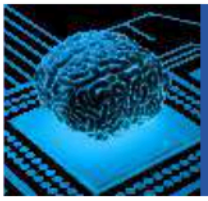

Sountra et al LGU (IJECI) 2019

LGU (IJECI)
ISSN: 2522-3429 (Print)
ISSN: 2616-6003 (Online)

LGU International Journal for
Electronic Crime Investigation

Research Article | Vol. 4 Issue 2, April - June 2020

A Literature Review of 2D Face Recognition Systems: Attacks and Defenses

Muhammad Shairoze Malik*
13beemmalik@seecs.edu.pk
National University of Science
and Technology Islamabad

Abstract:

In a world where biometric identification and authentication is increasingly growing, it is critical for IT Security personnel to have an understanding of how these systems they use work, and weaknesses that lay with them. Face recognition is a very common form of biometric identification and authentication, where the face of the user is used to both identify and authenticate the user. This paper serves as a literature review for a preliminary understanding of face recognition systems. First the basics of how a facial recognition system works is reviewed. After this, selected articles demonstrating a variety of facial recognition spoofing attacks are summarized to demonstrate the attack surface. This is then followed by a review of studies conducted for defenses against these attacks, focusing primarily on the key concepts used to help identify spoofing attempts. This paper then concludes by summarizing the findings, and explaining the importance of understanding these different aspects of facial recognition from a business perspective.

Keywords: Facial Recognition, Biometrics Attacks, Biometric Defenses

1. Introduction:

Biometrics is one of the most commonly used methods for identification and authentication that we have in use today. To humans this comes as second nature, as we use faces, voices, and other biometric features of an individual to identify them in our everyday lives. Thus, the inclination to adopt these methods of who we are to help authenticate us in the digital world has become very popular. That being said the system when converted, just as in real life, is not perfect. Most of these identification and authentication methods are prone to false positives where an individual is able to spoof the identity of another through a variety of means.

In this paper we explore different methods of attacking and protecting against facial recognition attacks. Facial recognition has also been around for quite some time and grown in popularity due to the uniqueness of each individuals' facial features [1]. However, over time attacker have reverse engineered the methodologies used to extract these unique features of an individual's face, so that they can then spoof these features to bypass the authentication method. In return, there have also been several attempts to mitigate the success rates of these attacks against facial recognition systems through a variety of means.

2. Basics of Facial Recognition:

Facial recognition in digital applications

involves the selection and identification of key unique characteristics. At high level, this is completed using algorithms to calculate distance between different facial features and create a uniquely identifiable profile or "face print" for different users of the system [1].

Every face has numerous, distinguishable landmarks, the different peaks and valleys that make up facial features. Facial recognition system defines these landmarks as nodal points. Each human face has approximately 80 nodal points. Some of these measured by the software are [1]:

- Distance between the eyes
- Width of the nose
- Depth of the eye sockets
- The shape of the cheekbones
- The length of the jaw line

These nodal points are measured creating a numerical code, called a face print, representing the face in the database. An example of this can be seen in figure 1, where the nodes and lines are used to visualize different the different vectors used to construct that face print.

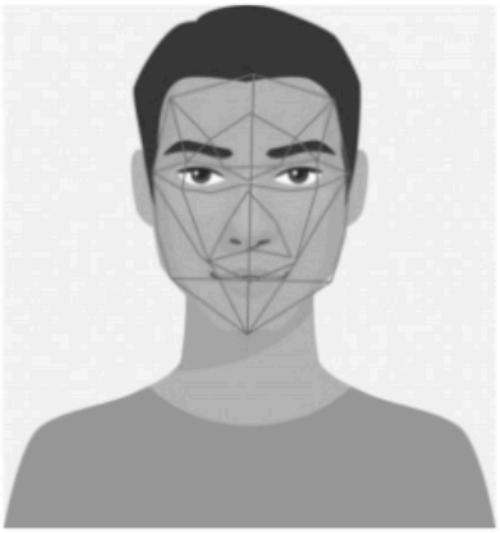

Figure 1. Visualization of a Face Print Construct [2]

After the basics of creating a face print another fundamental concept to understand about facial recognition is the process used to complete the authentication or identification.

A prerequisite step to facial recognition is creating a database of users. This database would contain a collection of images of each user, to be used when completing the verification. Once the users have been enrolled in the system, its usage can begin.

