Universe
Scientific Publishing

Review of Facial Recognition and Liveness Detect

Yuxuan Wang

Glasgow College, University of Electronic Science and Technology of China, Chengdu,611731, China

Email: yuxuanwang0822@163.com

Abstract:Facial recognition technology has been dramatically integrated into almost all the aspects of human life, such as mobile payment, identification applications, security management, and criminal cases, etc. However, these applications can be easily fooled by deliberate spoofing strategies. To ensure the identifications of users and avoid being spoofed are the central cores of this technology. As a result, its safeness and accuracy issues attract researchers to dig into this field. In terms of present existing deception and spoofing strategies, liveness detection plays a significant role in improving the robustness of facial recognition techniques. This paper will summarize the current mainstream facial recognition technology methods. The basic ideas, methods, implementations, and corresponding drawbacks of current facial recognition methods are in this paper. The future trends of facial recognition and liveness detection are also discussed and concluded.

Keywords: Facial Recognition, Liveness Detect, Deception, Spoofing Strategy.

1 INTRODUCTION

Due to the high convenience, durable stability, non-contact, and natural recognition characteristics, facial recognition technology is favored by tons of applications such as mobile payment, identification application, security management, etc. However, malicious deception and spoofing strategies are proposed deliberately to form an attack on the facial recognition and liveness detection system to gain personal information and possess it illegally.

Since the biological features of human faces are prone to be captured and accessed, the robustness and accuracy of this technology still require to be improved to certify users' identifications. The primary attack forms include (1) photo attack; (2) video attack; (3) 3-dimensional masks are made of bionic material to assure the verisimilitude of the spoofing face. When the attacker wearing the mask, the facial recognition system is not capable of distinguishing whether it is a natural face or not[1]; (4) attack aimed at neural network recognition method. Small perturbations [2-4] interfere with the performance of the recognition due to its vulnerability. Aiming at present deception methods, liveness detection has become a popular anti-deception system. It refers to ensure whether the captured face from the recognition system camera is a real human face or fake deception like printed photos, recorded videos, or 3D masks. In recent years, related research and liveness detection methods in the face recognition field have developed rapidly.

In this paper, we classify the existing anti-spoofing methods into interactive and non-interactive methods, which means whether the recognition system requires user cooperation. Then, describe each in the following sections. The remainder of this paper is organized in describing different categories as follows. We illustrate the interactive methods in Section 2 and non-interactive methods in Section 3. In Section 3, different methods in the subdirectory of non-interactive methods are summarized. Finally, the strengths and weaknesses of current facial anti-spoofing methods, as well as the analysis of their future, are concluded in Section 4.

2 INTERACTIVE METHODS

One way to confirm the liveness of the system user is through motion detection. It requires users to make corresponding actions within a specified time following the instructions displayed on the user interface, like turning the head, eye blinking, etc. In the interactive system of the preset action set, the subject can prove that it is a real face by making corresponding actions within the specified time according to the system requirements. This method prevents the 2D spoofing strategies successfully since static printed photos and pre-recorded videos cannot perform the randomly selected behaviors. Literature[5] calculates the area of the eye area and measures the teeth' HSV (Hue Saturation Value). The system then determines whether the user blinks, opens the mouth or not. Experiments have proved that attacker must use dynamic mouth and eyes to bypass the recognition which led to drastic changes in the facial structure. Although this method has a high recognition rate and can prevent photo and video deceptions, it requires a high degree of user cooperation without good user experience.

3 NON-INTERACTIVE METHODs

In reality, human faces are not absolutely static. It has various motions and micro-expressions. In terms of non-interactive methods, the liveness detection is executed without users' cooperation and consciousness. Many types of non-interactive methods are based on human face physical features[6], like texture, geometry, reflection rate, etc. These sub-classified strategies are illustrated in the following sections.

3.1 OPTICAL FLOW METHODS

The optical flow is the homeopathic movement of pixels. After relevant calculations, a collection of motion vectors (i.e., sports fields) of scenes or objects in the three-dimensional world coordinates in the image sequence can be obtained. Since the human face is a three-dimensional structure, its optical flow motion is quite different from that of the two-dimensional plane, so that the optical flow method can distinguish a fake human face from a real one. In literature[7], the Farneback Algorithm is first used to calculate the optical flow of the face area, and then the motion information, including the direction and angle, is converted into displacement data. Finally, SVM (support vector machine) is utilized to distinguish real faces and photos. Literature[8] modifies the ordinary approach by adding restrictions, and literature[9] uses Lucas Kanade Algorithm to reach a better performance. Although this approach is simple to be achieved, environmental factors like lighting are prone to be affected. Besides, the effect of anti-3D mask attacks is poor.

