JOURNAL OF IMAGE AND GRAPHICS

主办：中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2021
06
VOL.26

ISSN1006-8961
CN11-3758/TB

中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）

第26卷第6期（总第302期）
2021年6月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

主管单位 中国科学院
主办单位 中国科学院空天信息创新研究院
中国图象图形学学会
北京应用物理与计算数学研究所

主　　编 吴一戎
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮　　编 100190
电子信箱 jig@aircas.ac.cn
电　　话 010-58887035
网　　址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订　　购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
（邮政信箱：北京399信箱　邮编：100048）
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Aerospace Information Research Institute, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@aircas.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048,P.R.China))
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元

2021年6月 第26卷 第6期（总第302期）
中国图象图形学报
Zhongguo Tuxiang Tuxing Xuebao

目次

生物特征识别学科发展报告(第1254页)

图像处理与通信技术

大规模室外图像3维重建技术研究进展(第1429页)

三维视觉和图形技术

虚实融合场景中的深度感知研究综述(第1503页)

Volume 26, Number 6
Published June 16, 2021

CONTENTS
JOURNAL OF IMAGE AND GRAPHICS

Overview of biometrics research (P1254)

Progress in the large-scale outdoor image 3D reconstruction (P1429)

Review of depth perception in virtual and real fusion environment (P1503)

E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035

中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS

中图法分类号:TP389.1　　**文献标识码**:A　　**文章编号**:1006-8961(2021)06-1216-11

论文引用格式:Li X L, Yu N H, Zhang X P, Zhang W M, Li B, Lu W, Wang W and Liu X L. 2021. Overview of digital media forensics technology. Journal of Image and Graphics,26(06):1216-1226(李晓龙,俞能海,张新鹏,张卫明,李斌,卢伟,王伟,刘晓龙. 2021. 数字媒体取证技术综述. 中国图象图形学报,26(06):1216-1226)[DOI:10.11834/jig.210081]

数字媒体取证技术综述

李晓龙[1],俞能海[2],张新鹏[3],张卫明[2],李斌[4],卢伟[5],王伟[6],刘晓龙[1]

1. 北京交通大学信息科学研究所, 北京　100044; 2. 中国科学技术大学网络空间安全学院, 合肥　230027;
3. 上海大学通信与信息工程学院, 上海　200444; 4. 深圳大学电子与信息工程学院, 深圳　518060;
5. 中山大学计算机学院, 广州　510006; 6. 中国科学院自动化研究所, 北京　100190

摘　要:　面对每天有数以百万计通过网络传播的多媒体数据,到底哪些内容是真实可信的,虚假内容的背后又经历了哪些篡改?数字取证技术将给出答案。该技术不预先嵌入水印,而是直接分析多媒体数据的内容,达到辨别真实性的目的。任何篡改和伪造都会在一定程度上破坏原始多媒体数据本身固有特征的完整性,由于其具有一致性和独特性,可作为自身的"固有指纹",用于鉴别篡改文件。随着篡改媒体的数量与日俱增,社会稳定甚至国家安全受到了严重威胁。特别地,随着深度学习技术的快速发展,虚假媒体与真实媒体之间的感官差距越来越小,这对媒体取证研究提出了巨大挑战,并使得多媒体取证成为信息安全领域一个重要的研究方向。因此,目前迫切需要能够检测虚假多媒体内容和避免危险虚假信息传播的技术和工具。本文旨在对过去多媒体取证领域所提出的优秀检测取证算法进行总结。除了回顾传统的媒体取证方法,还将介绍基于深度学习的方法。本文针对当今主流的多媒体篡改对象:图像、视频和语音分别进行总结,并针对每种媒体形式,分别介绍传统篡改方法和基于 AI(artificial intelligence)生成的篡改方法,并介绍了已公开的大规模数据集以及相关应用的情况,同时探讨了多媒体取证领域未来可能的发展方向。

关键词:多媒体取证;多媒体溯源;篡改检测;篡改定位;虚假人脸

Overview of digital media forensics technology

Li Xiaolong[1], Yu Nenghai[2], Zhang Xinpeng[3], Zhang Weiming[2], Li Bin[4], Lu Wei[5], Wang Wei[6], Liu Xiaolong[1]

1. Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China; 2. School of Information Science and Technology, University of Science and Technology of China, Hefei 230027, China; 3. School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China; 4. School of Information Engineering, Shenzhen University, Shenzhen 518060, China; 5. School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China; 6. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Abstract: Internet and social networks have become the main platforms for people to access and share various digital media. Among them, media based on images, videos, and audio carry more information and are the most eye-catching. With the rapid development of computer technology, image and video editing software and tools have appeared one after another, such as Photoshop, Adobe Premiere Pro, and VideoStudio. These editing software can be faster and easier to modify