At high level the following is the process for facial recognition [1]:

I. Face detection: in either an image or video the detection of the face being processed.
II. Face alignment: once the face has been detected, it has to be aligned in comparison to the images stored.
III. Feature extraction: extract the unique nodal points in creating the face print.
IV. Face matching: comparing the extracted features against the registered users in the database for either identification or authentication, or both.

Of course, there is far more details and variations that can occur in this process when considering specific implementations, however the basic concepts remain the same. The basis provided in this section should be sufficient to understand the usage of the attacks and defenses shown in the remainder of the paper.

1. Facial Recognition Attacks:

Spoofing attack is the demonstration of introducing a fake biometric proof to a framework with the end goal to accomplish verification. Forging such an attack is generally simple for facial recognition systems since the photos or recordings of authorized users can be downloaded from internet or taken from a distance. Attackers can get access of system by showing printed photographs or replaying recorded recordings before the sensors. Since these are the most well-known, easiest and least expensive strategies to evade face recognition systems [3].

3.1 2D Photo Attack

In this attack fraudulent show hard copy photograph of authorized user to facial recognition system to gain the access. The photograph may have been taken earlier by the fraudulent or may be victim shared his photo on

social network site. Attacker can create 2D mask (a high-resolution photograph prints out) as he gets the victim photo. The created photo or mask is then shown on the screen of the devices that could be a mobile or tablet to bypass authentication. The said attack is performed using the help of digital camera to create high-resolution picture of victims to bypass the recognition system and to gain the secured access [4].

3.2 2D Video Attack

In this attack fraudulent using video clips instated of photo image. This technique is more advance and realistic compare to still images as this capture the motions valid user. Attacker will use video clip in front of facial recognition device to make it difficult for system to identify between fraudulent and valid user. This way attacker bypasses the countermeasures because video clips are more efficient and realistic against photos [4].

3.3 3D Mask Attack

This attack is one more step ahead from 2D photographic and 2D video attacks as fraudulent wearing a 3D mask and imitating the actual personal face.

A mask is an object on the face, use for entertainment or masks can also be used for spoofing intention. There are different methods for designing masks. Mask of a person can be designed even by using a paper. Organizations have standard 3D face models against every race or origin and masks are made by mapping one frontal and one profile photo of the target person on this model. The model depends on ethnic shape, so it does not show exact 3D face traits of the target person. The mask used by attacker for 3D face spoofing purposes needs to be exactly same or similar to 3D face shape characteristics of the victim to carry out successful attack [3].

Researchers focus on studying face spoofing attacks that are done by using different types of masks. For this they find advance way of designing 3D masks out of printed 2D face images. There is a service named ThatsMyFace.com that is used for facial reformation and to transform 2D image to 3D shape. To get 3D sculpture of a person, one just needs to upload 2D mugshots (one frontal and one profile) of a person to ThatsMyFace.com service and he would get it in his mail box as a nice picture or a toy or as wearable mask that can easily be used for attack [3].

For research purpose, four types of masks from ThatsMyFace.com are used. First type M1 is a paper make design of face that is printed, cut, folded and glued together to create 3D face mask. The other three types (M2, M3, M4) of masks are designed of a hard resin composite. M2 is ½ of real size of face and other two M3 and M4 are of real size. M4 is photographed in better lightening conditions with respect to M3 and it is real life-size mask with holes at the eyes and nostrils. Beside these masks, researcher take two more mask types to analyze. M5 is printed from a real 3D scan of a person and last type M6 is made of silicon that is created using face shape and painted realistically. Researchers use open source framework with two Gabor wavelet-based algorithms to test 3D face masks spoofing performance.3D Mask database was designed to study this. For testing purpose, database have 4 subjects with 4 gallery samples. Different types of masks (M1, M2, M3, M4, M5, M6) for each subject are designed in order to perform attack [3].

3.4 Spoofing performances of 3D Masks:

The likeness scores for masks evaluated separately. For M1, which is paper based mask have very low chances to spoof recognition system as it has distorted and noisy texture and patch edges are clearly visible. M2 is ½ of real face size. Scale shows that M2 mask reach 78.12% success rate which is very close to photo attacks. Attacker can use this mask to attack successfully. According to Scale, M3-realistic size mask does not meet the expectations and perform inadequately. Both M1 and M3 depends on the input image, quality and correctness of end product to reconstruct 3D shape [3].

