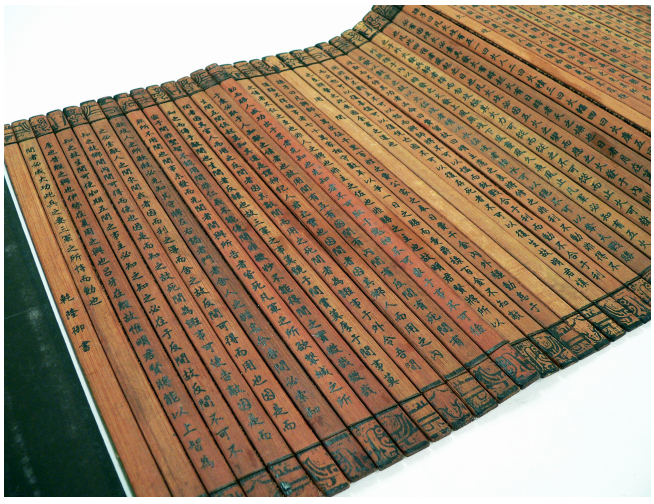
March 2021, Issue Code NL-2021-5-C

DDD (Digital Data Deception) Technology Watch Newsletter: Chinese Section

Table of Contents



Source: `https://www.flickr.com/photos/bluefootedbooby/370460130/`

“兵者，诡道也。故能而示之不能，用而示之不用，近而示之远，远而示之近。利而诱之，乱而取之，实而备之，强而避之，怒而挠之，卑而骄之，佚而劳之，亲而离之。攻其无备，出其不意。此兵家之胜，不可先传也。”

— 孙武：《孙子兵法·始计篇》

(The above is the original Chinese version of the English quotation shown on the cover page of the newsletter's main issue covering English papers.)

Editors: Li Qin, Shujun Li, Sanjay Bhattacherjee, Enes Altuncu, and Virginia Franqueira
Affiliation: Institute of Cyber Security for Society (iCSS), University of Kent, UK
Contact Us: ddd-newsletter@kent.ac.uk

Editorial

In this fifth issue of the Digital Data Deception (DDD) newsletter, we continue to include a Chinese section covering two selected topics related to DDD: biometrics spoofing (presentation attack) and anti-spoofing (liveness detection), and DDD-related psychology.

The papers covered in this Chinese section were identified via a mixed method: some were identified via a keyword-based search into Scopus, and others by manually inspecting the tables of contents of some selected journals. For the second method, the journals were selected based on the following criteria: if they provided fulltext access to papers, and if they are top-tier or highly technically relevant journals. In total 9 papers were selected, 6 on biometrics spoofing and liveness detection, and 3 on DDD-related psychology. One selected paper is a survey paper, and the other 8 papers are about original research.

For each paper covered in this section, we provide an objective summary of the research based on our own reading and understanding of the paper. We also provide our own editorial comments for each paper, including our recommendations and opinions on the research reported. For all papers covered, we paid attention to datasets and source code the authors used or developed, but none of the papers provided URLs of such resources. The same was observed for the Chinese section of the last (fourth) issue of the DDD newsletter. This may indicate that making source code and datasets used open source or public is less popular among Chinese researchers (at least when they publish in Chinese journals), but at this point we do not have sufficient evidence to arrive at a conclusion.

For this issue, we decided to include only papers we covered in the References. For research papers, datasets and source code cited in those papers, we directly embed their citation information (with a URL) to the summary. We hope that this new approach will help improve readability of the newsletter.

We hope you enjoy reading the Chinese section of this issue. Feedback is always welcome, and should be directed to ddd-newsletter@kent.ac.uk.

List of Acronyms

- ACER: Average Classification Error Rate
- aIAT: autobiographical Implicit Association Test
- APCER: Attack Presentation Classification Error Rate
- AUC: Area Under Curve
- BPCER: Bona Fide Presentation Classification Error Rate
- CNN: Convolutional Neural Network
- CNV: Contingent Negative Variation
- CVF: Computer Vision Foundation
- CVPR: Conference on Computer Vision and Pattern Recognition
- DFD (dataset): Deep Fake Detection
- DFDC: Deepfake Detection Challenge
- DQ_LBP: Difference Quantization Local Binary Pattern
- EEG: Electroencephalography
- EER: Equal Error Rate
- ERP: Event-Related Potentials
- Face-IoUP: Face-Intersection over Union with Penalty
- FCN: Fully Convolutional Network
- FFW (dataset): Fake Face in the Wild
- FPR: False Positive Rate
- HTER: Half Total Error Rate
- FGS: Fast Gradient Sign
- ITP: Implicit Theories of Personality
- LBP: Local Binary Pattern
- LPQ: Local Phrase Quantization
- RNN: Recurrent Neural Network
- ROC: Receiver Operating Characteristic
- rPPG: Remote Photoplethysmography
- SVM: Support Vector Machine
- TPR: True Positive Rate

Biometric Spoofing (Presentation Attack) and Anti-Spoofing (Liveness Detection)

Introduction

Biometric spoofing or presentation attack has been a major threat to biometrics-based applications such as user authentication and identification. The term "presentation attack" comes from the fact that such an attack is normally conducted by presenting a fake biometric signal in order to spoof the system. For instance, for face spoofing, an attacker can use a printed picture of the target user's facial image, a 3D printed mask, or even a dynamic video showing a 3D video, in front of the camera. A main technical solution to biometric spoofing is liveness detection, which tries to detect if a presented biometric signal is from a living person. The importance of biometric anti-spoofing has led to the creation of a new series of international standards (ISO/IEC 30107) and its first part ISO/IEC 30107-1 was published in 2016 (`https://www.iso.org/standard/53227.html`). The ISO/IEC standards use the term "presentation attack" for spoofing and "presentation attack detection" for anti-spoofing. In this section we use such terms interchangeably.

This section covers six recently published papers in Chinese journals on biometric spoofing and anti-spoofing. The six papers cover three sub-topics: adversarial sampling for face spoofing, face anti-spoofing, detecting deepfake faces and iris anti-spoofing. One of the six papers is a survey paper on face anti-spoofing.

Performance metrics of biometric anti-spoofing methods have been recently standardised in ISO/IEC 30107-3:2017 "Information technology – Biometric presentation attack detection – Part 3: Testing and reporting" (`https://www.iso.org/standard/67381.html`). The three most important metrics are the following:

- Attack Presentation Classification Error Rate (APCER): "proportion of attack presentations ... incorrectly classified as bona fide presentations".
- Bona fide Presentation Classification Error Rate (BPCER) or Normal Presentation Classification Error Rate (NPCER): "proportion of bona fide presentations incorrectly classified as attack presentations".
- Average Classification Error Rate (ACER): (APCER + BPCER)/2

In addition to research papers, there have been a number of challenges, competitions and benchmarks related to biometric anti-spoofing. For face anti-spoofing, the annual CVPR Face Anti-spoofing (Presentation Attack Detection) Challenge (since 2019, `https://sites.google.com/qq.com/face-anti-spoofing/`) deserves monitoring. It provides a platform for benchmarking newly proposed face anti-spoofing methods. CVPR (Conference on Computer Vision and Pattern Recognition) is one of most established research conferences on pattern recognition, co-sponsored by the IEEE Computer Society and the CVF (Computer Vision Foundation). For deepfake detection, the 2020 DFDC (Deepfake Detection Challenge, `https://www.kaggle.com/c/deepfake-detection-challenge/`) and the FaceForensics Benchmark (`http://kaldir.vc.in.tum.de/faceforensics_benchmark/` are of interest.