3.2 HYPERSPECTRAL, MULTISPECTRAL ANALYSIS

The hyperspectral and multispectral analysis has been proposed targeted at anti-spoofing the 3D mask attack. Since the materials used to make 3D mask figures special and have different physical characteristics with real human face skin, their imaging reflection rates vary in different wavelengths. In literature[10], the facial recognition system is implemented by combining the spectral, visual face, near infrared and thermal image to analyze the extracted human face features. Literature[11] used a fiber optic spectrometer to measure and compare the multi-spectral reflectance characteristic curves of human skin with common skin-like objects. The results show that the skin reflectance curve with a wavelength of 520nm- 600nm presents a "W" shape, which is obviously different from that of the photo. In literature[12], facial characteristics are collected by a hyperspectral imager. After analyzing the spectral characteristics under various lighting conditions and feature positions, the facial geometric features, the expression and posture changes, and the appropriate band is selected to perform anti-spoofing. The method based on multi-spectrum has higher accuracy and a more comprehensive range compared with the previously mentioned approaches, but it needs to be equipped with active light sources of different bands, such as infrared thermal imaging (spectral energy, and requires higher equipment. Besides, the stability under different ambient temperatures is not ideal, and it is costly to equip every facial recognition system with corresponding facilities.

3.3 CONVOLUTIONAL NEURAL NETWORKS

CNN (Convolutional neural network), as a widespread technique in the artificial intelligence field, appeals to tons of researchers to dig into its further applications. This technique is also used in the facial recognition region. Since the structure of the CNN is vulnerable to deliberate attacks, some defense ways like enhancing the network structure and feature extraction steps are proposed to improve its performance. In literature[13], the CNN is initially imposed for face anti-spoofing. In the next phase, improved methods based on CNN are then emerging. Literature[14] illustrates an approach that combined the DCT with CNN to deal with the extracted features of human faces. The face feature value is first fused with the CNN. After weighted fusion of global features and local features, the resulting image is input into CNN. Except for the conventional CNN methods, literature[15] also proposes a methodology that combined motion detection (eye blinking) with CNN for anti-spoofing. In literature[16], researchers added an extra lip motion detection into the system and combined it with CNN. Similar multi-method fusion strategies are also proposed: Another way for anti-spoofing is to improve the existing deep learning algorithms to raise the defense ability. Literature[17] compares the performance with CNN and an adjusted model named (LRF)-ELM (local receptive field). Literature[18] combines the method of nonlinear diffusion with CNN to enhance the anti-fraud performance. It is now proved by tons of researchers that CNN promotes facial anti-spoofing methods through various approaches with satisfying results. However, due to the excessive number of parameters, the neural network often has an over-fitting phenomenon, which causes the accuracy of the test to be reduced. Secondly, the use of CNN requires a large amount of data, which costs a lot in gathering, labeling, cropping, and other pretreatments. This leads to increasing research costs.

3.4 TEXTURE DETECTION

The texture feature contains the regular distribution of gray values formed by the repeated arrangement of objects on the image. The fake face image is generally made by more than one collection. Therefore, the local highlights, shadow changes, and blur degrees of the image will differ from the real one. In the texture detection method, the image of the human face is first converted to the frequency domain using the two-dimensional Fourier transform, and the high-frequency components of the two-dimensional photo are less than the real human face. In the literature[19], by analyzing the joint information of texture and color, studying brightness and chroma channels, it is found that the research method based on color texture is better than gray texture. This method is susceptible to noise, more sensitive to changes in illumination, and the accuracy of high-quality photo recognition is reduced.

3.5 MICRO MOTION DETECTION

Real human faces are not absolutely static. Micro-expression, including eyeball rotation and facial muscle contraction, can be used as a basis for live detection. It can be analyzed by zooming in on the micro-movements of the joint. Literature[20] proposed a combination of DMD, local binary pattern (LBP), and a classification pipeline with a histogram intersection core composed of support vector machines (SVM). DMD captures visual dynamics in fixed-size images, LBP can effectively capture dynamic patterns, and SVM is considered an ideal general classification tool. This mode can effectively extract dynamic information to obtain temporal dynamic characteristics and capture coherent spatial structure.

3.6 DEPTH INFORMATION

To enhance the system's robustness against attacks, more sophisticated anti-spoofing techniques can verify the three-dimensionality of the face captured by the device, such as by laser scanning. Literature[21] proposes a method based on three-dimensional photoelectric scanning. About 8000 feature points are obtained and curvature of features are calculated. However, the accuracy of this method decreases when folding the image, and the recognition effect of the 3D mask deception method is flawed. Its prominent advantage is that there is no need for excessive user interaction, and the user experience is greatly improved compared to the interactive recognition mode.

3.7 PUPIL RECOGNITION

Literature[22][23] developed a pupil direction observation system for anti-spoofing in the face recognition system. First, the Haar cascade classifier is combined with a specially trained eye region detection classifier to extract the eye region from the real-time camera. The Kanade-Lucas-Tomasi (KLT) algorithm is used to extract and track feature points to minimize head movement and obtain a stable eye area. This method is a low-cost security solution. The KLT algorithm is used to solve the problem of providing a stable eye frame for the algorithm while also reducing the calculation time and improving the algorithm's performance.