M4 is the most successful attack among all types, reaching 100% success because it is taken under ideal conditions with respect to other masks. M5 has exact shape of user face as it is printed directly from a real face scan using 3D printer so it is different from previous masks and it achieved high attack success rate in both 3D and 2D recognition systems.M6 mask is made of silicon and it has least attack success rate. Silicon

masks do not manage to break into system. Silicon mask attacks are ineffective and impractical. In order to have a silicon mask of authorized user, attacker needs to know the 3D shape of user and moreover attacker requires 3D printing skills to create realistic mask. This evaluation cleared that spoofing face recognition system with 3D masks is highly dependent on degree of accuracy [3].

3D Face Recognition System is introduced in counter defense of spoofing against 2D Face Recognition System but 3D Face Recognition System is also vulnerable to face spoofing attacks. 3D Face Recognition system can be spoofed using masks of type M4 [3].

3.5 Replay Attack:

Reply attacks has been used against low-end finger scanners, iris scanners and facial recognition systems. Researchers focus on replay attack on biometric face recognition system. Replay attack possibility comes once the biometric digital signatures are not validated. Digital Signature defines when biometric signals were sampled and they are not taken before the sample date. Biometric signals do not have enough data to perform appropriate authentication [5].

Consumer devices like mobile phone, tablets do not have intelligent biometric sensors and also users do not have much knowledge to preform transactions with security processes. People leave traces of their biometric on the public places or offices like finger print, DNA, facial photos. Security is compromised by these facts. If system is compromised during communication or activity of system is captured by attacker, this can be used at later time. If attacker get access of the system, he can use some other individual's biometric data to masquerade with another individual. Attacker can also automate fraudulent biometric data by using the victim computer [5].

3.6 Replay Attack Demonstration against Face Recognition System:

In this demonstration researchers assumed that attacker have access of system and he replaced video clips or images with his own video data or images. The demonstration was performed against a desktop face recognition system. System was configured to use either the physical camera or virtual webcam to authenticate user. The demonstration attempted to authenticate user to the lock screen is in four phases:

I. Use a printed photo of a face to the physical camera.
II. Replay a video of the printed face to the virtual webcam.
III. Use a real face to the physical camera.
IV. Replay a video of a live face to the virtual webcam.

Table 1: Demonstrated Results

	Physical Camera	Virtual Webcam
Printed Face	Not Authenticate	Not Authenticate
Real Face	Authenticated	Authenticated

Outcome: Above table shows the results of demonstration. It can be seen that the system would not authenticate the user for phase 1 and 2 using printed photo and replay video of printed face. Phase3 is expected as a normal use case for FRS (facial recognition system). However, in phase4, user video which was taken in past is replayed to FRS using the virtual webcam, system will authenticate that user, even user is imposter and showing false video [5].

3.7 Invisible Mask Infrared Attack:

Infrared light has a longer wavelength compare to the normal light. Infrared light cannot be seen by human eyes due to its longer wavelength and it can be produced with infrared LEDs available in market. Infrared can be captured by camera sensors as they are built on three types of units R, G, B which are sensitive to red, green, blue colors. Camera's which do not have infrared unit produce a very different image of the person that is exposed to infrared light from what people see. This attack considering two attacking possibilities that are possible to take place in our daily life. First, running away from surveillance system by dodging technique and second is bypassing the authentication system. Infrared technique is invisible for human eyes. It is also called Invisible Mask Attack (IMA) in the author research [6].

A criminal can easily escape from surveillance

camera, if he knows about the working way of surveillance system, dodging techniques and about face searching patterns. Video face searching can be done by splitting the video into frames, so that the searching can happen on images. Images are prepared to extract the face part. In the pre-processing stage, images are sent to a land marking model that recognize set of land mark points of each face on the image, land mark points of the face are identified and cropped out based on the positions of the land mark points for later use. Image as input is then added to a face embedding model that converts it into a fixed length vector for future searching. To dodge face searching, an attacker can either fail the land marking model or avoid the embedding model with enough amount of infrared light on the face to bypass the face land marking points [6].

Researchers designed a dodging device that have infrared LED and is small in size so that it can easily fit in cap and can be mounted on the person head to emit enough infrared to fail the pre-processing step. This way attacker no longer needs to worry about being identified. See Figure 2.

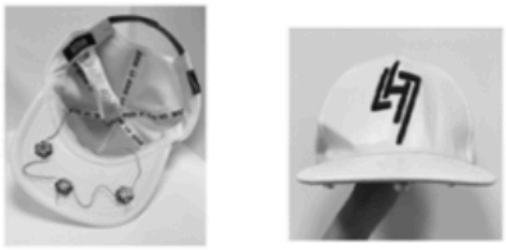

Figure 2. Retrofitted LED Cap for evading detection

To bypass authentication, attacker needs at least one victim photo and attacking device designed by researcher (cap with infrared LEDs). This infrared light produces dots on attacker face wearing infrared cap. This way light alters the attacker face features and provide a wrong classification to facial recognition system. By improving the size, positions, strengths of the dots, attacker get the image that is likely recognized by embedding model to impersonate a different person. This way attacker can avoid the detection system and by pass the facial recognition-based authentication system [6].