Adversarial Samples for Face Spoofing

Ma (马玉琨) et al. [6] proposed an adversarial sample generation method for face spoofing purposes, based on two intuitive principles: choosing the input dimensions that can cause less distortion in the output first, and adding perturbation constraints to avoid modifying too many pixels in a small region and to only allow the changing of one colour channel for each (RGB) pixel. The first principle implies that the method is a white-box method as it needs to know the structure of the underlying machine learning model to calculate the impact of each input dimension on the output signal. The purpose of the second principle is to reduce the perceptibility of the added noises to generate adversarial samples. In their main experiments, a CNN (Convolutional Neural Network) based face classifier was used as the underlying classifier. For the liveness detection dataset, they used the 2011 Print Attack dataset (`https://www.idiap.ch/dataset/printattack`), which includes 400 (200 real and 200 fake) videos. These videos were processed to get 72,806 training images and 48,451 testing images. The authors analysed the effect of the perturbation

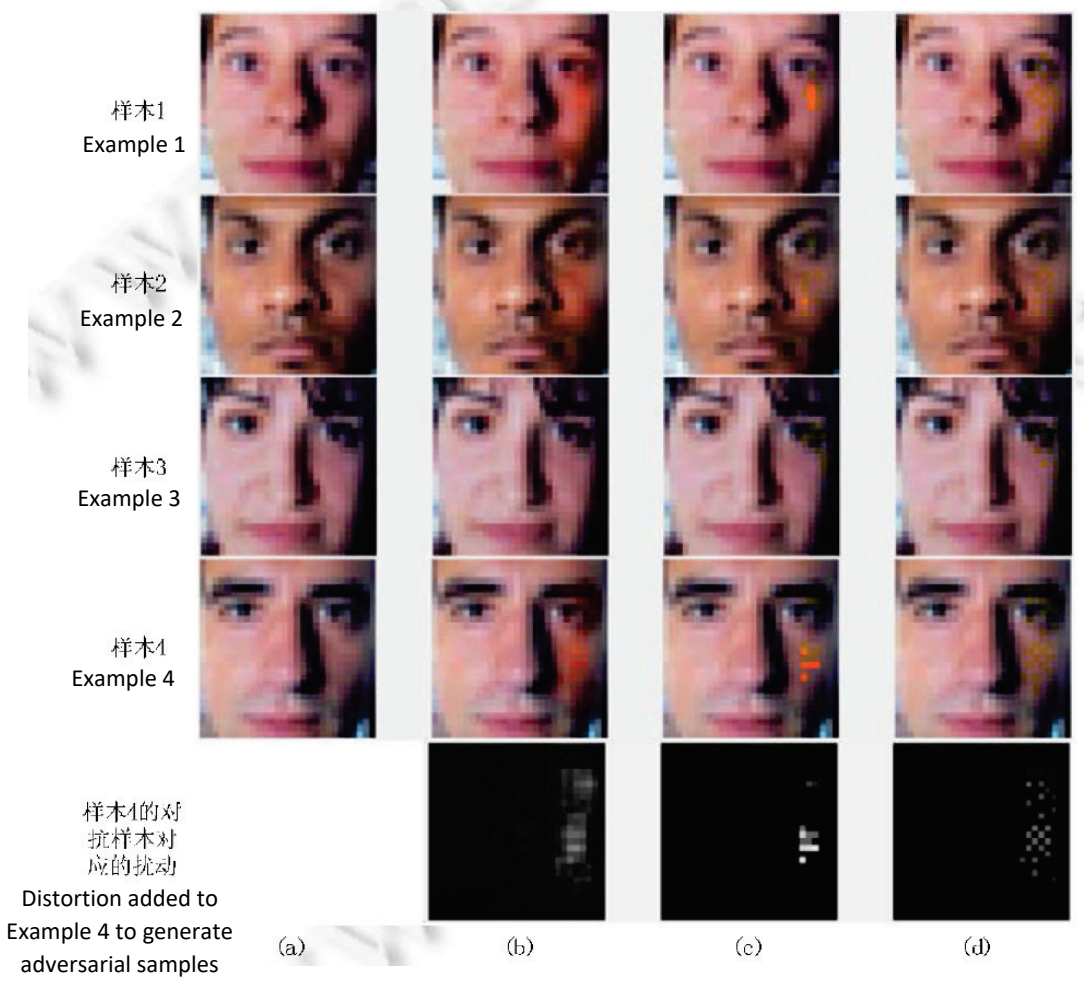


Figure 1: The comparison of Ma (马玉琨) et al.'s adversarial sample generation method with DeepFool (Figure 7 in [6]): (a) the original image; (b) the adversarial sample generated using DeepFool; (c) the adversarial sample generated by Ma (马玉琨) et al.'s method without perturbation constraints; (d) the adversarial sample generated by Ma (马玉琨) et al.'s method with perturbation constraints.

amplitude μ on the spoofing success rate against the CNN-based classifier and human observers, and concluded that 90 is the best value of μ to have the lowest perceptibility by 20 recruited human observers and a good spoofing success rate of over 90%. The authors compared the proposed method with two previously proposed methods: DeepFool (Moosavi-Dezfooli et al., 2016, `https://doi.org/10.1109/CVPR.2016.282`) and FGS (fast gradient sign) (Goodfellow et al., 2015, `https://arxiv.org/abs/1412.6572v3`). The experiments with recruited human observers showed that the proposed method can reduce the perceptibility by 20% and 21% comparing with DeepFool and FGS, respectively. Figure 1 shows four examples of adversarial samples for the proposed method and DeepFool. The authors also showed that the two principles they proposed could indeed help reduce the number of dimensions and pixels changed. They also tested the generalisability of their method by applying it to a different classifier LeNet-5 (`http://yann.lecun.com/exdb/lenet/`), showing a good spoofing success rate of 85.75% (slightly lower than DeepFool which achieved a rate of 87.63%).

