chakka, 2011), **SLR** (Peixoto et al., 2011), **SLRBLR** (Tan et al., 2010), **PLS** (Schwartz et al., 2011; Yang et al., 2015; Pinto et al., 2012, 2015b, UNICAMP Chakka, 2011, Muvis Chingovska, 2013), **KDA** (Arashloo et al., 2015) |
| Distance metric | $\chi^2$ (Kose and Dugelay, 2012, IDIAP Chakka, 2011), **Cosine** (Housam et al., 2014b, a) |
| Heuristic | **Blink count** (Pan et al., 2007), **Thresholding** (Kollreider et al., 2008; Garcia and de Queiroz, 2015), **Weighted sum** (Kollreider et al., 2009) |

blur-insensitive characteristic of the Fourier phase spectrum (Arashloo et al., 2015). Multiscale binarized statistical image feature descriptor on three orthogonal planes (MBSIF-TOP) use filters based on statistical learning that represent spatio-temporal texture descriptors (Arashloo et al., 2015). Dynamic texture descriptor can be found in local derivative pattern on three orthogonal planes (LDP-TOP) (Phan et al., 2017, 2016). LDP-TOP analyzes discriminative textures of spectrum videos, where subtle face movements occur over frames.

Histograms of oriented gradients (HOG) (Maatta et al., 2012; Schwartz et al., 2011; Yang et al., 2013, 2015; Komulainen et al., 2013a) is another texture descriptor that represents the variation of gradient orientations in different parts of the image in an illumination-invariant fashion. As such, the magnitude of the gradients in different orientations are summed in cells, which are lately combined in blocks. Bins, cells and blocks are normalized at the end to compose the final feature vector.

Gabor wavelets have been also applied in multiple scales and orientations in order to extract texture information in image cells. Usually, Gabor wavelets are calculated using mean and standard deviation of the magnitude of the coefficients at multiple scales and orientation (Maatta et al., 2012, Muvis Chingovska, 2013).

A compact and discriminant global representation can be achieved in the gray level co-occurrence matrices (GLCM) (Haralick et al., 1973).

GLCM describes the joint probability of neighboring pixels, and different Haralick features can be extracted from each matrix (Schwartz et al., 2011; Kim et al., 2012a; Pinto et al., 2012, 2015b, MaskDown Chingovska, 2013, UNICAMP Chingovska, 2013).

Edge information can also be considered for texture representation. In order to describe edges, difference of Gaussians (DoG) are used to remove lighting variations while preserving high frequency components (Peixoto et al., 2011; Zhang et al., 2012), and histograms of shearlet coefficients (HSC) are used to estimate the distribution of edge orientations in a multi-scale analysis (Feng et al., 2016; Schwartz et al., 2011). Nonlinear diffusion based on additive operator splitting (AOS) (Alotaibi and Mahmood, 2017) is also used to extract edge information for spoofing detection, applying a large time step to speed up the diffusing process and to distinguish the edges and surface texture in the input image.

Finally, following a very recent trend in computer vision field, deep neural networks (DNN) (Menotti et al., 2015) are trained in order to provide adaptive features which describe trainable texture.

#### 2.3.2. Motion

Table 1 clearly shows that motion descriptors are the second in importance for face spoofing detection, and there are two different ways of considering motion for this purpose. One way is to detect and describe intra-face variations, such as eye blinking, facial expressions and head rotation. Conditional random fields (CRF) have been recently used to determine eye closity and consequently detect blinking (Pan et al., 2007); for global facial movements, optical flow of lines (OFL) is used to measure spatio-temporal variations of face images in horizontal and vertical orientations (Feng et al., 2016; Kollreider et al., 2008, 2009), histogram of oriented optical flow (HOOF) and histogram of magnitudes of optical flows (HMOF) are applied to create a binned representation of facial motion directions and magnitudes (CASIA (Chingovska, 2013)); and robust alignment by sparse and low-rank decomposition (RASL) tries to align faces in multiple frames and measure non-rigid motion (Yan et al., 2012, CASIA Chakka, 2011).

Another way of using motion is to evaluate the consistency of the user interaction within the environment. In light of that, motion correlation between face and background regions is computed (Komulainen et al., 2013b, CASIA Chingovska, 2013), as well as, traditional background subtraction based on Gaussian mixture models (GMM) (Yan et al., 2012; Pinto et al., 2015a, CASIA Chakka, 2011, LNMIIT Chingovska, 2013).

Facial texture of an individual within a sequence of frames is explored by using the dynamic mode decomposition (DMD) (Tirunagari et al., 2015), which extracts features by means of eigenfaces in the snapshots displaced on temporal–spatial. DMD is used in combination with LBP technique as a texture descriptor, which is applied to capture evidences of human presence in a video sequence, such as eye blink and movements of the lips. Finally, in competitions, other types of methods have been used as a motion descriptor to analyze face spoofing attacks (SIANI Chakka, 2011, IGD Chingovska, 2013).

#### 2.3.3. Frequency

Frequency-based countermeasures take advantage of certain image artifacts that occur in spoofing attacks. 2D discrete Fourier transform (2D-DFT), and 1D and 2D fast Fourier transform (1D-FFT, 2D-FFT) are calculated to find these artifacts in single (Kim et al., 2012b) or multiple images (Phan et al., 2017; Pinto et al., 2012, 2015b, a; Garcia and de Queiroz, 2015, CASIA Chingovska, 2013, LNMIIT Chingovska, 2013, UNICAMP Chingovska, 2013). When one considers multiple images, the concept of Visual Rhythms is used to merge multiple Fourier spectra in a single map that represents spatial frequency information over time, and then HOG, LBP and/or GLCM can be used for final face representation. When specifically considering color banding, which concerns abrupt changes caused by inaccurate print or screen flicker, Haar wavelets decomposition can be applied to find large unidirectional variations (Yan et al., 2012, CASIA Chakka, 2011).

#### 2.3.4. Color

Although colors do not remain constant due to lighting variations, certain dominant characteristics are considerable clues to discriminate impostors from genuine faces. In this context, color frequency (CF) histograms describe the distribution of colors in an image (Schwartz et al., 2011, UNICAMP Chakka, 2011). These histograms are computed for different blocks of the image, as performed by HOG, using three bins to encode the number of pixels with the highest gradient magnitude in each color channel. Image moments globally describe face liveness by means of image distortion analysis (IDA) (Kim et al., 2012a; Wen et al., 2015), image quality assessment (IQA) (Galbally and Marcel, 2014) and image quality measures (IQM) (ATVS Chingovska, 2013). IDA was proposed to extract characteristics through the HSV and RGB color spaces, smoothing and light intensity. IQA allows to maximize both critical performance measures in a complete face spoofing detection. IQM aims to show that the lowest values, obtained by quality measurements, produced with Gaussian filtering are samples of impostor face. YCbCr and HSV color spaces are used as color descriptors in Boulkenafet et al. (2015, 2016). In Lakshminarayana et al. (2017), each channel of RGB color space was used for feature extraction. Other methods have been used as color descriptors, used in competitions, to analyze face spoofing attacks (AMILAB Chakka, 2011, PRA Lab Chingovska, 2013).

#### 2.3.5. Shape

Shape information is very useful to deal with printed photo attacks, since facial geometry cannot be reproduced in a planar surface. Active contours based on constrained local models (CLM) are used to detect facial landmarks in a video sequence. These landmarks define then a sparse 3D structure that describes the planarity of the face (Wang et al., 2013).

#### 2.3.6. Reflectance

Considering that genuine and impostor faces behave differently in the same illumination conditions, it is possible to use the reflectance information to distinguish them. To accomplish that, the Variational Retinex method decomposes an input image into reflectance and illumination components (Tan et al., 2010; Kose and Dugelay, 2013b, 2014) in order to analyze the entire image.

### 2.4. Classifiers

#### 2.4.1. Discriminant

The idea behind discriminant techniques is to distinguish different classes by minimizing intra-class variation and/or maximizing inter-class variation. This type of classifier is explored in approximately 64% of the gathered works.

As evidenced in Table 1, works use a discriminant classifier alone or along with others in their frameworks. Support vector machines (SVM) are the most common classification technique in spoofing detection, and often presents superior performance. In order to achieve that, SVM finds optimal hyperplanes to separate descriptors from genuine and impostor faces. When these classes are not linearly separable, different kernel functions can be used to obtain a nonlinear classifier. Although linear SVM has been extensively used in different countermeasures (Maatta et al., 2012; Kose and Dugelay, 2013c, a, b, 2014; Boulkenafet et al., 2016; Komulainen et al., 2013a; Pinto et al., 2012, CASIA Chingovska, 2013), radial basis function kernel (Maatta et al., 2011; Bharadwaj et al., 2013; Pereira et al., 2012; Chingovska et al., 2012; Erdogmus and Marcel, 2013; Kim et al., 2012b; Boulkenafet et al., 2015; Pinto et al., 2012; Wen et al., 2015), and histogram intersection kernel (Tirunagari et al., 2015; Phan et al., 2017, 2016) have also been applied to increase the classification accuracy. Different SVM versions can also be considered, such as Hidden Markov Support Vector Machines (LNMIIT Chingovska, 2013). In (Yang et al., 2013; Wang et al., 2013; Komulainen et al., 2013b; Kim et al., 2012a; Zhang et al., 2012; Pinto et al., 2015a, UOULU Chakka, 2011, AMILAB Chakka, 2011, PRA Lab Chingovska, 2013), however, the authors do not describe the type of SVM kernel used in the experiments.