1. Facial Recognition Spoofing Defenses:

As facial recognition has become a commonly used solution for many authentication platforms, ensuring its validity is critical to limit attacks to the systems being protected by such a method. As such in this section we review selected articles to better understand the different existing protection methods studied and tested against facial recognition spoofing.

The articles presented cover a diverse range of topics to be explored. The first paper competes 6 different countermeasures against facial recognition spoofing [7]. This article was the primary focus in this section as it provided a high-level overview of the exact information being sought after. The second article investigates the use a new algorithmic technique to be used in a multimodal face and fingerprint-based authentication system [8]. The last article looks at reviews current motion-based counter measures against spoofing and suggests an improved technique [9].

A. Competition on Counter Measures to 2-D Facial Spoofing Attacks

This paper, by a diverse team of students from several universities, explores different counter measures against spoofing attacks and compares them to one another [7]. The competition conducted was in regards to only 2-D spoofing attacks and had clear rules to ensure consistency between tests. For video input a publicly available database was used, and for hard copy images a color laser printed image was printed on a standard A4 size sheet of paper. For printed copies user could either hold the paper by hand or have it fixed against the wall. For evaluating the different techniques, the lowest value of a combination of equal error rate (ERR), False Rejection Ratio (FRR), and False Acceptance Ratio (FAR) was determined.

There were three overall themes in the competition for anti-spoofing methods used: motion analysis, texture analysis, and liveness detection. Motion based analysis in this paper compares the difference between the movements of 3-D face to those of a 2-D printed version of the same face. Texture analysis observes the limitations of texture difference on printed sheets

and photographed blurriness in comparison to live captures. Lastly, liveness detection, as the name implies, tried to detect the liveness of an individual. This can be done three means facial movements such as blinking, breathing, etc. The following is a brief summary of each of the different algorithms used in the 6 methods explored [7]:

a. AMILAB

This method used a combination of all three techniques. Color differentiation and other techniques were used for texture differentiation. The result of the score was then captured alongside general movements of the face and number of blinks to determine whether the authentication attempt was legitimate or a spoof attempt.

b. CASIA

This method primarily observed facial recognition spoofing attempts through a video input. To determine a spoofed or legitimate attempt the team used three main assumptions about real videos in comparison to recordings: real faces have non-rigid movements, real videos have less noise, real videos have less background movement.

c. UNICAMP

This method explored primarily a motion-based technique to determine spoofed attempts. The two key factors in determining legitimacy was the movement of facial features, and the movement of background environments. This method considered printed copies as the spoof source, and thus could assume a printed facial movement would also be associated with the same background movement.

d. IDIAP

This method compares the usage of printed copies to properties observable when using such a method. Such as static print artifacts in a history of captured images. Thus, this solution captures a video of the attempts and uses stills from it in its algorithm to detect fake attempts.

e. SIANI

This method explores the usage of face area and its position in respects to the rest of the image to determine actual vs spoofed attempts. The key measurement areas made in the determination process involve: face, non-face, distance between eyes, and mouth.

f. UOULU

This method was very similar to the IDIAP method discussed, in that it used key print artifacts that can be identified when using printed copies of images.

Overall each of the 6 methods performed well in determining 2-D spoofing attempts against a facial recognition system. The most effective methods however were: CASIA, IDIAP and UOULU. Thus, the most effective observation to use when determining 2-D Spoofing attacks can then be associated to the identification of printed artifacts.

This paper presented a great high-level overview of the different methods currently being used to identify spoofing attempts in a facial recognition software. It also tested the different methods in a consistent environment to determine effectiveness of each of the methods proposed. The findings were encouraging in being able to determine 2-D spoofing attempts from legitimate ones.

B. Robust Multimodal Face and Fingerprint Fusion in The Presence of Spoofing Attacks

This paper, by Wold, Radu, Chen, and Ferryman, provides a new algorithmic technique for evaluating a multimodal face and fingerprint authentication system [8]. This paper was studied as it provided detailed context as to the different calculations and metrics considered when looking at biometric evaluation methods. In turn providing a better understanding of the feature matching component explained the Basics of facial recognition component.