Editorial Comments

This paper has a narrow focus on print attack and the proposed method is more pixel-based, so we do not anticipate it can be directly generalised to other types of presentation attacks. However, the two principles used are quite general so similar methods can be derived. The performance evaluation does not

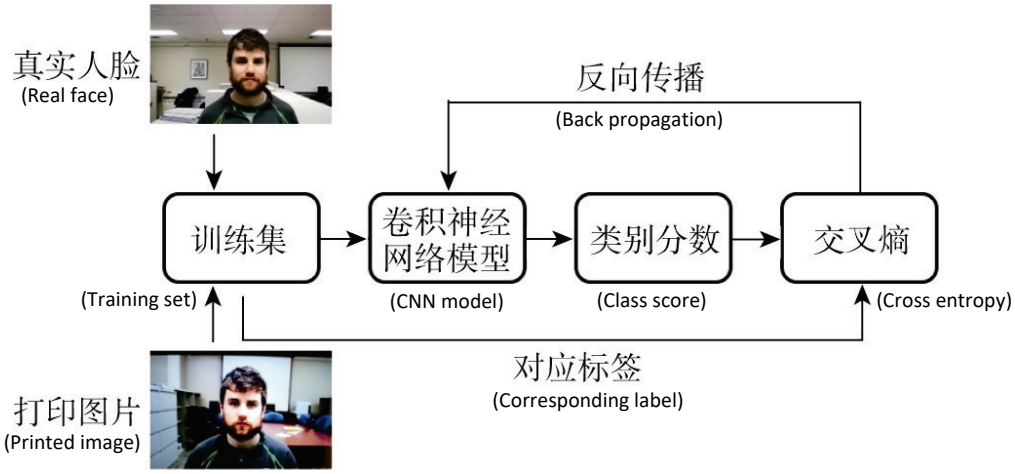


Figure 2: The general process of deep learning based face anti-spoofing using a CNN as its core (Figure 5 in Lu (卢子谦) et al. [5]).

look very comprehensive, but the use of human observers to evaluate the perceptibility of the adversarial samples is very good. Note that the authors did not give sufficient details of the experiments involving human participants, which will make reproducing their work harder.

Face Anti-Spoofing

Lu (卢子谦) et al. [5] reviewed recent research progress on anti-spoofing of facial images. The authors mentioned different methods of face anti-spoofing including those based on remote photoplethysmography (rPPG), texture analysis, optical flow, and those requiring additional input(s) such as speech, cooperative user action(s), infrared and depth signals. According to how machine learning is used, they reviewed face anti-spoofing research in the following two categories: those based on traditional machine learning methods such SVM (Support Vector Machine), and those based on deep learning models such as CNN, FCN (Fully Convolutional Network) and RNN (Recurrent Neural Network). They considered only application scenarios where the user does not need to do anything additional. Traditional methods are mostly based on manually defined features, but deep learning based methods normally leverage the deep learning architecture to automatically extract features. Unsurprisingly, more recent deep learning based methods have proven more effective and become the main stream of research in face anti-spoofing. Figure 2 shows the general process of deep learning based face anti-spoofing methods with a CNN as its core. Many deep learning based face anti-spoofing methods use information from multiple channels to improve their performance, e.g., multiple video frames, depth information, spatial information extracted from rPPG features, infrared and thermal images. The performance of modern deep learning based face anti-spoofing methods has reached a very high level, e.g., at a challenge organised at the CVPR 2019 conference, for 1,000 individuals and 21,000 face videos (with RGB, infrared and thermal imaging data), all top three teams achieved a very high TPR (true positive rate) under a very small FPR (false positive rate). The authors also pointed out that a new trend in this topic – more light-weighted models such as FeatherNets (ranked No. 3 in the CVPR 2019 competition) have proven to be capable of achieving a very high TPR with a very fast and lightweight model. In addition to reviewing related research, the authors also reviewed some commonly used datasets for testing face anti-spoofing methods.

Editorial Comments

While being generally informative, this review paper contains less structured information about different performance metrics used by researchers and does not refer to the highly relevant ISO/IEC 30107 standard series.
The actual performance figures of teams participating the CVPR 2019 challenge do not match the detailed results re-

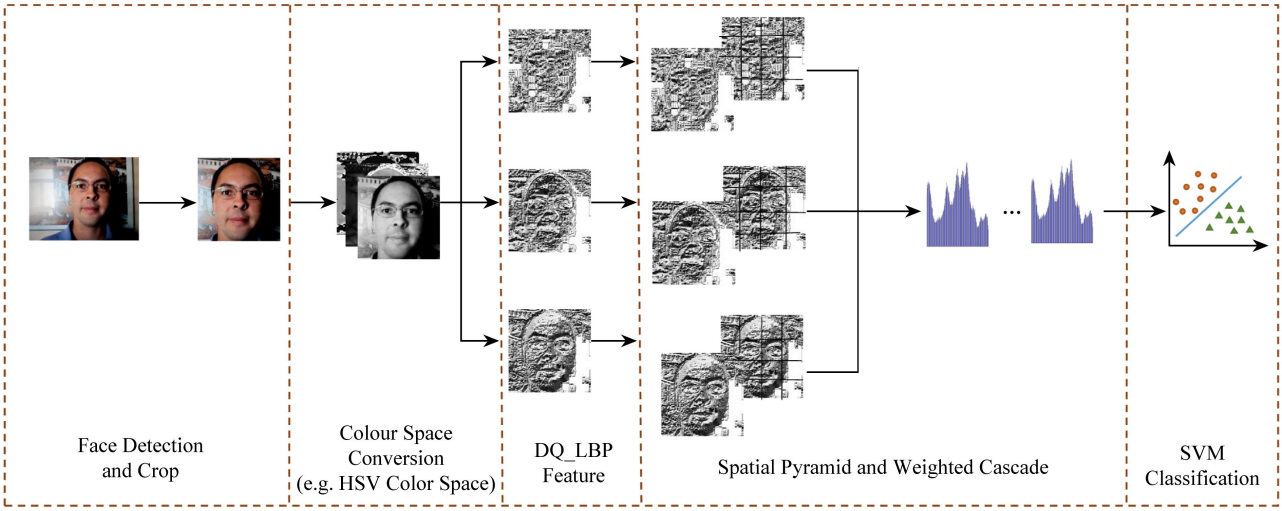


Figure 3: The DQ_LBP-based face anti-spoofing method proposed by Shu (束鑫) et al. (Figure 1 in [7]).

leased on the official competition website: `https://sites.google.com/qq.com/face-anti-spoofing/winners-results/challengecvpr2019`. The challenge was repeated in 2020 and the new results based on the more standardised metrics can be found at `https://sites.google.com/qq.com/face-anti-spoofing/winners-results/challengecvpr2020`. We contacted the organiser of the challenge and was told that they would continue to run a 2021 version as well.