As an alternative to linear approaches, the linear discriminant analysis (LDA) (Bharadwaj et al., 2013; Erdogmus and Marcel, 2013; Galbally and Marcel, 2014; Erdogmus and Marcel, 2014, MaskDown Chingovska, 2013, ATVS Chingovska, 2013) explicitly models the difference between classes and within classes to address the classification task, with an advantage of being used for dimensionality reduction.

Other types of classifiers use discriminant procedures to accomplish face spoofing detection: multilayer perceptron (MLP) (Komulainen et al., 2013b) was used to evaluate whether excessive movement (flat printed photo-strike by hand) or no movement (flat printed photo strike attached to a media) had variations during an $N$ video sequence; neural network (NN) (Feng et al., 2016) is good at learning implicit patterns, which is able to recognize motion cues for spoofing detection with proper training. NN is trained by a backpropagation procedure using a labeled data set through an autoencoder, which is treated as a pre-training process; convolutional neural networks (CNN) (de Souza et al., 2017; Asim et al., 2017; Alotaibi and Mahmood, 2017; Menotti et al., 2015; Lakshminarayana et al., 2017) uses trainable features with shared weights and local connections between different layers, where all weights in all layers of a CNN network are learned through training. A CNN aims to learn invariance representations of scale, translation, rotation and related transformations on a trainable-based feature framework. Bayesian network (BN) (SIANI Chakka, 2011) provides a probabilistic method by extending Bayes's rule for updating probabilities in the light of new evidences. Adaboost is a type of ensemble classifier that speeds up the process of finding discrimination of impostor and genuine users (IGD Chingovska, 2013).

### 2.4.2. Regression

The regression-based classification maps use input descriptors directly into their class labels considering a predictive model obtained from known pairs of descriptors and labels. They have been widely used for spoofing detection due to their simplicity, accuracy and efficiency. Different regression methods are referred in the literature: linear logistic regression (LLR) (Komulainen et al., 2013b, MaskDown Chingovska, 2013) was used for the combination of information extracted through two descriptors; for the combination of correlative motion and texture (*e.g.*, LBP) applications, using MLP and SVM classifiers, respectively; logistic regression (LR) (Yan et al., 2012, CASIA Chakka, 2011) is a confidence quantification of every feature representation, which in the final scores are fused by weight sum rule, and the weights of different classifiers are learned on validation set using grid search; sparse logistic regression (SLR) (Peixoto et al., 2011) analyzes different lighting conditions and regions of high frequency for detecting images made by impostors; sparse low rank bilinear logistic regression (SLRBLR) (Tan et al., 2010) was explored in images with prominent reflectance and illumination, using two techniques to extract these characteristics of the image, being reflectance based on variational Retinex, and the illumination based on the DoG technique in the identification of medium–high frequency bands; partial least square (PLS) (Schwartz et al., 2011; Yang et al., 2015; Pinto et al., 2012, 2015b, UNICAMP Chakka, 2011, Muvis Chingovska, 2013) is calculated from a linear transformation on the features extracted by descriptors using weighting methods.

Kernel discriminant analysis (KDA) is a regression classifier, which projects the input data onto a discriminative spectral subspace, avoiding the computational time found in eigen-analysis (Arashloo et al., 2015). In other words, KDA uses projective functions (vectors) based on eigendecomposition of kernel matrix, being costly when applied to a large number of training samples.

### 2.4.3. Distance metric

The use of distance metrics is supposed to improve the performance in face spoofing detection systems, with the goal of measuring the dissimilarities among samples. However, these approaches usually require an exhaustive search to accomplish the classification task, which may lead to a high cost in large reference sets. Chi-square ($\chi^2$) (Kose and Dugelay, 2012, IDIAP Chakka, 2011) and cosine distance (Housam et al., 2014b, a) are common choices to this end, and they are used to compute the cumulative distance of a probe face (that one to be identified) and the entire reference set to decide whether the face is genuine or impostor.

### 2.4.4. Heuristic

Different heuristics have been used to decide whether a face is real or fake. As a drawback, heuristics may lead to overfitting, specially when only self-collected data is considered. Number of eye blinks (Pan et al., 2007), motion measurements thresholding (Kollreider et al., 2008), average pixel ratio thresholding (Garcia and de Queiroz, 2015) and weighted sum of motion measurements (Kollreider et al., 2009) are examples of heuristics found in the literature.

## 3. Quantitative evaluation of the surveyed works

Results reported in the surveyed papers were grouped according to the data sets used in their experiments. All numerical values in our study are the exact same values presented in their original works, which followed the same evaluation protocol.

### 3.1. Data sets

Nine publicly available data sets were chosen to evaluate the methods: Concerning 2D attacks, NUAA Photograph Imposter (Tan et al., 2010), Yale Recaptured (Peixoto et al., 2011), Print-Attack (Anjos and Marcel, 2011), Replay-Attack (Chingovska et al., 2012), Casia Face Anti-Spoofing (Zhang et al., 2012), MSU-MFSD (Wen et al., 2015) and UVAD (Pinto et al., 2015b; da Silva Pinto, 2013) are the most known and used in the literature; with respect to mask attacks, Kose and Dugelay's data set (Kose and Dugelay, 2013c) and 3D Mask Attack data set (Erdogmus and Marcel, 2013) are the only two found in the literature. General characteristics of each data set are summarized in Table 2, and more details can be found in Peixoto et al. (2011), Tan et al. (2010), Kose and Dugelay (2013c), Chingovska et al. (2012), Erdogmus and Marcel (2013), Zhang et al. (2012) and Anjos and Marcel (2011).

NUAA Photograph *Imposter* data set[5] (Tan et al., 2010) is one of the first publicly available data sets for face spoofing detection evaluation. Images in NUAA were collected by cheap webcams in three sessions on different environments and under different illumination conditions, with an interval of two weeks between each session. The evaluated attack is a printed photo, which can be flat or warped. These photo attacks were prepared using A4 paper and a color printer.

The main goal of the Yale Recaptured data set[6] (Peixoto et al., 2011) was to have impostor images in multiple illumination conditions. As such, texture-based methods are commonly employed over this data set. Static images were collected with a distance of 50 centimeters between the display and the camera.

The Print-Attack data set[7] (Anjos and Marcel, 2011) was used to benchmark different works in the first spoofing detection competition (Chakka, 2011). This data set was created by showing a flat printed photo of a genuine user to an acquisition sensor in two ways: Hand-held (*i.e.*, the impostor holds the photo using the hands) or fixed support (*i.e.*, photos are stuck on a wall).

Replay-Attack data set[8] (Chingovska et al., 2012) is an extension of the Print-Attack data set to evaluate spoofing in videos and photos, and it was used in the second spoofing detection competition (Chingovska, 2013). It consists of 1300 video clips of photo and video attacks. All

[5] http://parnec.nuaa.edu.cn/xtan/NUAAImposterDB_download.html.

[6] http://ic.unicamp.br/rocha/pub/downloads/2011icip/.

[7] https://www.idiap.ch/dataset/printattack/downloadproc.

[8] https://www.idiap.ch/dataset/replayattack/downloadproc.

**Table 2**
Summary of available face spoofing data sets.

| Year | Data set | #Subjects | #Real/Fake | Type of attack |
|---|---|---|---|---|
| 2010 | NUAA Photograph Imposter (Tan et al., 2010) | 15 | 5105/7509 | 1. Flat printed photo<br>2. Warped photo |
| 2011 | Yale Recaptured (Peixoto et al., 2011) | 10 | 640/1920 | 1. Flat printed photo |
| 2011 | Print-Attack (Anjos and Marcel, 2011) | 50 | 200/200 | 1. Flat printed photo |
| 2012 | Replay-Attack (Chingovska et al., 2012) | 50 | 200/1000 | 1. Flat printed photo<br>2. Video playback |
| 2012 | Casia Face Anti-Spoofing (Zhang et al., 2012) | 50 | 150/450 | 1. Warped photo<br>2. Eye-cut photo<br>3. Video playback |
| 2013 | Kose and Dugelay (Kose and Dugelay, 2013c) | 20 | 200/198 | 1. Mask |
| 2013 | 3D Mask Attack (Erdogmus and Marcel, 2013) | 17 | 170/85 | 1. Mask |
| 2014 | MSU-MFSD (Wen et al., 2015) | 35 | 70/210 | 1. Flat printed photo<br>2. Video playback |
| 2015 | UVAD (Pinto et al., 2015b; da Silva Pinto, 2013) | 404 | 808/16268 | 1. Video playback |

**Table 3**
Summary of non-public face spoofing data sets.

| Year | data set | #Subjects | #Real/Fake | Types of attacks |
|---|---|---|---|---|
| 2012 | Pinto et al. (2012) | 50 | 100/600 | 1. Video playback |
| 2012 | BERC Webcam (Kim et al., 2012b) | 25 | 1408/7461 | 1. Flat printed photo |
| 2012 | BERC ATM (Kim et al., 2012b) | 20 | 1797/5802 | 1. Flat printed photo |
| 2013 | Wang et al. (2013) | 50 | 750/2250 | 1. Flat printed photo<br>2. Warped photo |

images and videos were collected under different lighting conditions, and three different attacks modes were considered: printed photo in high-resolution and video playbacks, using a mobile phone with low-resolution screen, and a 1024 × 768 pixels ipod screen.