4 CONCLUSION

Although there are many methods of facial recognition and liveness detection in the current state, they still own advantages and disadvantages. However, these issues can be inevitably tackled in both practical and theoretical ways. In the process of analyzing existing methods, it is found that the appropriate and efficient technology should satisfy the following requirements: (1) less user interaction to improve users' experience; (2) the restricted cost and requirements for the practical equipment; (3) improved accuracy and speed to distinguish real and fake faces; (4) the better safety performance in the applications; (5) less influenced by the environmental factors, such as light, with better performance in multi-scene usage. Therefore, it can be concluded that the future of facial recognition has excellent potential, and research based on the above conditions will also become the mainstream direction of facial recognition technology.

References

[1] N. Erdogmus and S. Marcel, "Spoofing Face Recognition With 3D Masks," in IEEE Transactions on Information Forensics and Security, vol. 9, no. 7, pp. 1084-1097, July 2014.

[2] Z. Zhu, Y. Lu and C. Chiang, "Generating Adversarial Examples By Makeup Attacks on Face Recognition," 2019 IEEE International Conference on Image Processing (ICIP), 2019.

[3] Y. Zhong and W. Deng, "Towards Transferable Adversarial Attack Against Deep Face Recognition," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 1452-1466, 2021.

[4] X. Lin et al., "Exploratory Adversarial Attacks on Graph Neural Networks," 2020 IEEE International Conference on Data Mining (ICDM), 2020, pp. 1136-1141.

[5] A. K. Singh, P. Joshi and G. C. Nandi, "Face recognition with liveness detection using eye and mouth movement," 2014 International Conference on Signal Propagation and Computer Technology (ICSPCT 2014), Ajmer, 2014, pp. 592-597.

[6] E. Jiang, "A review of the comparative studies on traditional and intelligent face recognition methods," 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2020, pp. 11-15.

[7] Huang Jiankai. Research on living detection technology of face recognition [D]. Wuhan: Central China Normal University, 2018.

[8] W. Guojiang, Y. Guoliang and F. Kechang, "Facial Expression Recognition Based on Extended Optical Flow Constraint," 2010 International Conference on Intelligent Computation Technology and Automation, 2010, pp. 297-300.

[9] B. K. Dehkordi and J. Haddadnia, "Facial expression recognition in video sequence images by using optical flow," 2010 2nd International Conference on Signal Processing Systems, 2010, pp. V1-727-V1-730.

[10] B. K. Dehkordi and J. Haddadnia, "Facial expression recognition in video sequence images by using optical flow," 2010 2nd International Conference on Signal Processing Systems, 2010, pp. V1-727-V1-730.

[11] Hu Miaochun. Robust Multispectral Features for Face Liveness Detection [D]. ejing: Beijing Jiaotong University,2015.

[12] Liu Yifei. Face Liveness Detection Based on Spectrum Analysis and Depth Information [D]. Bejing: Beijing Jiaotong University,2017.

[13] Yang Jianwei, Lei Zhen, Li S Z. Learn Convolutional Neural Network for Face Anti-spoofing [EB/OL]. (2014-08-26).

[14] Wang Jiaxin and LeiZhichun. A Convolutional Neural Network Based on Feature Fusion for Face Recognition [J/OL]. Laser and Optoelectronics Progress. 2020, 57(10),339-345.

[15] M. M. Hasan, M. S. U. Yusuf, T. I. Rohan and S. Roy, "Efficient two stage approach to detect face liveness : Motion based and Deep learning based," 2019 4th International Conference on Electrical Information and Communication Technology (EICT), 2019, pp. 1-6.

[16] R. B. Hadiprakoso, H. Setiawan and Girinoto, "Face Anti-Spoofing Using CNN Classifier & Face liveness Detection," 2020 3rd International Conference on Information and Communications Technology (ICOIACT), 2020, pp. 143-147.

[17] Y. Akbulut, A. Şengür, Ü. Budak and S. Ekici, "Deep learning based face liveness detection in videos," 2017 International Artificial Intelligence and Data Processing Symposium (IDAP), 2017, pp. 1-4.

[18] R. Koshy and A. Mahmood, "Enhanced Anisotropic Diffusion-based CNN-LSTM Architecture for Video Face Liveness Detection," 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA), 2020, pp. 422-425.

[19] Boulkenafet Z, Komulainen J, Hadid A. Face spoofing detection using colour texture analysis [J]. IEEE Trans on Information Forensics and Security, 2016, 11 (8): 1818-1830.

[20] S. Tirunagari, N. Poh, D. Windridge, A. Iorliam, N. Suki and A. T. S. Ho, "Detection of Face Spoofing Using Visual Dynamics," in IEEE Transactions on Information Forensics and Security, vol. 10, no. 4, pp. 762-777, April 2015.

[21] A. Lagorio, M. Tistarelli, M. Cadoni, C. Fookes and S. Sridharan, "Liveness detection based on 3D face shape analysis," 2013 International Workshop on Biometrics and Forensics (IWBF), Lisbon, 2013, pp. 1-4.

[22] M. Killioğlu, M. Taşkiran and N. Kahraman, "Anti-spoofing in face recognition with liveness detection using pupil tracking," 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI), Herl'any, 2017, pp. 000087-000092.

[23] Deng Xiong, Wang Hongchun, Zhao Lijun, Wu Zhiyou, Pi Jiatian. Review of research methods for face recognition and live detection[J/OL]. Computer Application Research: 1-7[2019-12-12].