Shu (束鑫) et al. [7] proposed a new face anti-spoofing method based on DQ_LBP (difference quantization local binary pattern) features extracted in a colour space and from different channels of a multi-resolution spatial pyramid (see Figure 3). Compared with the more traditional LBP features, DQ_LBP also encodes the differential values between adjacent pixels, so can provide more information about local texture patterns. By considering the actual distribution of differential pixel values, DQ_LBP is designed in such a way to avoid expanding the dimensionality of the feature set. For the classification method, the authors used the more traditional SVM. Four colour spaces were considered: gray scale, RGB, HSV and YCbCr. The performance was compared with six other methods for extracting similar features from a colour space. DQ_LBP without spatial pyramid was also used as a base-line option of the proposed method. For performance metrics, the authors mainly used EER (equal error rate) and HTER (half total error rate). The experimental results showed that DQ_LBP with spatial pyramid features achieved the best performance for two datasets CASIA FASD (`http://www.cbsr.ia.ac.cn/english/FaceAntiSpoofDatabases.asp`) and Replay-Attack (`https://www.idiap.ch/dataset/replayattack`), all colour spaces and both performance metrics (EER and HTER). The authors also studied how the number of layers of the spatial pyramid affects the performance, and the AUC (area under curve) of the ROC (receiver operating characteristic) curves showed that using two (the first and the last) of a three-layered pyramid can keep the performance with a significantly reduced dimensionality of the feature set. In their experiments the authors noticed that the best-performing colour space varied, so they also proposed to fuse features from different colour spaces, leading to the discovery that HSV+YCbCr is a good overall choice for three datasets, CASIA FASD, Replay-Attack and Replay-Mobile (`https://www.idiap.ch/dataset/replay-mobile`), and for both performance metrics (EER and HTER). In addition, the authors also tested the performance of the proposed method for a number of different types of presentation attacks in the three datasets used, demonstrating a very good performance overall. Finally, the authors compared the proposed method with 14 previously proposed ones (ten for CASIA FASD, Replay-Attack and five for Replay-Mobile, with one being used for both datasets) and showed that their method is among the best for most exper-

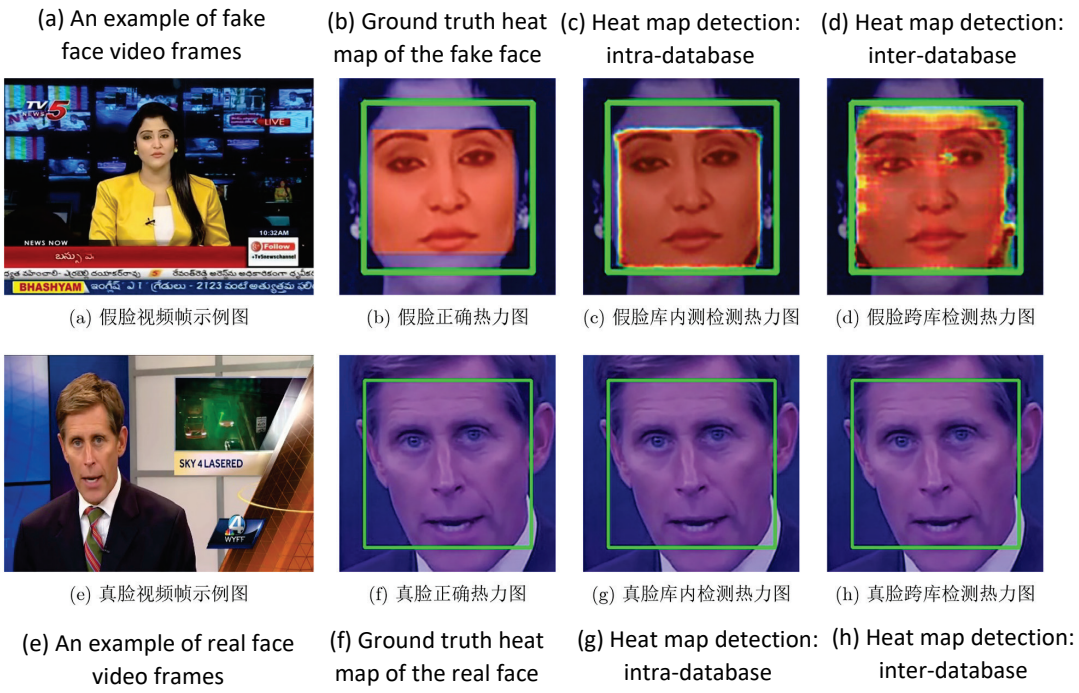


Figure 4: The detection results of the proposed fake face detection method proposed by Hu (胡永健) et al., applied to the database FaceForensics++ (Figure 2 in [2]).

imental conditions, particularly for Replay-Mobile. The authors mentioned that the proposed method did not generalise well across databases, and they planned to improve its performance by combining it with CNN-based approaches.

Editorial Comments

As a whole this paper provided very comprehensive experimental results to demonstrate its performance. As a traditional method not based on deep learning, the method looks interesting as it outperformed a number of deep learning based methods.

Detecting Deepfake Faces

Hu (胡永健) et al. [2] studied a related but different problem from biometric face spoofing: detecting deepfake faces in videos. The main difference is that deepfake faces are normally used to create fake videos, but not used to impersonate a target user in the context of biometric user authentication or identification. Deepfake face detection is often discussed in the context of digital or multimedia forensics, where the aim is to detect computer-generated photo-realistic faces. Despite the contextual difference, deepfake faces can be used for impersonation purposes and therefore deepfake detection methods can still find applications in face anti-spoofing. The main idea of the proposed method is to consider a deepfake face image as a spliced image using two sources, a computer-generated deepfake face and the background, with different textural structures at the pixel level. Based on this assumption, the authors proposed to use a deep neural network based image segmentation method to segment the input image into two parts: a suspected tampering mask M indicating a possible deepfake face, and the background. The tampering mask M is "denoised" to remove isolated small regions. Then, by comparing the tampering mask M with the detected face region, a tampering indicator called Face-IoUP (Face-Intersection over Union with Penalty) is proposed to calculate the intersection between the tampering mask M and the face region. A higher overlap more likely indicates a case of tampering (deepfake face) and a penalty is applied to discount overlaps falling outside of the detected face region. The denoising and the Face-IoUP step each uses a threshold, whose optimal value is determined by testing a range of possible values using a validation dataset.