Casia Face Anti-Spoofing data set[9] (Zhang et al., 2012) contains seven scenarios with different types of attack and a variety of image qualities. This data set presents three types of attacks: warped photo, eye-cut photo and video playback. Kose and Dugelay's data set[10] (Kose and Dugelay, 2013c) is a paid-mask data set created by the MORPHO company. Subjects were captured by a 3D scanner that uses a structured light technology to obtain genuine images of facial shape and texture. After that, masks for those images were manufactured by Sculpteo 3D Printing,[11] and then recaptured by the same sensor to obtain impostor images.

3D Mask-Attack data set (3DMAD)[12] (Erdogmus and Marcel, 2013) was the first publicly available data set for mask attacks, and it consists of video sequences recorded by an RGB-D camera. Masks were manufactured using the services of ThatsMyFace,[13] and one frontal and two profile images of each subject for that purpose were required.

MSU Mobile Face Spoofing data set (MSU MFSD)[14] (Wen et al., 2015) consists of 280 video clips of print photo and video attack attempts to 35 participants. All printed photos used for attacks were created with a state-of-the-art color printer on larger sized paper. In order to perform an attack, video playback from each participant was taken under the similar conditions as in their authentication sessions.

Unicamp Video-Attack data set (UVAD)[15] (Pinto et al., 2015b; da Silva Pinto, 2013) is comprised of videos of valid accesses and attacks of 404 subjects, all built at Full HD quality, recorded at 30 frames per second and nine seconds long. All videos were created by filming each person in two sections under different lighting conditions, backgrounds and places (indoors and outdoors).

**Table 4**
Metrics commonly applied on face spoofing evaluation.

| Metric | Stand for | Equation | Type |
|---|---|---|---|
| FAR | False acceptance rate | $FAR = \frac{NFA}{\#Impostor}$ | Error |
| FRR | False rejection rate | $FRR = \frac{NFR}{\#Genuine}$ | Error |
| EER | Equal error rate | $EER = (FAR = FRR)$ | Error |
| HTER | Half total error rate | $HTER = \frac{FAR+FRR}{2}$ | Error |
| ACC | Accuracy | $100 \times \left(1 - \frac{FAR \times \#Impostor + FRR \times \#Genuine}{\#Impostor + \#Genuine}\right)$ | Hit |
| AUC | Area under curve | $Area = \int_a^b f(x)dx$, where $f : [a,b] \to \mathbb{R}$ | Hit |

Other non-public data sets are proposed in Wang et al. (2013), Kim et al. (2012b) and Pinto et al. (2012). In Pinto et al. (2012), the data set was introduced to detect video-based spoofing. Kim et al. (2012b) proposed two different data sets called BERC Webcam and BERC ATM, which consist of images taken from live people and four types of 2-D paper masks (photo, print, magazine, caricature). Wang et al. (2013) created an image data set for photo attack evaluation using flat-printed and warped photos. Table 3 summarizes the main characteristics of the aforementioned data sets.

### 3.2. Performance metrics

A spoofing detection system is subject to two types of errors: an impostor can be accepted as a genuine user (*i.e.* number of false acceptance — NFA), or a genuine user can be considered as an impostor (*i.e.,* number of false rejection — NFR). The probability of these errors to occur are respectively called false acceptance rate (FAR) and false rejection rate (FRR). These rates present an inversely proportional relation. A receiver operating characteristics (ROC) curve is obtained by computing all possible pairs of FAR and FRR values, as illustrated in Fig. 3. The integral of a ROC curve is known as the area under curve (AUC), *i.e.,* the gray-filled area in Fig. 3. Also, the point of the ROC curve where FAR equals FRR is called equal error rate (EER), and the point where the average of FAR and FRR is minimal is called half total error rate (HTER). Finally, the overall accuracy (ACC) considers both genuine users and impostors along with the FAR and FRR. Table 4 summarizes all the aforementioned metrics.

[9] http://www.cbsr.ia.ac.cn/english/FaceAntiSpoofdatasets.asp.

[10] http://www.morpho.com/.

[11] http://www.sculpteo.com/en/.

[12] https://www.idiap.ch/dataset/3dmad/download-proc.

[13] http://www.thatsmyface.com/Products/products.html.

[14] http://www.cse.msu.edu/rgroups/biometrics/Publications/datasets/MSUMobileFaceSpoofing/index.htm.

[15] http://figshare.com/articles/visualrhythm_antispoofing/1295453.

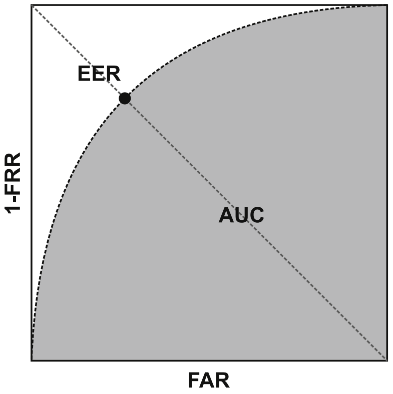

**Fig. 3.** Relation among the metrics on the ROC curve.

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 a biased performance analysis. All other metrics are based on a separate evaluation of FAR and FRR, so they are more reliable for a comparative analysis. For these reasons, surveyed works were compared using metrics according to the following order of preference: EER, HTER, AUC and ACC.

### 3.3. *Performance analysis of the existing spoofing detectors*

Comparing different works is a difficult task, since most of the time we do not have access to the original source codes, and reproducing codes and experimental results are very complicated. For that reason, we have decided to perform this comparison by using the results reported in the gathered papers. However, 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, specially if this data set is prone to be biased (Torralba and Efros, 2011). Strictly speaking, besides a common available data set, it is of underlying importance to follow the same methodology and to have the same metrics when comparing different countermeasures.

Given the data sets presented in Section 3.1, some criteria were adopted to select which works should be considered in our analysis: (i) they must follow the same data set protocol; (ii) they must report its results using at least one of the metrics discussed in Section 3.2; and (iii)they must be comparable to other works using the same data set (*i.e.* use the most common metrics for that data set). On that account, some works were then removed from the analysis for NUAA (Chingovska et al., 2012; Housam et al., 2014b, a), Print-Attack (Yan et al., 2012; Yang et al., 2013), Replay-Attack (Yang et al., 2015; Garcia and de Queiroz, 2015), Casia Face Anti-Spoofing (Tirunagari et al., 2015; Chingovska et al., 2012; Galbally and Marcel, 2014; Yang et al., 2015; Pinto et al., 2015a; Wen et al., 2015; Lakshminarayana et al., 2017), Kose and Dugelay's (Kose and Dugelay, 2013a), 3DMAD (Erdogmus and Marcel, 2014; Pinto et al., 2015a) data sets and other works in Pan et al. (2007), Kollreider et al. (2008, 2009). Therefore, such works are not presented in Table 1 and Fig. 4. Tables 5–12 summarize the selected results. It is noteworthy that sometimes it was necessary to take conclusions by indirectly comparing different metrics, as in Table 5.

Table 5 summarizes the results of the methods concerning the NUAA data set, and the most common metrics were EER, AUC and ACC. This data set presents a relatively balanced number of positives and negatives samples, which avoids biased results when using ACC. Peixoto et al. (2011) did not report both EER and AUC, but their ACC shows that they did not achieve the best performance. As it can be observed in Table 5, methods with high AUC have low EER. Although AUC does not allow us to differ between the methods proposed by Maatta et al. (2011, 2012) and Yang et al. (2013), EER clearly shows that Maatta et al. (2012) achieved the best performance for the NUAA data set.

A summary of the results considering the Yale Recaptured data set is presented in Table 6. Works using this data set were compared by means of ACC, the only metric in common to all of them. Since this data set is highly unbalanced (*i.e.,* a ratio of 1:3), ACC would not be the most recommended metric. For this comparison, however, the use of ACC is not an issue due to the perfect performance reported (Maatta et al., 2012), which means that both classes were perfectly classified.

As stated in Section 3.1, the Print-Attack data set was used as a benchmark in the first spoofing detection competition (Chakka, 2011), wherein three works achieved perfect score (*i.e.*, IDIAP Chakka, 2011, UOULU Chakka, 2011 and CASIA Chakka, 2011). Later works (Maatta et al., 2012; Tirunagari et al., 2015) also achieved 0% of HTER (see Table 7), and Maatta et al. (2012) reached the best performance in a third data set. As shown in Table 2, NUAA, Yale Recaptured and Print-Attack data sets are solely based on printed photo attacks. Given the work in Maatta et al. (2012) achieved the lowest error rates in all of the attacks using the same approach, it is safe to assume that multiple texture features (*i.e.,* LBP, Gabor wavelets and HOG) and an SVM classifier are enough to detect printed photo attacks over that data set.

The Replay-Attack data set was used in the second spoofing detection competition (Chingovska, 2013), and both CASIA and LNMIIT obtained 0% of HTER. There is a proposed method in Feng et al. (2016), which also achieved a perfect HTER (see Table 8). This data set has an uneven number of real and fake images (*i.e.,* a ratio of 1:5), but it does not influence the analysis since all works report their results using the same metric.