收稿日期:2021-02-08;**修回日期**:2021-03-19;**预印本日期**:2021-03-26

基金项目:国家重点研发计划项目(2020AAA0140003);国家自然科学基金项目(U2001202, 62072480, U1736213)

Supported by: National Key Research and Development Program of China (2020AAA0140003); National Natural Science Foundation of China (U2001202, 62072480, U1736213)

the media. The effect of image forgery is realistic, and the effect of video editing and synthesis is natural and smooth. In recent years, the image generation technology has also been greatly developed, and the visual effects of the generated images may be fake. The problem of multimedia forgery attracts people's attention. The purpose of forgery may be entertainment (such as beautifying images), malicious modification of the content of images and videos (such as deliberately modifying photos of political figures or deliberately exaggerating the severity of news events), and malicious copying. Image forgery incidents in recent years also remind people to focus on the security of media content. The authenticity of visual media content decreases and is increasingly being questioned. At present, millions of multimedia data are transmitted via the Internet every day. What type of content is true? What tampering was made behind the wrong content? The digital forensics technology proposed in recent years provides the answer. This technology does not embed a watermark in advance but directly analyzes the content of multimedia data to achieve the purpose of authenticity recognition. The basic principle is that the inherent characteristics of the original multimedia data are consistent and unique and can be used as its own "intrinsic fingerprint". Any tampering or forgery destroys its integrity to a certain extent. In recent years, media tampering has been increasing and has seriously threatened social stability and even national security. Especially with the rapid development of deep learning technology, the perceived gap between fake media and real media decreases. This finding poses a serious challenge to media forensic research and makes multimedia forensics an important issue in the field of information security research direction. Therefore, technologies and tools that can detect erroneous multimedia content are urgently required, and the spread of dangerous erroneous information is avoided. This article aims to summarize the excellent detection and forensics algorithms proposed in the previous multimedia forensics field. In addition to reviewing traditional media forensics methods, we introduce methods based on deep learning. This article summarizes the current mainstream multimedia tampering objects, namely, images, videos, and sounds. Each media form includes traditional tampering methods and artificial intelligence (AI)-based tampering methods. Among them, video tampering is mainly divided into intraframe tampering and interframe tampering. Intraframe tampering takes the video frame as a unit to delete objects on the screen or performing "copy and move" operations, and interframe tampering takes the video sequence as a unit to add or delete frames. Traditional methods for detecting fake videos can be divided into video encoding tracking detection, video content inconsistency detection, video frame repeated tampering, and copy and paste detection. AI-based error video detection technology focuses on detecting artifacts left over from the network generated in the imaging network, which is different from the imaging process of a real camera. The purpose of digital image forensics technology is to verify the integrity and authenticity of digital images. Image forensic methods can be divided into active methods and passive methods. Active image forensics includes embedding watermarks or signatures in digital images. The passive blind forensic (blind forensics) method is not limited by these factors. It distinguishes images by detecting traces of tampering in the image. Common image forgery and tampering include enhancement, modification, area duplication, splicing, and synthesis. The detection of partial replacement image is divided into the following: 1) Area copy and tamper detection, which copies and pastes part of the area in the image to other areas. During the copying process, the copied area may undergo various geometric transformations and postprocessing. 2) Image processing fingerprints detection. The visual difference caused by simple area copying, splicing, and tampering is still evident. The forger performs postprocessing, such as zooming, rotating, and blurring the image, to eliminate these traces. 3) In recompression fingerprint detection, tampered images inevitably undergo recompression; thus, digital image recompression detection can provide a powerful auxiliary basis for digital image forensics. For the traceability detection technology of forged images, most images are captured by the camera. The general physical structure of the camera and the physical differences between different cameras leave traces on the captured images. These traces (camera fingerprints) appear as a series of features on the image, and the acquisition device of this image can be identified by examining the fingerprint of the device embedded in the image. The detection technology for the overall image generated by AI also focuses on detecting the artifacts left by the network generated in the imaging network. In the previous decades, some digital audio forensic studies have focused on detecting various forms of audio tampering. These methods check the metadata of audio files. In addition, the publicly available large-scale data sets and related applications are introduced, and the possible future development directions of the multimedia forensic field is discussed.