To test the performance of the proposed method, Hu (胡永健) et al. used three base-line image segmentation models, FCN-8s, FCN-32s, and

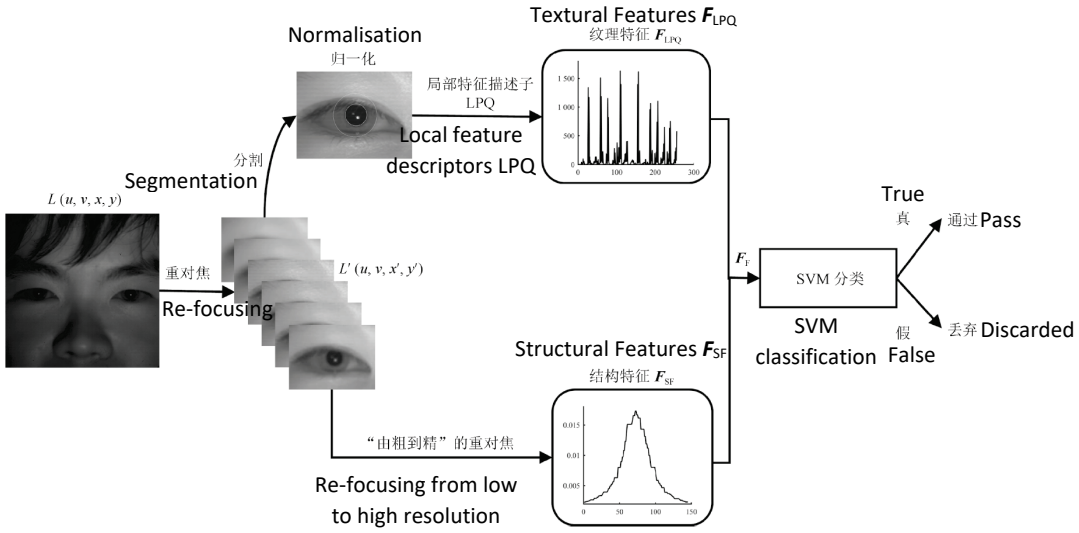


Figure 5: The architecture of light-field iris liveness detection method proposed by Song (宋平) et al. (Figure 1 in [8]).

DeepLabv3 (Chen et al., 2017, `https://github.com/tensorflow/models/tree/master/research/deeplab`). The experiments were conducted on four datasets, TIMIT (i.e., DeepfakeTIMIT, `https://www.idiap.ch/dataset/deepfaketimit`), FaceForensics++ (`https://github.com/ondyari/FaceForensics`), FFW (Fake Face in the Wild, `http://ali.khodabakhsh.org/research/ffw/`) and DFD (Deep Fake Detection, `https://github.com/ondyari/FaceForensics/tree/master/dataset/DeepFakeDetection`), each split into the training, validation and testing sets following a 7:2:1 ratio. Because FFW does not include real videos, 50 videos were randomly selected from FaceForensics++. Both intra- and inter-dataset detection experiments were conducted to validate the generalisability of the proposed method. The proposed method was compared with four recently proposed methods, MesoInception-4 (Afchar et al., 2018, `https://github.com/DariusAf/MesoNet`), ShallowNetV1 (Tariq et al., 2018, `https://doi.org/10.1145/3267357.3267367`), MISLnet (Bayar & Stamm, 2018, `https://gitlab.com/MISLgit/constrained-conv-TIFS2018`), ResNet-50 (He et al., 2016, `https://doi.org/10.1109/CVPR.2016.90`) and Xception (Chollet, 2017, `https://keras.io/api/applications/xception/`). The experimental results showed that the proposed method worked very well under all conditions, particularly on reducing the inter-dataset detection error rate. The results are largely stable across all three datasets used, demonstrating that the proposed method is very robust. The time complexity of the proposed method is higher than some other methods, but not significantly so (e.g., around 50% slower than MesoInception-4). This is considered less of an issue for digital forensics, but could lead to usability issues for user authentication.

Editorial Comments

The paper reported that for a 10s video the average time for detection was 37.8s. This time complexity is still acceptable if generalised to detect a single facial image. However, if it is used to detect a short video, the method will need some significant optimisation to improve its efficiency.

Iris Liveness Detection for Anti-Spoofing

Song (宋平) et al. [8] proposed a light-field imaging based iris liveness detection method in more complex imaging environments, including a longer distance, greater depth of field, half-controlled, more noisy and complicated background and lighting conditions, etc. Figure 5 shows the general architecture of the proposed method, which include three large steps. Firstly, the original image is processed to obtain a four-dimensional light-field data, and then a re-focusing step is applied to obtain a number of focal stack images. Next, two sets of supplementary features are calculated separately: one set reflecting

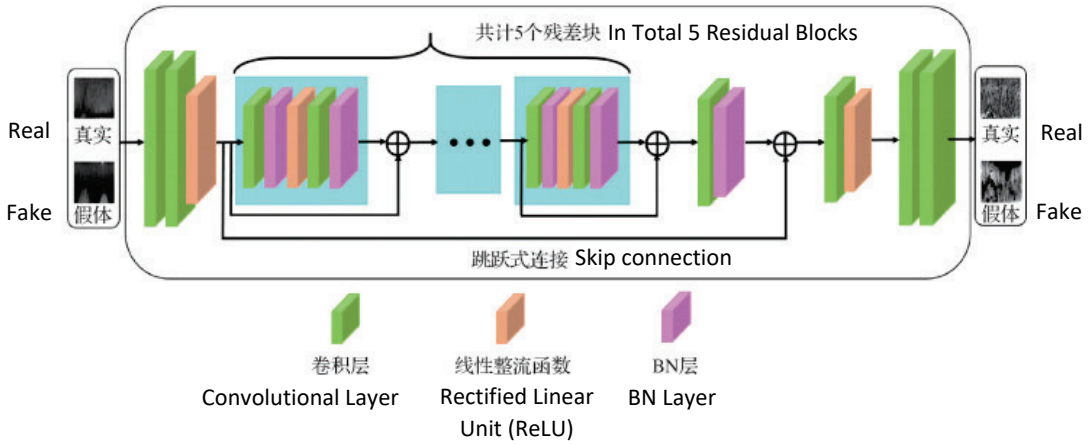


Figure 6: The architecture of the enhanced gray-level space generator for the iris anti-spoofing method proposed by Liu (刘明康) et al. (Figure 1 in [4]).

the 3D structures of the focal stack images, and the other set reflecting the LPQ (local phrase quantization) texture of the sharpest focal stack image via a number of steps including iris localisation, segmentation, normalisation and feature extraction. Finally, the two sets of features are fused and then an SVM is used to classify iris images. To test the performance of the proposed method, the authors built a medium-sized iris liveness detection dataset with 504 valid iris samples, including 230 real iris images collected from 14 human participants and 274 fake iris images of two types (printed iris images on plain and photo paper, and iris images shown on an iPad screen). The dataset was split into the training set with real and fake samples from five participants, and the testing set with samples from the other nine participants. The performance of the proposed method was compared with ten other iris liveness detection methods, and the proposed method achieved the best performance (ACER = 3.69%).

Editorial Comments

This paper's main highlight is the handling of the complex imaging environment. Most other studies on biometric liveness detection were based on "easier" (and less realistic) scenarios. The authors used their own light-field imaging cameras and also produced their own dataset for their experiments. We however did not see a URL of the dataset they produced.
The authors claimed that the iris liveness dataset they constructed is the first long-distance, near-infrared light-field dataset in the research literature. They planned to further enrich the dataset to increase the number of samples. The size of their dataset used in [8] is actually quite small (just 14 participants), so the expansion of the dataset is indeed useful.