The Casia Face Anti-Spoofing data set is characterized by the highest number of types of attacks, as presented in Table 2, but presenting a low number of samples. Table 9 presents the results on this data set, and Boulkenafet et al. (2016) proposed the method with the best performance, reaching perfect results (*i.e.,* 3.20% EER).

Tables 10 and 11 present, respectively, the spoofing detection results for mask attacks over Kose and Dugelay's and 3D Mask Attack data sets. The best performance over that data set was achieved by the method based on texture and reflectance descriptors, and an SVM classifier. The number of fake images in the 3D Mask Attack data set was greater than the number of real ones, but works were evaluated using HTER; so unbalancing is not a problem. Menotti et al. (2015) obtained the best performance by combining deep learning and SVM. Methods dealing with video playback and mask attacks did not rely only on texture descriptors, exploring different features (*i.e.,* motion, frequency and reflectance) to reduce the classification error. SVM is still the most preferred classifier.

Table 12 summarizes the results over UVAD and MSU-MFSD data sets. The first one was only used by one work, while in the second Boulkenafet et al. (2016) achieved the lowest EER. Table 13 shows results over non-public data sets, and it is only presented for completeness, since it is not easy reproducible.

## 4. Discussion and analysis

Most of the effort to address the problem of face spoofing detection have been carried out over the past decade. Henceforth we provide a big picture of the field (trends), as well as the analysis of open issues and future perspectives that could be tackled and followed in order to leverage the face spoofing systems.

### 4.1. *Timeline and trends of the state-of-the-art works*

Fig. 4 depicts a chronological arrangement of the surveyed works in order to demonstrate the convergence of descriptors and classifiers over the time.

**Table 5**
Results over NUAA Imposter data set.

| Reference | Features | Classifier | EER (%) | AUC | ACC (%) |
|---|---|---|---|---|---|
| 2010, Tan et al. (2010) | Variational Retinex | SLRBLR | – | 0.94 | – |
| 2011, Maatta et al. (2011) | LBP | SVM | 2.90 | 0.99 | 98.00 |
| 2011, Peixoto et al. (2011) | DoG | SLR | – | – | 93.20 |
| 2011, Schwartz et al. (2011) | CF + HOG + HSC + GLCM | PLS | 8.20 | 0.96 | – |
| **2012**, Maatta et al. (2012) | **LBP + Gabor Wavelets + HOG** | **SVM** | **1.10** | **0.99** | **–** |
| 2012, Kose and Dugelay (2012) | LBPV | $\chi^2$ | 11.97 | – | – |
| 2013, Yang et al. (2013) | LBP + LPQ + HOG | SVM | 1.90 | 0.99 | 97.70 |
| 2015, Arashloo et al. (2015) | MLPQ-TOP + MBSIF-TOP | KDA | 1.80 | – | – |
| 2017, Alotaibi and Mahmood (2017) | AOS | CNN | – | – | 99.00 |
| 2017, de Souza et al. (2017) | LBP | CNN | 1.80 | 0.99 | 98.20 |

**Table 6**
Results over Yale Recaptured data set.

| Reference | Features | Classifier | ACC (%) |
|---|---|---|---|
| 2011, Peixoto et al. (2011) | DoG | SLR | 91.70 |
| **2012**, Maatta et al. (2012) | **LBP + Gabor Wavelets + HOG** | **SVM** | **100.00** |

**Table 7**
Results over Print-Attack data set.

| Reference | Features | Classifier | HTER (%) |
|---|---|---|---|
| **2011, IDIAP** (Chakka, 2011) | **LBP** | $\chi^2$ | **0.00** |
| **2011, UOULU** (Chakka, 2011) | **LBP** | **SVM** | **0.00** |
| **2011, CASIA** (Chakka, 2011) | **RASL + GMM + Haar Wavelets** | **LR** | **0.00** |
| 2011, AMILAB (Chakka, 2011) | Color + Texture | SVM | 0.63 |
| 2011, SIANI (Chakka, 2011) | Motion | BN | 10.63 |
| 2011, UNICAMP (Chakka, 2011) and Schwartz et al. (2011) | CF + HOG + HSC + GLCM | PLS | 0.63 |
| **2012**, Maatta et al. (2012) | **LBP + Gabor Wavelets + HOG** | **SVM** | **0.00** |
| 2013, Bharadwaj et al. (2013) | HOOF | LDA | 0.62 |
| **2015**, Tirunagari et al. (2015) | **DMD + LBP** | **SVM** | **0.00** |

**Table 8**
Results over Replay-Attack data set.

| Reference | Features | Classifier | HTER (%) |
|---|---|---|---|
| 2012, Chingovska et al. (2012) | LBP | SVM | 15.16 |
| 2012, Pereira et al. (2012) | LBP-TOP | SVM | 7.60 |
| 2013, Komulainen et al. (2013b) | Motion Correlation + LBP | LLR + SVM + MLP | 5.11 |
| 2013, Bharadwaj et al. (2013) | HOOF + LBP | LDA | 1.25 |
| **2013, CASIA** (Chingovska, 2013) | **LBP + 1D-FFT + HMOF + Motion Correlation** | **SVM** | **0.00** |
| 2013, IGD (Chingovska, 2013) | Motion | Adaboost | 9.13 |
| 2013, MaskDown (Chingovska, 2013) | LBP + GLCM + LBP-TOP | LLR + LDA | 2.50 |
| **2013, LNMIIT** (Chingovska, 2013) | **LBP + GMM + 2D-FFT** | **SVM** | **0.00** |
| 2013, Muvis (Chingovska, 2013) | LBP + Gabor Wavelets | PLS | 1.25 |
| 2013, PRA Lab (Chingovska, 2013) | Color + Texture | SVM | 1.25 |
| 2013, ATVS (Chingovska, 2013) | IQM | LDA | 12.00 |
| 2013, UNICAMP (Chingovska, 2013) | 2D-DFT + GLCM | SVM | 15.62 |
| 2014, Galbally and Marcel (2014) | IQA | LDA | 15.20 |
| 2015, Menotti et al. (2015) | DNN | CNN | 0.75 |
| 2015, Tirunagari et al. (2015) | DMD + LBP | SVM | 3.75 |
| 2015, Wen et al. (2015) | IDA | SVM | 7.41 |
| 2015, Pinto et al. (2015b) | 2D-DFT | PLS | 14.27 |
| 2015, Boulkenafet et al. (2015) | LBP + Color | SVM | 2.90 |
| 2015, Arashloo et al. (2015) | MLPQ-TOP + MBSIF-TOP | KDA | 1.00 |
| 2015, Pinto et al. (2015a) | GMM + 2D-DFT | SVM | 2.75 |
| 2016, Boulkenafet et al. (2016) | LPQ + Color | SVM | 3.30 |
| **2016**, Feng et al. (2016) | **HSC + Optical Flow** | **NN** | **0.00** |
| 2016, Kim et al. (2012a) | MLBP + GLCM + IDA | SVM | 5.50 |
| 2016, Phan et al. (2016) | LDP-TOP | SVM | 1.75 |
| 2017, Alotaibi and Mahmood (2017) | AOS | CNN | 10.00 |
| 2017, Lakshminarayana et al. (2017) | Color | CNN | 0.80 |

From 2007 to 2010, spoofing detectors were mostly focused on the analysis of motion or reflectance, since both types of descriptors are based on a quite straightforward observation: printed faces do not behave or reflect light as real faces do. Although such countermeasures have persisted to date, another image cue has grown in importance in the literature: face texture. As pointed out by Tan et al. (2010), an impostor face is captured by a camera twice, while a genuine-user once. The former consequently produces artifacts that are not presented in real face acquisition. These artifacts are very perceptible in texture images, and texture coding techniques seem to be an effective way to capture and describe them, as evidenced in the number of works relying on traditional texture description approaches and their variations from 2011 to 2017.

In terms of classification, SVM-based works became more and more popular to the point of dominating the face spoofing literature in recent years, which is somehow expected, since SVM has gained a wide attention in many other machine learning tasks, such as medical diagnosis (Sweilam et al., 2010), object recognition (Muralidharan and

**Table 9**
Results over Casia face anti-spoofing data set.

| Reference | Features | Classifier | EER (%) |
|---|---|---|---|
| 2012, Zhang et al. (2012) | DoG | SVM | 17.00 |
| 2013, Komulainen et al. (2013a) | HOG | SVM | 3.30 |
| 2013, Yang et al. (2013) | LPQ + LBP + HOG | SVM | 11.80 |
| 2015, Boulkenafet et al. (2015) | LBP + Color | SVM | 6.20 |
| **2016**, Boulkenafet et al. (2016) | **LPQ + Color** | **SVM** | **3.20** |
| 2016, Feng et al. (2016) | HSC + Optical Flow | NN | 5.83 |
| 2016, Kim et al. (2012a) | MLBP + GLCM + IDA | SVM | 4.89 |
| 2016, Phan et al. (2016) | LDP-TOP | SVM | 8.94 |
| 2017, Asim et al. (2017) | LBP-TOP | CNN | 8.02 |