The preliminary concern addressed in this paper is the usage of advanced spoofing techniques against biometrics systems even in the presence of multimodal recognition system. The two types of biometric recognition systems used in this study were for face and fingerprints. Previous studies showed the usage of these systems using a sum-rule based fusion technique in regards to

identifying spoofing events. The focus of this study involved using a newly proposed 1-mediam filtering alternative to the sum rule, and the usage of liveness scores and patterns of spoofed material.

The two methods of identifying spoofing attempts were describing in further detail. Liveness score is the usage of sensors in the biometric systems to help determine the "liveness" of the person or object being identified. As we saw from the previous paper, in facial recognition this could include recognition of blinking [7]. In fingerprint detection systems, sensors can be added to record pulse. For patterns of spoofed material, it could be as simple as lack of surrounding detail detection, or excessive noise. Nevertheless, of the usage of both or one of these methods, the usage of two biometric authentication methods makes spoofing a difficult task.

When introducing a multimodal biometric system, the algorithm used to combine the input must be considered. In this paper they suggest the use of a 1-median filtration method. The details of the method are out of scope for this article. However, the benefits include better tolerance of outliers when combining and face and fingerprint recognition scores, and a comprise between 0-spoof and mspoof robustness [8]. This method also allows for more flexibility in choosing the comprise level between accuracy and security for specific implementation methods based on the sensitivity of the protection reasoning [8].

The method was then tested against multiple experiments. Three different databases of spoofing attacks were used, and fed to the system to evaluate scores for the 1-mediam filtering approach. An approach to evaluate the system involved a combination of the liveness score and recognition score. The recognition score relating to how accurately the users are identified. The experiment found value in the combination of both scores in assessing the effectiveness [8]. Overall, the method proposed showed strong signs of effectiveness and benefits in comparison to the traditional sum-based approach.

C. C. Motion-Based Countermeasure Against Photo and Video Spoofing Attacks in Face Recognition

This paper studies previous motion-based countermeasures against spoofing in facial recognition, and proposes a new training set based approach [9]. The purpose of this paper was to also review approaches used to determine liveness features in more detail as this was a common trend amongst anti-spoofing techniques.

The first study reviewed in this paper, reflected on cue based motion detection to determine attacks. It depended on planner effect, face-background motion, and liveness. The 3 methods of detecting planner effect were further explained. Face part motion consistency is used to detect the consistency in different components of a face while in motion. This relation is not so consistent in actual face movements as compared to planner objects and thus can be used as a detection mechanism. Geometric invariants are simple to identify while measuring a 3D versus 2D (planner) object [9]. Structure from motion attempts to recreate the 3d objects from different images, and in turn can identify 2d spoofing attempts. For face-background motion we saw similar explanations in the previous articles, which simply attempt to correlate the movement of the face to the movement of the back ground [8]. We have also seen liveness detection using eye blinking, however this study also discussed use of lip movements and facial expressions.

Before explaining the method proposed in this paper an explanation of rigid and non-rigid facial movements can also help gain a better understanding of the context. As the paper describes, rigid and non-rigid movements of face can be used to help differentiate legitimate vs spoofed attacks. Rigid movements can be viewed as the movement of the entire face, such as head rotations. Non-rigid movements can be identified by movements within the face region, such as facial expressions. The focus on using the combination of the two is to help identify both 2D and 3D spoofing attempts. Non-rigid movements are of focus here as most 3D masks, unless of very high quality, are incapable of reproducing the movements of the facial features when, for example, an individual smile. This paper assumes the use of 2D attacks and generic 3D masks in its experiments and does not factor in high quality masks that are not readily available and very costly to produce [9].

The method and experiments conducted in this study are very detailed, and so an overview of the

approach will be provided here. The method proposed used rigid and no rigid motions as studied before, but places them through a machine learning based algorithm to detect differences. The purpose of this being the available key indicators of spoofing attempts explained earlier, and readily available datasets to train with. Using detection algorithms from previous studies and publicly available libraries, the system was developed to learn from 4 different datasets [9]. The system was then tested against different evaluation metrics to determine effectiveness of the proposed method. Overall it found the study of discriminant motion cues to be effective classifiers for the system to learn and predict spoofing attempts.