Liu (刘明康) et al. [4] proposed an enhanced gray-scale space in which real and fake iris images are separated further to allow easier classification between the two classes. The new space is defined using a ResNet as shown in Figure 6. Based on this new gray-scale image space, the authors proposed to use a pre-trained LightCNN-4 network for feature extraction. After that, a triplet loss function is defined based on three classes of inputs: true samples, fake samples, and the so-called central anchor (true) sample that is calculated as the true sample closest to the average sample in the feature space. The aim of the triplet loss function is threefold: 1) to minimise the distance between true samples and the anchor sample; 2) to maximise the distance between false samples and the anchor sample; 3) to keep a safe margin between the two classes. The triplet loss function is combined with the softmax loss function to balance speed and performance. To test the performance of the proposed method, the authors conducted experiments on three datasets: ND-Contact (Doyle et al., 2013, `https://cvrl.nd.edu/projects/data/#nd-cosmetic-contact-lenses-2013-data-set`), CRIPAC-Printed-Iris and CASIA-Iris-Fake (Sun et al., 2014, `https://doi.org/10.1109/TPAMI.2013.234`). Four previously proposed methods were used

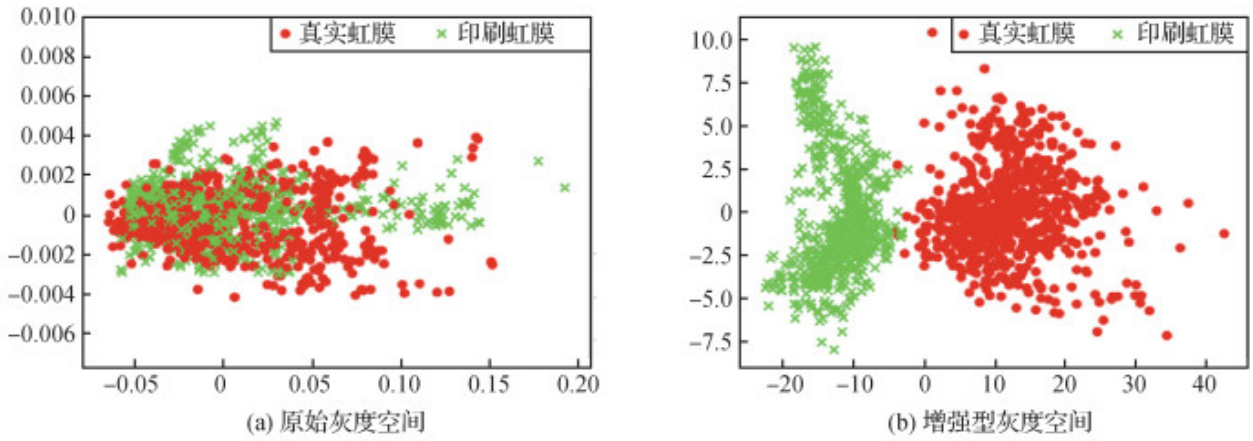


Figure 7: The feature space distributions of real and false samples in the CRIPAC-Printed-Iris dataset (Figure 13 in [4]): (a) the original gray-scale space; (b) the enhanced gray-scale space. Red and green dots represent true and false samples, respectively.

as the benchmarks: HMC (Yan et al., 2018, `https://doi.org/10.1109/ICB2018.2018.00018`), Iris + Ocular (Hoffman et al., 2018, `https://doi.org/10.1109/CVPRW.2018.00213`), MCNN (He et al., 2016, `https://doi.org/10.1109/BTAS.2016.7791186`), and MT-PAD (Chen & Ross, 2018, `https://doi.org/10.1109/WACVW.2018.00011`). For all experiments, 2,500 images (1,500 true samples and 1,000 false samples) were used for the training set and 1,000 images (600 true samples and 400 false samples) for the testing set. The authors tested the performance under both intra- and inter-dataset conditions, and also under the inter-attack condition, i.e., when false samples in the training set are of a different type from false samples in the testing set. The inter-dataset and inter-attack conditions were used to test the generalisability of the proposed method. For all experiments, the proposed method achieved the best performance among all five methods, particularly under the inter-dataset and inter-attack conditions – the FRR (false rejection rate) was reduced from over 10% for the other four methods to below 0.2%, while reducing the FAR (false accept rate) to below 1.7% from between 2.53% to 13.37% for the other four methods. In addition to showing the experimental results on performance, the authors also used visualisations to show how the proposed enhanced gray-scale space can help separate true and false iris images in the new feature space (see Figure 7 for one such visualisation).

Editorial Comments

The experimental results in this paper are very comprehensive and provide strong evidence that the proposed method is very promising for iris liveness detection under many conditions. The visualisations in the paper are particularly interesting as they provide more explainable evidence on why the proposed method performed so well. The core of the proposed method is the enhanced gray-scaled space, which seems quite general and can potentially be generalised to other biometric modalities such as face anti-spoofing. This paper was published in July 2020 and we recommend following up the authors' future work on related topics.

NB: The paper's PDF version could not be downloaded without creating an account but when we tried to register an account the web page was broken. The HTML fulltext was however available at `http://www.cjig.cn/html/jig/2020/7/20200711.htm`.

DDD-related Psychology

Introduction

Since deception is mostly created by a human creator for the purpose of affecting behaviours of other human receiver(s), studying the human psychology behind DDD can provide useful insights on detecting and preventing DDD.

This section covers two papers studying psychological aspects of deception and a third paper proposing a conceptual framework of deception detection based on human tasks with cognitive load manipulation. One paper [9] is of particular interest: it looked at how Chinese-English bilinguals behaved when lying and telling the truth using EEG (electroencephalography) signals and provided useful insights about the role of foreign languages in lying behaviours.