**Table 10**
Results of different methods over Kose and Dugelay's data set.

| Reference | Features | Classifier | AUC | ACC (%) |
|---|---|---|---|---|
| 2013, Kose and Dugelay (2013b) | Variational Retinex | SVM | 0.97 | 94.47 |
| 2013, Kose and Dugelay (2013c) | LBP | SVM | 0.98 | 93.50 |
| **2014**, Kose and Dugelay (2014) | **LBP + Variational Retinex** | **SVM** | **0.99** | **98.99** |

**Table 11**
Results over 3D mask attack data set.

| Reference | Features | Classifier | HTER (%) |
|---|---|---|---|
| 2013, Erdogmus and Marcel (2013) | LBP | LDA | 0.95 |
| **2015**, Menotti et al. (2015) | **DNN** | **CNN** | **0.00** |
| **2016**, Feng et al. (2016) | **HSC + Optical Flow** | **NN** | **0.00** |

**Table 12**
Results over UVAD and MSU-MFSD data set.

| Data set | Reference | Features | Classifier | HTER (%) | EER (%) |
|---|---|---|---|---|---|
| UVAD | 2015, Pinto et al. (2015a) | GMM + 2D-DFT | SVM | 29.87 | – |
| UVAD | 2017, Phan et al. (2017) | 2D-DFT + LDP-TOP | SVM | 23.69 | – |
| MSU-MFSD | 2015, Wen et al. (2015) | IDA | SVM | – | 5.82 |
| MSU-MFSD | 2016, Boulkenafet et al. (2016) | LPQ + Color | SVM | – | 3.50 |
| MSU-MFSD | 2016, Phan et al. (2016) | LDP-TOP | SVM | 7.70 | 6.54 |
| **MSU-MFSD** | **2016**, Kim et al. (2012a) | **MLBP GLCM + IDA** | **SVM** | – | **2.44** |

**Table 13**
Results of different methods over other non-public data sets.

| Data set | Features | Classifier | EER (%) | ACC (%) |
|---|---|---|---|---|
| BERC Webcam (Kim et al., 2012b) | LBP + 2D DFT | SVM | 8.43 | – |
| BERC ATM (Kim et al., 2012b) | LBP + 2D DFT | SVM | 4.42 | – |
| **Self collected** (Pinto et al., 2012) | **GLCM + 2D DFT** | **PLS** | – | **100.00** |
| **Self collected** (Pinto et al., 2012) | **GLCM + 2D DFT** | **SVM** | – | **100.00** |
| **Self collected** (Wang et al., 2013) | **CLM** | **SVM** | – | **100.00** |

Chandrasekar, 2011) and market analysis (Huang et al., 2005). In fact, even if we consider only face processing applications, there are several ways of exploring SVM: face recognition (Tefas et al., 2001), face detection (Osuna et al., 1997), facial landmark extraction (Rapp et al., 2011), facial expression analysis (Kotsia and Pitas, 2007), and so forth. Although SVM provides very accurate results, the two most researched and up-to-date of these applications (*i.e.,* detection and recognition) are recently getting better results using deep learning methods (Zhang and Zhang, 2014; Taigman et al., 2011). Thus, we expect to see an attention shift towards deep learning in spoofing detection works for the next few years, which can already be seen in the most recent literature (Menotti et al., 2015; Alotaibi and Mahmood, 2017; Lakshminarayana et al., 2017; de Souza et al., 2017; Asim et al., 2017).

### 4.2. *Open issues*

After surveying all existing face spoofing detection methods in the last decade, it is still difficult to establish if there was a remarkable progress in this field. The main points that support this view are the following:

1. Automatic detection of impostors by face still follows the same recipe as many other computer vision problems: first extracting some features for further classifying them by a supervised predictor. Moreover, most works follow the same architecture of popular face recognition systems, using similar feature sets and classification methods. This is even more evident if one observes Table 14, where the best performing works over all data sets considered in our review are shown. As stated in Section 4, texture-based descriptors and SVM-based classification have prevailed in the face spoofing literature. The combination texture + SVM has reached the best performance in five out of the nine data sets analyzed. For the remaining two, texture and SVM are still there, but combined with other descriptors (*i.e.*, motion, frequency or reflectance).
2. Most of the time, spoofing detection follows face recognition trends. For instance, deep learning techniques are becoming very popular in face recognition (Zhi-Peng et al., 2014) and have consistently outperformed other existing methods, just like what happened to LBP + SVM few years before. As demonstrated, Menotti et al. (2015) recently employed deep learning for face

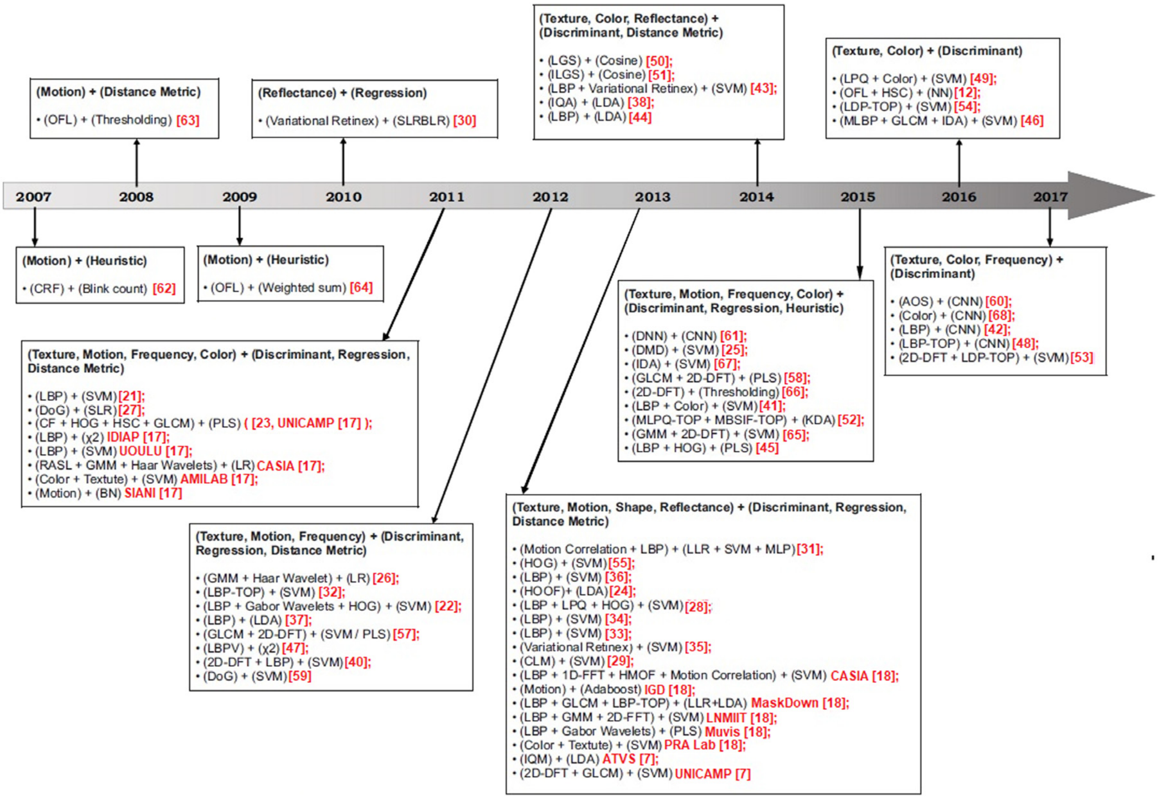
(Motion) + (Distance Metric)
• (OFL) + (Thresholding) [63]
(Reflectance) + (Regression)
• (Variational Retinex) + (SLRBLR) [30]
(Texture, Color, Reflectance) + (Discriminant, Distance Metric)
• (LGS) + (Cosine) [50];
• (ILGS) + (Cosine) [51];
• (LBP + Variational Retinex) + (SVM) [43];
• (IQA) + (LDA) [38];
• (LBP) + (LDA) [44]
(Texture, Color) + (Discriminant)
• (LPQ + Color) + (SVM) [49];
• (OFL + HSC) + (NN) [12];
• (LDP-TOP) + (SVM) [54];
• (MLBP + GLCM + IDA) + (SVM) [46]
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
(Motion) + (Heuristic)
• (CRF) + (Blink count) [62]
(Motion) + (Heuristic)
• (OFL) + (Weighted sum) [64]
(Texture, Motion, Frequency, Color) + (Discriminant, Regression, Heuristic)
• (DNN) + (CNN) [61];
• (DMD) + (SVM) [25];
• (IDA) + (SVM) [67];
• (GLCM + 2D-DFT) + (PLS) [58];
• (2D-DFT) + (Thresholding) [66];
• (LBP + Color) + (SVM) [41];
• (MLPQ-TOP + MBSIF-TOP) + (KDA) [52];
• (GMM + 2D-DFT) + (SVM) [65];
• (LBP + HOG) + (PLS) [45]
(Texture, Color, Frequency) + (Discriminant)
• (AOS) + (CNN) [60];
• (Color) + (CNN) [68];
• (LBP) + (CNN) [42];
• (LBP-TOP) + (CNN) [48];
• (2D-DFT + LDP-TOP) + (SVM) [53]
(Texture, Motion, Frequency, Color) + (Discriminant, Regression, Distance Metric)
• (LBP) + (SVM) [21];
• (DoG) + (SLR) [27];
• (CF + HOG + HSC + GLCM) + (PLS) ( [23, UNICAMP [17] );
• (LBP) + (χ2) IDIAP [17];
• (LBP) + (SVM) UOULU [17];
• (RASL + GMM + Haar Wavelets) + (LR) CASIA [17];
• (Color + Textute) + (SVM) AMILAB [17];
• (Motion) + (BN) SIANI [17]
(Texture, Motion, Frequency) + (Discriminant, Regression, Distance Metric)
• (GMM + Haar Wavelet) + (LR) [26];
• (LBP-TOP) + (SVM) [32];
• (LBP + Gabor Wavelets + HOG) + (SVM) [22];
• (LBP) + (LDA) [37];
• (GLCM + 2D-DFT) + (SVM / PLS) [57];
• (LBPV) + (χ2) [47];
• (2D-DFT + LBP) + (SVM) [40];
• (DoG) + (SVM) [59]
(Texture, Motion, Shape, Reflectance) + (Discriminant, Regression, Distance Metric)
• (Motion Correlation + LBP) + (LLR + SVM + MLP) [31];
• (HOG) + (SVM) [55];
• (LBP) + (SVM) [36];
• (HOOF)+ (LDA) [24];
• (LBP + LPQ + HOG) + (SVM) [28];
• (LBP) + (SVM) [34];
• (LBP) + (SVM) [33];
• (Variational Retinex) + (SVM) [35];
• (CLM) + (SVM) [29];
• (LBP + 1D-FFT + HMOF + Motion Correlation) + (SVM) CASIA [18];
• (Motion) + (Adaboost) IGD [18];
• (LBP + GLCM + LBP-TOP) + (LLR+LDA) MaskDown [18];
• (LBP + GMM + 2D-FFT) + (SVM) LNMIIT [18];
• (LBP + Gabor Wavelets) + (PLS) Muvis [18];
• (Color + Textute) + (SVM) PRA Lab [18];
• (IQM) + (LDA) ATVS [7];
• (2D-DFT + GLCM) + (SVM) UNICAMP [7]