2. Conclusion

In this paper we reviewed attacks and defenses against facial recognition systems that were studied and tested. We began with studying the fundamentals of facial recognition works. This was followed by a literature review of multiple articles on spoofing 2D face recognition systems. We observed 5 different spoofing techniques, and on overview of the execution of one study. The next section conducted a literature review on the different types of defense systems that have been studied and tested. A comparison between 6 different mitigation techniques was reviewed, followed by an in-depth analysis and experimentation to new innovations for anti-spoofing face recognition systems. This literature review considered the potential of spoofing of only 2d face recognition systems with only computer aided interpretation on 3D rendering for spoofing detection.

Understanding the concepts explained in this article are critical for any business that is potentially using this type of authentication method, either directly or indirectly. The indirect justification can be understood through the increasing trend of BYOD work environments. If a company, for example, allows users to store and view emails on their personal mobile devices, they may be susceptible to attacks against facial recognition. Many phones, now support and promote the use of face detection for unlocking the device, such as iPhone X. In these situations, companies still have to consider the impact of having face recognition systems as point of attack for intruders to access confidential material. This concept is why fundamentally all security departments should train their employees on the basics of biometric systems. The material presented in this paper can be used to understand the basics of facial recognition, how it can be spoofed, and defenses against these attacks.

References

[1] S. Li and A. Jain, Handbook of Face Recognition. New York, NY: Springer Science+Business Media, Inc., 2005.

[2] Face Print. Vigilant Solutions.

[3] N. Erdogmus and S. Marcel, "Spoofing 2D face recognition systems with 3D masks," in International Conference of the BIOSIG Special Interest Group, 2013.

[4] R. Tronci, D. Muntoni, G. Fadda, M. Pili, N. Sirena and G. Murgia, "Fusion of multiple clues for photo-attack detection in face recognition systems," in International Joint Conference on Biometrics (IJCB), 2011.

[5] D. F. Smith, A. Wiliem and B. C. Lovell, "Face Recognition on Consumer Devices: Reflections on Replay Attacks," in IEEE Transactions on Information Forensics and Security, 2015.

[6] D. T. X. W. W. H. X. L. K. Z. Zhe Zhou, "Invisible Mask: Practical Attacks on Face Recognition with Infrared," 2018.

[7] M. Chakka, A. Anjos, S. Marcel, R. Tronci, D. Muntoni, G. Fadda, M. Pili, N. Sirena, G. Murgia, M. Ristori, F. Roli, Junjie Yan, Dong Yi, Zhen Lei, Zhiwei Zhang, S. Li, W. Schwartz, A. Rocha, H. Pedrini, J. Lorenzo-Navarro, M. Castrillon-Santana, J. Maatta, A. Hadid and M. Pietikainen, "Competition on counter measures to 2-D facial spoofing attacks", 2011 International Joint Conference on Biometrics (IJCB), 2011.

[8] P. Wild, P. Radu, L. Chen and J. Ferryman, "Robust multimodal face and fingerprint fusion in the presence of spoofing attacks", Pattern Recognition, vol. 50, pp. 17-25, 2016.

[9] T. Edmunds and A. Caplier, "Motion-based countermeasure against photo and video spoofing attacks in face recognition", Journal of Visual Communication and Image Representation, vol. 50, pp. 314-332, 2018.

[10] A. Anjos, L. El-Shafey, R. Wallace, M. Gunther, C. McCool, ¨ and S. Marcel. Bob: a free signal processing and machine learning toolbox for researchers. In ACM International Conference on Multimedia, pages 1449–1452, 2012.

[11] J. Bai, T.-T. Ng, X. Gao, and Y.-Q. Shi. Is physics-based liveness detection truly possible with a single image? In IEEE International Symposium on Circuits and Systems, pages 3425 –3428, June 2010.

[12] N. Kose and J.-L. Dugelay. Countermeasure for the protection of face recognition systems against mask attacks. In IEEE International Conference on Automatic Face and Gesture Recognition, April 2013.

[13] J. Maatta, A. Hadid, and M. Pietikainen. Face spoofing detection from single images using micro-texture analysis. In International Joint Conference on Biometrics, pages 1–7, 2011.

[14] K. Nixon, V. Aimale, and R. Rowe. Spoof detection schemes. In A. Jain, P. Flynn, and A. Ross, editors, Handbook of Biometrics, pages 403–423. Springer US, 2008.

[15] G. Pan, L. Sun, Z. Wu, and S. Lao. Eyeblink-based anti-spoofing in face recognition from a generic web camera. In IEEE International Conference on Computer Vision, pages 1 –8, October 2007.

[16] J. Trefny and J. Matas. Extended set of local binary patterns ` for rapid object detection. In Proceedings of the Computer Vision Winter Workshop, 2010.