Psychology of Deceivers

Zhang (张积家) et al. [9] conducted some EEG-based experiments to explore the differences in the neural mechanism of Chinese-English bilinguals when lying in Chinese and English. Their work mainly focused on the effects of two psychological factors: cognitive load and emotion. They recruited 34 Chinese participants who speak Chinese (Mandarin) as the first language and English as the second language, with a balanced gender ratio. Each participant's English proficiency was tested following the China Standards of English (CSE, `http://cse.neea.edu.cn/`). For the main part of the experiment, each human participant was asked to do four tasks (each for 60 times) defined by a 2×2 experimental design: language (2 values – Chinese and English) and honesty (2 values – telling the truth and lying). For each task, the exact procedure is illustrated in Figure 8. During the tasks, participants were sitting in a soundproof electromagnetic shielding room, facing a computer screen, and their EEG signals were collected using a Neuroscan ERP (event-related potential) workstation. After conducting all the four tasks, participants were asked to take a questionnaire covering three aspects: 1) perceived emotional strength using a simpler questionnaire derived from Caldwell-Harris & Dinn's work in 2009 (`https://doi.org/10.1016/j.ijpsycho.2008.09.006`); 2) the preferred language for telling the truth and lying and the perceived reason; 3) rapid naming of pictured objects, colours, letters and numbers following Denckla & Rudel's work in 1974 (`https://doi.org/10.1016/S0010-9452(74)80009-2`). The analysis of the EEG signals was focused on two ERP components: P200 (an indicator of early anxiety) and CNV (contingent negative variation). As a whole, the results of the EEG signals and the self-reported data led to the

Figure 8: The five-stepped procedure of the task "lying in Chinese" in the experiment conducted by Zhang (张积家) et al. (Figure 1 in [9]), showing what a human participant saw during the task on a computer screen. The human participant told the truth or a lie in Step 4 after seeing a word shown below the picture in Step 3. For other three tasks, the word shown in Step 3 will differ: "真话" ("truth" in Chinese) for telling the truth in Chinese, "truth" for telling the truth in English and "lie" for lying in English. The picture shown in Steps 2 and 3 was selected from 15 possible ones defined by three different colours (red, green and blue) and five different animals (cat, dog, fish, bird, and goat) corresponding to high-frequency words in both Chinese and English. The pictures were shown randomly to participants and the random order was reshuffled after every two participants. The final step was used to separate two consecutive tasks and it was skipped if a human participant did not say anything in Step 4.

following main conclusions: 1) lying led to a heavier cognitive load than telling the truth; 2) speaking a foreign language (English) led to more anxiety, 3) lying in a foreign language (English) led to a heavier cognitive load than in the native language (Chinese); 4) when speaking in a foreign language (English) the anxiety took so much cognitive load that the tension is more dominated by the foreign language component than by lying. The results in this paper re-confirmed an observation reported in some previous studies: when speaking a foreign language there can be more signs of lying in a person's speech.

Editorial Comments

This paper may look less related to DDD, but the results can provide useful insights regarding how to detect deceptive data created by humans, especially when they are created in real time and in a context where a foreign language is involved. Particularly, since more signs of lying could be observed when a person speaks in a foreign language, the effect of the foreign language should be considered when designing deception detection methods to avoid false positives and biases.

Ding (丁倩) et al. [1] studied the effect of psychological entitlement on online cheating behaviours and its inner mechanism among university students. Two other factors were also studied in the context of psychological entitlement: ego depletion and ITP (implicit theories of personality). The term "ego depletion" refers to a reduced level of self-control after depletion of resources managing self-control. The term ITP refers to two different types of personality according to how a person explains human behaviours: entity theorists who tend to explain human behaviours using intrinsic and static characteristics, and incremental theorists who tend to explain human behaviours using a dynamic and more external-affected manner. The study involved 800 undergraduate students recruited from two Chinese universities, who were asked to take a survey. After removing invalid and incomplete responses, in total 738 responses were considered valid, including responses from 300 males and 438 females. The survey questionnaire includes four parts covering psychological entitlement, ego depletion, ITP and online cheating behaviours, respectively. The results showed that: 1) psychological entitlement can positively predict online cheating behaviour (the gender as the control variable); 2) ego depletion can partially mediate the interaction between psychological entitlement and online cheating behaviour; 3) ITP plays a role on the relationship between psychological entitlement and ego depletion – the relationship was stronger for entity theorists than for incremental theorists. Based on the findings, the authors explained their practical implications: education and other social interventions can be used to reduce the level of psychological entitlement of people, help people to recover resources for self-control, and nudge entity theorists towards incremental theorists, in order to help prevent people from engaging with online cheating behaviours.

Editorial Comments

Although the paper focussed on university students, the main results obtained are largely aligned with previous studies on other sub-populations of people. Considering that university students are more active online users and often more technical savvy with new ICT technologies, we believe that studies on this particular sub-population are important.

Psychology in Deception Detection

Liang (梁静) et al. [3] proposed a conceptual framework for deception detection based on cognitive load manipulation. The framework is based on the observation that people tend to have a higher cognitive load when lying due to the memory-response conflict – they have to try to conceal critical information for which they need to lie while behaving normally for other irrelevant information. The higher cognitive load comes from the need to process two different types of information differently. Since truth tellers do not have to manage the memory-response conflict, some researchers hypothesised that by employing some human tasks with cognitive load manipulation we may be able to enlarge the observable gap between their behaviours and truth tellers'. The authors stated that this approach is relatively less studied and results from past studies are not all consistent. A visual illustration of the proposed framework is shown in Figure 9. The authors have not conducted actual experiments, but described a number of concrete experimental designs following their proposed framework. All experiments employ the Con-