**Fig. 4.** Timeline of face spoofing detection in the last decade.

spoofing detection, evaluating the performance of the proposed method on two data sets: Replay-Attack and 3D Mask Attack. Their results were comparable to the state-of-the-art in the first one, and are the best performance so far in the second one. This practically shows that any face recognition breakthroughs will lead to improvements on texture-based spoofing detection as well.

3. All surveyed works perform training and testing using the same data set (although in a non-overlapping way). They presented their results with different metrics (*i.e.* ACC, AUC, HTER and EER) and near perfect results were found for each of the nine publicly available data sets considered in this work. Far from showing that face spoofing detection is a solved problem, this fact actually indicates the lack of a challenging data set that allows a thoroughly analysis of the proposed methods. Other computer vision problems have been conducted in this direction, like person re-identification with VIPER data set (Ma et al., 2015) and object recognition with Caltech-256 Object Category data set (Griffin et al., 2007), both with state-of-the-art accuracy below 50%. We believe that a large data set in a wild scenario is more likely to promote breakthroughs. In addition to a large amount 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.
4. A lack of a standard evaluation protocol for spoofing detection methods is also an issue. Currently, most of the researchers use HTER and EER for detection results to avoid biased results when a data set is unbalanced, but these metrics do not show the effects of spoofing detection on the recognition step. Chingovska et al. (2014) introduced an evaluation protocol for biometric systems under spoofing attacks that simultaneously analyzes both recognition and spoofing detection results through expected performance and spoofability curves (EPSC) by dividing a data set in three categories: genuine users, zero-effort impostors and spoofing attacks. However, the proposed evaluation method depends on a prior probability of the spoofing attacks, or a cost relation between the ratio of incorrectly accepted zero-effort impostors and the ratio of incorrectly accepted spoofing attacks. The latter ones could vary for different systems, adding more variables to the problem. Hence, a more intuitive and self-explanatory evaluation metric is also required to instigate future efforts in this research topic. This also extends to benchmarks that rigorously evaluate both recognition and spoofing detection, which are currently not available in the literature.

In general, existing works seem to be going towards data set tuning (*i.e.* overfitting) instead of designing more effective and flexible solutions. This is corroborated by the works of de Freitas Pereira et al. (2013) and Pinto et al. (2015a), which show initial cross-data set performance analyses using Casia Face Anti-Spoofing and Replay-Attack data sets. Different methods were evaluated by de Freitas Pereira et al. (2013), and Tables 15 and 16 show the results for the two most interesting ones, respectively, (1) Motion Correlation + MLP; and (2) LBP-TOP + SVM. While LBP-TOP + SVM presents the best performance in experiments within a data set, Motion Correlation + MLP performs better in experiments across different data sets, which seems to indicate that not necessarily the best performing works for a specific data set, like the ones shown in Table 14, are actually the best countermeasures. On the other hand, they probably miss in terms of generalization power. Similar results can be also found in Pinto et al. work (Pinto et al., 2015a). Therefore, countermeasures with good performance in cross-data set experiments – in the absence of a truly challenging data set – are expected to be more effective in real world scenarios. Current countermeasures, however, hardly beat a random classifier (*i.e.* 50% HTER).

**Table 14**
Best performing works over different data sets.

| Reference | Features | Classifier | Data set |
|---|---|---|---|
| Maatta et al. (2012) | LBP + Gabor Wavelets + HOG | SVM | NUAA Imposter<br>Yale Recaptured<br>Print-Attack |
| IDIAP (Chakka, 2011) | LBP | $\chi^2$ | Print-Attack |
| UOULU (Chakka, 2011) | LBP | SVM | Print-Attack |
| CASIA (Chakka, 2011) | RASL + GMM + Haar wavelets | LR | Print-Attack |
| Tirunagari et al. (2015) | DMD + LBP | SVM | Print-Attack |
| CASIA (Chingovska, 2013) | LBP + 1D-FFT + HMOF + Motion Correlation | SVM | Replay-Attack |
| LNMIIT (Chingovska, 2013) | LBP + GMM + 2D-FFT | SVM | Replay-Attack |
| Feng et al. (2016) | HSC + Optical Flow | NN | Replay-Attack |
| Boulkenafet et al. (2016) | LPQ + Color | SVM | Casia Face Anti-Spoofing |
| Kose and Dugelay (2014) | LBP + Variational Retinex | SVM | Kose and Dugelay's |
| Menotti et al. (2015) | DNN | CNN | 3D Mask Attack |
| Feng et al. (2016) | HSC + Optical Flow | NN | 3D Mask Attack |
| Kim et al. (2012a) | MLBP + GLCM + IDA | SVM | MSU-MFSD |

**Table 15**
Cross-data set results (HTER) for Motion Correlation + MLP, as presented by Wen et al. (2015).

| Train | Test | |
|---|---|---|
| | Casia face anti-spoofing | Replay-Attack |
| Casia face anti-spoofing | 30.33% | 50.25% |
| Replay-Attack | 48.28% | 11.79% |

**Table 16**
Cross-data set results for LBP-TOP + SVM in HTER, as presented by Wen et al. (2015).

| Train | Test | |
|---|---|---|
| | Casia face anti-spoofing | Replay-Attack |
| Casia face anti-spoofing | 23.75% | 50.64% |
| Replay-Attack | 61.33% | 8.51% |

### 4.3. Future perspectives

Given the actual state of researches in face spoofing detection and the observed trends, we would like to point some future directions that could help other authors to address challenges that still need to be solved.

First, although texture-based solutions imported from face recognition systems have the best results in experiments within a data set, their performance rapidly degrade in experiments across different data sets (de Freitas Pereira et al., 2013). Thus, designing solutions specifically for spoofing detection like the initial works based on motion and reflectance seems to be a more promising way of achieving reasonable generalization. This topic has being understudied in the last years, but can find new stimuli in unexplored variations of deep learning that may benefit from this kind of information, such as long short-term memory networks (Hochreiter and Schmidhuber, 1997) and Fourier CNNs (Pratt et al., 2017).

Second, other learning frameworks could be explored to offer a different perspective on how to solve this problem. Principles of life-long (Fischer, 2000) and transfer learning (Yu et al., 2014) have not been explored so far. Such techniques would allow incorporating new samples into an existing model at any time, making it more flexible to cover further attacks in the future without retraining the entire classifier. In addition, clustering approaches (Cornujols et al., 2018) may be an option to analyze massive amounts of data in unsupervised or semi-supervised ways and could help to eventually discover unknown attacks without exhaustive manual annotation.

Third, a large web collected corpus for spoofing detection in uncontrolled scenarios would give an immediate boost to this field and would reduce overfitting problems related to data sets and/or attack types. More than that, this corpus could be created as extension of existing wild face recognition databases, such as Labeled Faces in the Wild (Learned-Miller et al., 2016), to allow evaluating both recognition and spoofing detection simultaneously. To this end, one may search the web looking for images of individuals from of a chosen face recognition data set containing printed faces or even elaborate attacks like silicone masks and makeup disguises.

Finally, multimodal biometric systems are less likely to be spoofed as impostors, since one has to forge multiple biometric features at the same time. For this reason, different works addressed the impostor problem by combining two or more human characteristics (Akhtar et al., 2012; Biggio et al., 2012; Johnson et al., 2010; Rodrigues et al., 2009, 2010; Farmanbar and Toygar, 2017b, a). With this in mind, facial biometrics can be seen as a special case, since multimodality can take advantage of multiple facial properties (*e.g.,* texture, shape and temperature) to avoid spoof attacks. Nowadays, different commercially available devices are able to capture color, depth and infrared images simultaneously at a reasonable price. These devices could be used to enhance current countermeasures, and possibly make them practicable in industrial applications (Litomisky, 2012).

## 5. Conclusion

In this survey, we presented a compilation of face spoofing detection works over the past decade, as well as, a thoroughly numerical and qualitative analysis. Spoofing attacks persist to be a security challenge for face biometric systems, and there were much effort in the field to find robust methods. However, all these efforts have been following the same recipe, not favoring breakthroughs in the field. Many works of face spoofing detection give emphasis on 2D attacks by presenting printed photos or replaying recorded videos, and 3D attacks have been recently studied due to the technological advancements in 3D printer and reconstruction. Although perfect results on public data sets have been achieved by many works, there is a considerable gap to move from academic researching to real-world applications in a effective way. As such, it is expected that researchers concentrate efforts to create more difficulty data sets and more unbiased evaluation methods, henceforth.

## Acknowledgments

J.P. Papa is grateful to FAPESP grants #2013/07375-0, #2014/12236-1, and #2016/19403-6, as well as CNPq grant #306166/2014-3.

## References

Akhtar, Z., Fumera, G., Marcialis, G.-L., Roli, F., 2012. Evaluation of serial and parallel multibiometric systems under spoofing attacks. In: Proc. BTAS, pp. 283–288.

Al Eidan, R.M., 2013. Hand biometrics: Overview and user perception survey. In: Proc. ICIA, pp. 252–257.

Alotaibi, A., Mahmood, A., 2017. Deep face liveness detection based on nonlinear diffusion using convolution neural network. Signal Image Video Process. 11 (4), 713–720 Springer.

Anjos, A., Marcel, S., 2011. Counter-measures to photo attacks in face recognition: A public data set and a baseline. In: Proc. IJCB, pp. 1–7.

Arashloo, S.R., Kittler, J., Christmas, W., 2015. Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features. IEEE TIFS 10 (11), 2396–2407.

Asim, M., Ming, Z., Javed, M.Y., 2017. CNN based spatio-temporal feature extraction for face anti-spoofing. In: Proc. ICIVC, pp. 234–238.

Bharadwaj, S., Dhamecha, T.I., Vatsa, M., Singh, R., 2013. Computationally efficient face spoofing detection with motion magnification. In: Proc. CVPRW, pp. 105–110.

Biggio, B., Akhtar, Z., Fumera, G., Marcialis, G.L., Roli, F., 2012. Security evaluation of biometric authentication systems under real spoofing attacks. IET Biom. 1 (1), 11–24.

Boulkenafet, Z., Komulainen, J., Hadid, A., 2015. Face anti-spoofing based on color texture analysis. In: Proc. ICIP, Quebec City, QC, pp. 2636–2640.

Boulkenafet, Z., Komulainen, J., Hadid, A., 2016. Face spoofing detection using colour texture analysis. IEEE TIFS 11 (8), 1818–1830.

Chakka, M.M., et al., 2011. Competition on counter measures to 2-d facial spoofing attacks. In: Proc. IJCB, pp. 1–6.

Chingovska, I., et al., 2013. The 2nd competition on counter measures to 2D face spoofing attacks. In: Proc. ICB, pp. 1–6.

Chingovska, I., Anjos, A., Marcel, S., 2012. On the effectiveness of local binary patterns in face anti-spoofing. In: Proc. BIOSIG, pp. 1–7.

Chingovska, I., Anjos, A.R., Marcel, S., 2014. Biometrics evaluation under spoofing attacks. IEEE TIFS 9 (12), 2264–2276.

Choi, D.J., Park, J.S., Oh, Y.H., 2015. Unsupervised rapid speaker adaptation based on selective eigenvoice merging for user-specific voice interaction. Eng. Appl. Artif. Intell. 40 (1), 95–102 Elsevier.

Cornujols, A., Wemmert, C., Ganarski, P., Bennani, Y., 2018. Collaborative clustering: Why, when, what and how. In: Information Fusion, Vol. 39. Elsevier, pp. 81–95.

de Freitas Pereira, T., Anjos, A., De Martino, J.M., Marcel, S., 2013. Can face anti-spoofing countermeasures work in a real world scenario? In: Proc. ICB, Madrid, pp. 1–8.

da Silva Pinto, A., 2013. A countermeasure method for video-based face spoofing attacks. (M.S. thesis), In: Inst. Comput.. UNICAMP Univ. Estadual Campinas, Campinas, Brazil.

de Souza, G.B., da Silva Santos, D.F., Pires, R.G., Marana, A.N., Papa, J.P., 2017. Deep texture features for robust face spoofing detection. IEEE TCS 64 (12), 1–5.

Dora, L., Agrawal, S., Panda, R., Abraham, A., 2017. An evolutionary single Gabor kernel based filter approach to face recognition. Eng. Appl. Artif. Intell. 62 (1), 286–301 Elsevier.

Erdogmus, N., Marcel, S., 2013. Spoofing in 2D face recognition with 3D masks and anti-spoofing with kinect. In: Proc. BTAS, pp. 1–6.

Erdogmus, N., Marcel, S., 2014. Spoofing face recognition with 3D Masks. IEEE TIFS 9 (7), 1084–1097.

Fan, H., Cao, Z., Jiang, Y., Yin, Q., Doudou, C., 2014. Learning deep face representation. In: Proc. ICCV, pp. 1–10.

Farmanbar, M., Toygar, O., 2017a. A robust anti-spoofing technique for face liveness detection with morphological operations. In: Optik - International Journal for Light and Electron Optics, Vol. 139. Elsevier, pp. 347–354.

Farmanbar, M., Toygar, O., 2017b. Spoof detection on face and palmprint biometrics. Signal, Image and Video Process. 11 (7), 1253–1260 Springer.

Feng, L., Po, Lai-Man, Li, Y., Xu, X., Yuan, F., Chun-Ho Cheung, T., Cheung, Kwok-Wai, 2016. Integration of image quality and motion cues for face anti-spoofing: A neural network approach. JVCIR 38, 451–460.

Fischer, G., 2000. Lifelong learning —more than training. J. Interact. Learn. Res. 11 (3/4), 265–294.

Galbally, J., Marcel, S., 2014. Face anti-spoofing based on general image quality assessment. In: Proc. ICPR, pp. 1173–1178.

Galbally, J., Marcel, S., Fierrez, J., 2014. Biometric antispoofing methods: A survey in face recognition. IEEE Access 1530–1552.

Garcia, D.C., de Queiroz, R.L., 2015. Face-Spoofing 2D-detection based on Moiré-Pattern analysis. IEEE TIFS 10 (4), 778–786.

Griffin, G., Holub, A., Perona, P., 2007. Caltech-256 Object Category Dataset. California Institute of Technology.

Haralick, R., Shanmugam, K., Dinstein, I., 1973. Texture features for image classification. IEEE TSMC 3 (6), 610–621.

Hasan, H., Abdul-Kareem, S., 2013. Fingerprint image enhancement and recognition algorithms: a survey. Neural Comput. Appl. 23 (6), 1605–1610 Springer.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

Housam, K.B., Lau, S.H., Pang, Y.H., Liew, Y.P., Chiang, M.L., 2014a. Face spoofing detection based on improved local graph structure. In: Proc. ICISA, pp. 1–4.

Housam, K.B., Lau, S.H., Pang, Y.H., Liew, Y.P., Chiang, M.L., 2014b. Face spoofing detection using local graph structure. In: Proc. ICISA, pp. 270–273.

Huang, W., Nakamori, Y., Wang, S.-Y., 2005. Forecasting stock market movement direction with support vector machine. Comput. Oper. Res. 32 (10), 2513–2522.

Jain, A.K., Flynn, P., Ross, A.A., 2008. Handbook of Biometrics. Springer.

Johnson, P., Tan, B., Schuckers, S., 2010. Multimodal fusion vulnerability to non-zero effort (spoof) imposters. In: Proc. WIFS, pp. 1–5.

Kah Ong Michael, G., Connie, T., Beng Jin Teoh, A., 2012. A contactless biometric system using multiple hand features. JVCIR 23, 1068–1084.

Kim, I., Ahn, J., Kim, D., 2016a. Face spoofing detection with highlight removal effect and distortions. In: Proc. SMC, pp. 4299–4304.

Kim, G., Eum, S., Suhr, J.K., Kim, D.I., Park, K.R., Kim, J., 2012b. Face liveness detection based on texture and frequency analyses. In: Proc. ICB, pp. 67–72.

Kollreider, K., Fronthaler, H., Bigun, J., 2008. Verifying liveness by multiple experts in face biometrics. In: Proc. CVPRW, pp. 1–6.

Kollreider, K., Fronthaler, H., Bigun, J., 2009. Non-intrusive liveness detection by face images. In: Image and Vision Computing, Vol. 27. Elsevier, pp. 233–244.

Komulainen, J., Hadid, A., Pietikainen, M., 2013a. Context based face anti-spoofing. In: Proc. BTAS, pp. 1–8.

Komulainen, J., Hadid, A., Pietikainen, M., Anjos, A., Marcel, S., 2013b. Complementary countermeasures for detecting scenic face spoofing attacks. In: Proc. ICB, pp. 1–7.