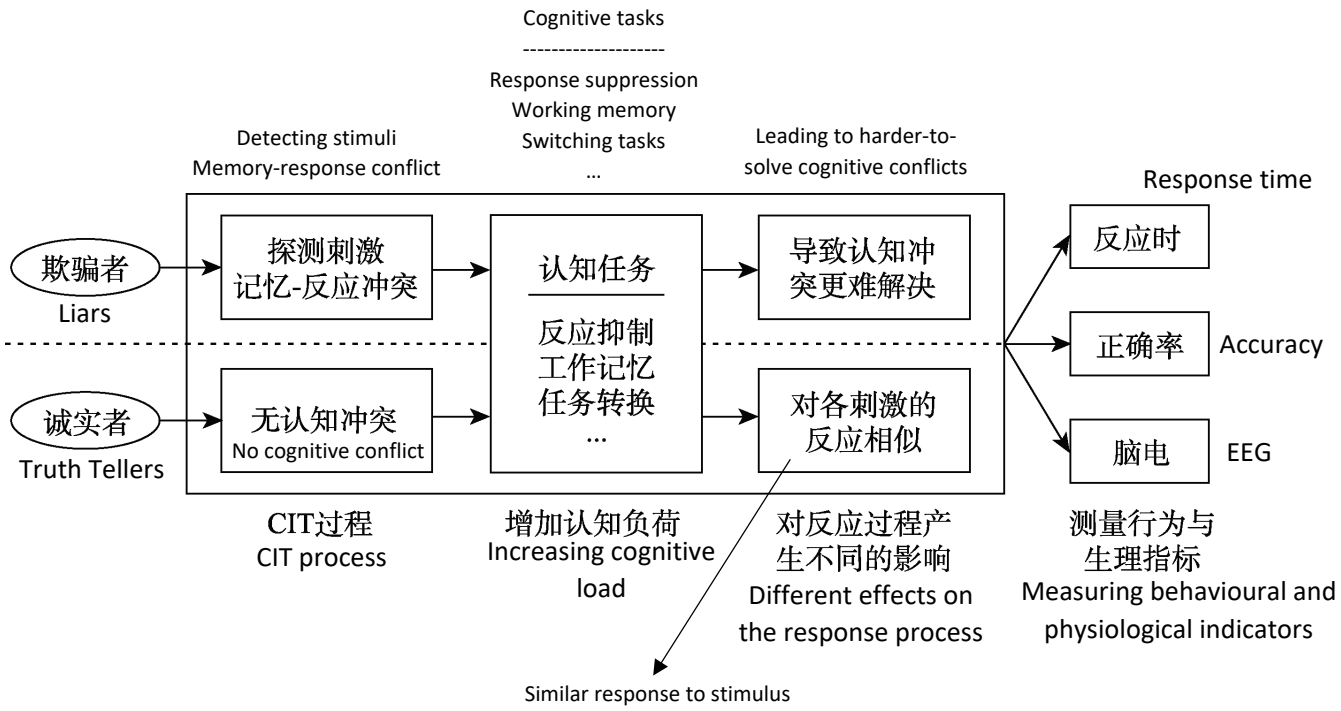


Figure 9: The deception detection process based on cognitive load manipulation proposed by Liang (梁静) et al. (Figure 2 in [3]).

cealed Information Test (CIT), in which human participants are asked to simulate a crime scene and then a number of stimuli (cognitive tasks) are used to test how they (simulated liars and truth tellers) respond to them. Three experiments are designed to study how the nature and difficulty level of the cognitive tasks can influence people's responses. The other two experiments employ both the CIT and the aIAT (autobiographical Implicit Association Test) to measure relevant response indicators (response time and ERP indicators) in the proposed deception detection framework. For the latter two experiments, the authors planned to use university students in one experiment and criminal suspects in the other. The experiment using criminal suspects will help increase the ecological validity of the results obtained since the human participants will be closer to those in real world scenarios. In addition to the psychological experiments, the authors also planned to study how to use machine learning to predict individual deceptive behaviours, taking the response indicators in the second class of experiments as the input.

Editorial Comments

Different from other papers we have covered in the DDD newsletter (this and past issues), this paper reports only research ideas and plans, not actual research results. The authors used the wording "project" throughout the paper without explicitly explaining what project they were referring to. According to the acknowledgements on the first page of the paper, the research was supported by six research projects (two funded by the National Natural Science Foundation of China and other four by provincial funders). We assume that the research is part of one or more of those projects.

Despite having no actual research results, the proposed deception detection framework looks technically sound and the experiments designed seem quite detailed. We recommend monitoring the authors' work in future. Since this paper was published in October 2020, we do not expect that any significant new results will be published before late 2021.

References

[1] Qian Ding (丁倩), Yilin Liu (刘祎琳), Yongxin Zhang (张永欣), and Zongkui Zhou (周宗奎). 2020. Psychological Entitlement and Online Cheating Behavior in College Students: A Moderated Mediation Model / 心理特权与大学生网络欺骗行为：有调节的中介效应. *Studies of Psychology and Behavior* / 《心理与行为研究》 18, 4 (2020), 537–543. http://psybeh.tjnu.edu.cn/CN/Y2020/V18/I4/537

[2] Yongjian Hu (胡永健), Yifei Gao (高逸飞), Beibei Liu (刘琲贝), and Guanjun Liao (廖广军). 2021. Deepfake Videos Detection Based on Image Segmentation with Deep Neural Networks / 基于图像分割网络的深度假脸视频篡改检测. *Journal of Electronics and Information Technology* /《电子与信息学报》 43, 1 (2021), 162–170. https://doi.org/10.11999/JEIT200077

[3] Jing Liang (梁静), Qiannan Ruan (阮倩男), He Li (李贺), Mengqing Ma (马梦晴), and Wenjing Yan (颜文靖). 2020. Deception detection based on memory-response conflict: A cognitive load approach / 认知负荷取向下基于记忆–反应冲突的欺骗检测. *Advances in Psychological Sciences* /《心理科学进展》 28, 10 (2020), 1619–1630. https://doi.org/10.3724/SP.J.1042.2020.01619

[4] Mingkang Liu (刘明康), Hongmin Wang (王宏民), Qi Li (李琦), and Sun (孙哲南) Zhenan. 2020. Enhanced gray-level image space for iris liveness detection / 增强型灰度图像空间实现虹膜活体检测. *Journal of Image and Graphics* /《中国图像图形学报》 25, 7 (2020), 1421–1435. http://www.cjig.cn/html/jig/2020/7/20200711.htm

[5] Ziqian Lu (卢子谦), Zheming Lu (陆哲明), Fengli Shen (沈冯立), and Zonghui Wang (王总辉). 2020. A Survey of Face Anti-Spoofing / 人脸反欺诈活体检测综述. *Journal of Cyber Security* /《信息安全学报》 5, 2 (2020), 18–27. http://jcs.iie.ac.cn/xxaqxb/ch/reader/view_abstract.aspx?file_no=20200203&flag=1

[6] Yu-Kun Ma (马玉琨), Li-Fang Wu (毋立芳), Meng Jian (简萌), Fang-Hao Liu (刘方昊), and Zhou Yang (杨洲). 2019. Algorithm to Generate Adversarial Examples for Face-spoofing Detection / 一种面向人脸活体检测的对抗样本生成算法. *Journal of Software* /《软件学报》 30, 2 (2019), 469–480. http://www.jos.org.cn/html/2019/2/5568.htm

[7] Xin Shu (束鑫), Hui Tang (唐慧), Xibei Yang (杨习贝), Xiaoning Song (宋晓宁), and Xiaojun Wu (吴小俊). 2020. Research on Face Anti-Spoofing Algorithm Based on DQ_LBP / 基于差分量化局部二值模式的人脸反欺诈算法研究. *Journal of Computer Research and Development* /《计算机研究与发展》 57, 7 (2020), 1508–1521. https://doi.org/10.7544/issn1000-1239.2020.20190319

[8] Ping Song (宋平), Ling Huang (黄玲), Yun-Long Wang (王云龙), Fei Liu (刘菲), and Zhe-Nan Sun (孙哲南). 2019. Iris Liveness Detection Based on Light Field Imaging / 基于计算光场成像的虹膜活体检测方法. *Acta Automatica Sinica* /《自动化学报》 45, 9 (2019), 1701–1712. http://www.aas.net.cn/cn/article/doi/10.16383/j.aas.c180213

[9] Jijia Zhang (张积家), Yutong Lu (陆禹同), Qirui Zhang (张启睿), and Jinqiao Zhang (张金桥). 2020. The effects of foreign language anxiety, nervousness and cognitive load on foreign language lying: Evidence from Chinese-English bilinguals / 外语焦虑、紧张情绪与认知负荷对外语说谎的影响：来自中–英双语者的证据. *Acta Psychologica Sinica* /《心理学报》 52, 7 (2020), 861–873. https://doi.org/10.3724/SP.J.1041.2020.00861