Kose, N., Dugelay, J.-L., 2012. Classification of captured and recaptured images to detect photograph spoofing. In: Proc. ICIEV, pp. 1027–1032.

Kose, N., Dugelay, J.-L., 2013a. Countermeasure for the protection of face recognition systems against mask attacks. In: Proc. FG, pp. 1–6.

Kose, N., Dugelay, J.-L., 2013b. Reflectance analysis based countermeasure technique to detect face mask attacks. In: Proc. DSP, pp. 1–6.

Kose, N., Dugelay, J.-L., 2013c. Shape and texture based countermeasure to protect face recognition systems against mask attacks. In: Proc. CVPRW, pp. 111–116.

Kose, N., Dugelay, J.-L., 2014. Mask spoofing in face recognition and countermeasures. In: Pattern Recognition, Vol. 32. Elsevier, pp. 779–789.

Kotsia, I., Pitas, I., 2007. Facial expression recognition in image sequences using geometric deformation features and support vector machines. IEEE TIP 16 (1), 172–187.

Lakshminarayana, N.N., Narayan, N., Napp, N., Setlur, S., Govindaraju, V., 2017. A discriminative spatio-temporal mapping of face for liveness detectio. In: Proc. ISBA, pp. 1–7.

Learned-Miller, E., Huang, G.B., RoyChowdhury, A., Li, H., Hua, G., 2016. Labeled faces in the wild: A survey. In: Advances in Face Detection and Facial Image Analysis. Springer, pp. 189–248.

Litomisky, K., 2012. Consumer Rgb-D Cameras and their Applications, Rapport Technique. University of California.

Ma, B., Li, Q., Chang, H., 2015. Gaussian descriptor based on local features for person re-identification. In: Proc. ACCV, pp. 505–518.

Maatta, J., Hadid, A., Pietikainen, M., 2011. Face spoofing detection from single images using micro-texture analysis. In Proc. IJCB, pp. 1–7.

Maatta, J., Hadid, A., Pietikainen, M., 2012. Face spoofing detection from single images using texture and local shape analysis. IET Biom. 1 (1), 3–10.

Marasco, E., Ross, A., 2015. A survey on antispoofing schemes for fingerprint recognition systems. ACM Comput. Surv. 47 (2), 28.

Meadowcroft, P., 2008. Card fraud - will PCI-DSS have the desired impact? Card Technol. Today 20 (3), 10–11.

Menotti, D., Chiachia, G., Pinto, A., Schwartz, W.R., Pedrini, H., Falcão, A.X., Rocha, A., 2015. Deep representations for iris, face, and fingerprint spoofing detection. IEEE TIFS 10 (4), 864–879.

Muralidharan, R., Chandrasekar Dr., C., 2011. Object recognition using support vector machine augmented by RST invariants. IJCSI 8 (5), 280–286.

Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on feature distributions. Pattern Recognit. 29 (1), 51–59 Elsevier.

Osuna, E., Freund, R., Girosi, F., 1997. Training support vector machines: an application to face detection. In Proc. CVPR, pp. 130–136.

Pan, G., Sun, L., Wu, Z., Lao, S., 2007. Eyeblink-based anti-spoofing in face recognition from a generic webcamera. In: Proc. ICCV, pp. 1–8.

Parveen, S., Ahmad, S.M.S., Hanafi, M., Adnan, W.A.W., 2015. Face anti-spoofing methods. Current Sci. 108 (8), 1491–1500.

Peixoto, B., Michelassi, C., Rocha, A., 2011. Face liveness detection under bad illumination conditions. In: Proc. ICIP, pp. 3557–3560.

Peralta, D., Galar, M., Triguero, I., Miguel-Hurtado, O., Benitez, J.M., Herrera, F., 2014. Minutiae filtering to improve both efficacy and efficiency of fingerprint matching algorithms. Eng. Appl. Artif. Intell. 32 (1), 37–53.

Pereira, T.F., Anjos, A., Martino, J.M., Marcel, S., 2012. LBP-TOP based countermeasure against facial spoofing attacks. In: Proc. ACCV, pp. 121–132.

Phan, Q.T., Dang-Nguyen, D.T., Boato, G., De Natale, F.G.B., 2016. FACE spoofing detection using LDP-TOP. In: Proc. ICIP, pp. 404–408.

Phan, Q.T., Dang-Nguyen, D.T., Boato, G., De Natale, F.G.B., 2017. Using LDP-TOP in video-based spoofing detection. In: Proc. ICIAP, pp. 614–624.

Pinto, A.S., Pedrini, H., Schwartz, W.R., Rocha, A., 2012. Video-based face spoofing detection through visual rhythm analysis. In: Proc. SIBGRAPI, pp. 221–228.

Pinto, A., Pedrini, H., Schwartz, W.R., Rocha, A., 2015a. Face spoofing detection through visual codebooks of spectral temporal cubes. IEEE TIP 24 (12), 4726–4740.

Pinto, A., Schwartz, W.R., Pedrini, H., Rocha, A.R., 2015b. Using visual rhythms for detecting video-based facial spoof attacks. IEEE TIFS 10 (5), 1025–1038.

Pratt, H., Williams, B., Coenen, F., Zheng, Y., 2017. FCNN: Fourier convolutional neural networks. In: Proc. ECML.

Rapp, V., Senechal, T., Bailly, K., Prevost, L., 2011. Multiple kernel learning SVM and statistical validation for facial landmark detection. In: Proc. FG, pp. 265–271.

Rodrigues, R.N., Kamat, N., Govindaraju, V., 2010. Evaluation of biometric spoofing in a multimodal system. In: Proc. BTAS, pp. 1–5.

Rodrigues, R.N., Ling, L.L., Govindaraju, V., 2009. Robustness of multimodal biometric fusion methods against spoof attacks. J. Vis. Lang. Comput. 20 (3), 169–179.

Sanmorino, A., Yazid, S., 2012. A survey for handwritten signature verification. In: Proc. URKE, pp. 54–57.

Schwartz, W.R., Rocha, A., Pedrini, H., 2011. Face spoofing detection through partial least squares and low-level descriptors. In: Proc. IJCB, pp. 1–8.

Sweilam, N.H., Tharwat, A.A., Moniem, N.K.A., 2010. Support vector machine for diagnosis cancer disease: A comparative study. Egyptian Inform. J. 11 (2), 81–92.

Taigman, Y., Yang, M., Ranzato, M., Wolf, L., 2011. DeepFace: Closing the gap to human-level performance in face verification. In: Proc. CVPR, pp. 1521–1528.

Tamrakar, D., Khanna, P., 2016. Kernel discriminant analysis of Block-wise Gaussian Derivative Phase Pattern Histogram for palmprint recognition. JVCIR 40, 432–448.

Tan, X., Li, Y., Liu, J., Jiang, L., 2010. Face liveness detection from a single image with sparse low rank bilinear discriminative model. In; Proc. ECCV, pp. 504–517.

Tefas, A., Kotropoulos, C., Pitas, I., 2001. Using support vector machines to enhance the performance of elastic graph matching for frontal face authentication. IEEE TPAMI 23 (7), 735–746.

Tirunagari, S., Poh, N., Windridge, D., Iorliam, A., Suki, N., Ho, A.T.S., 2015. Detection of face spoofing using visual dynamics. IEEE TIFS 10 (4), 762–777.

Torralba, A., Efros, A., 2011. Unbiased look at dataset bias. In: Proc. CVPR, pp. 1521–1528.

Wang, T., Yang, J., Lei, Z., Liao, S., 2013. Face liveness detection using 3D structure recovered from a single camera. In: Proc. ICB, pp. 1–6.

Wen, D., Han, H., Jain, A.K., 2015. Face spoof detection with image distortion analysis. IEEE TIFS 10 (4), 746–761.

Yadav, K.S., Mukhedkar, M.M., 2013. Review on speech recognition. IJSE 1 (2), 61–70.

Yan, J., Zhang, Z., Lei, Z., Yi, D., Li, S.Z., 2012. Face liveness detection by exploring multiple scenic clues. In: Proc. ICARCV, pp. 188–193.

Yang, J., Lei, Z., Liao, S., Li, S.Z., 2013. Face liveness detection with component dependent descriptor. In: Proc. ICB, pp. 1–6.

Yang, J., Lei, Z., Yi, D., Li, S.Z., 2015. Person-specific face antispoofing with subject domain adaptation. IEEE TIFS 10 (4), 797–809.

Yu, H., Chen, Y., Liu, J., Jiang, X., 2014. Lifelong and fast transfer learning for gesture interaction. J. Inf. Comput. Sci. 11 (4), 1023–1035.

Zhang, Z., Yan, J., Liu, S., Lei, Z., Yi, D., Li, S., 2012. A face antispoofing data set with diverse attacks. In: Proc. ICB, pp. 26–31.

Zhang, C., Zhang, Z., 2014. Improving multiview face detection with multi-task deep convolutional neural networks. In: Proc. WACV, pp. 1036–1041.

Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A., 2003. Face recognition: A literature survey. ACM Comput. Surv. 35 (4), 399–458.

Zhi-Peng, F., Yan-Ning, Z., Hai-Yan, H., 2014. Survey of deep learning in face recognition. In: Proc. ICOT, pp. 5–8.