Key words: multimedia forensics; multimedia traceability; forgery detection; forgery localization; fake face

0 引 言

互联网和社交网络已成为人们获取和分享各类数字媒体的主要平台。其中以图像、视频和音频为主的数字媒体承载着巨大的信息量,最为引人注目。目前,随着计算机视觉技术的发展,图像、视频的编辑软件和工具更是层出不穷。这些编辑软件不仅使多媒体修改的过程快速便捷、修改的方式千变万化,而且图像伪造的效果逼真、视频剪辑合成的效果自然流畅。此外,随着深度学习技术的日益成熟,图像生成技术也突飞猛进,生成图像的视觉效果能够以假乱真。可见,数字媒体内容的真实性越来越不能得到保证,其真伪越来越受到质疑,多媒体鉴伪已经成为信息安全领域的一个研究热点和难点。

多媒体伪造问题已经成为研究人员关注的重点。伪造的目的可能是为了娱乐(比如利用美图秀秀等工具对图像进行美颜修饰)、恶意更改图像或视频的内容(比如蓄意修改重要任务的照片或者故意夸大新闻事件的严重程度)、恶意配音等。近年出现的各种关于多媒体伪造的事件也不断提醒着人们关注媒体内容安全。比如,“华南虎”、“广场和平鸽”、青藏铁路“藏羚羊”以及广州“白云山雪景”等伪造图像严重误导了人们的认知。国际上,多媒体伪造的例子也层出不穷,美国战地记者关于伊拉克战争报道的照片拼接事件引发了民众的不信任(图 1 左图),伊朗发布的篡改导弹发射的照片对世界安全造成了威胁(图 1 中图)。多媒体伪造已经涉及政治、科学、新闻、战争和娱乐等诸多领域。据公开报道,美国研究诚信办公室(office of Research Integrity, ORI)主任 John Dahlberg 表示,多媒体伪造是一个“日益显著的问题,需要我们进行解决”。同时,心理学研究也表明有大约 30% 的人会被虚假信息欺骗,这将严重影响公众看待事物的观点(例如图 1 右图的换脸图片),甚至可能会引起严重后果。

图 1 多媒体篡改的典型示例

Fig. 1 Example of multimedia tampering

如今,每天有数以百万计的多媒体数据通过网络传播,到底哪些内容是真实可信的,虚假内容的背后又经历了哪些篡改?近年来提出的数字取证技术将会给出答案,该技术不预先嵌入水印,而是直接分析内容,达到真实性取证的目的。任何类型的篡改都会在媒体文件中留下“痕迹”,这些“痕迹”违背了相机成像的规则,通过检测“痕迹”实现鉴别篡改文件。

本文旨在对过去多媒体取证领域的科研工作者们提出的优秀检测取证定位算法进行总结。除了介绍基于传统方法的多媒体取证方法,还将总结基于深度学习的方法。本文针对当今主流的多媒体篡改对象:图像、视频和语音分别进行总结,并针对每种媒体形式,分别介绍传统篡改方法和基于 AI(artificial intelligence)生成的篡改方法,介绍了已公开的大规模数据集以及相关应用的情况。

1 视频伪造检测技术

视频篡改主要分为帧内篡改和帧间篡改。帧内篡改以视频帧为单位,删除画面中的某个物体,或是做“复制—移动”操作,如图 2 所示。随着硬件技术的发展,篡改者借助深度学习缩小了篡改后的视频与真实视频在视觉上的差距。最近备受关注的深度换脸技术(DeepFakes)就是利用深度学习将视频中的人脸替换为其他人脸,该技术的开源代码,包括详细的使用说明,都可在软件项目托管平台 GitHub 上获取。这样,深度换脸的学习成本和篡改成本极低,篡改者可通过简单的操作,或者借助深度生成网络,如生成对抗网络(generative adversarial network, GAN),直接生成人脸,或者修改人脸的表情和口型等属性信息。

(a) 原图 (b) 物体复制移动 (c) 篡改区域 (d) 篡改图

图 2 “复制—移动”篡改示例

Fig. 2 Example of copy-move forgery

((a) original image; (b) object copy-move; (c) manipulated regions; (d) manipulated image)

帧间篡改则以视频序列为单位,增加或删除帧。随着视频编辑技术的发展,视频篡改变得愈发容易,所以研究能够有效检测视频真伪的取证算法变得尤

为重要。

1.1 针对传统方法伪造视频的检测技术

有些传统方法伪造视频的检测技术通过关键点来对视频进行表示。Laptev(2005)的研究是记录视频在时域上变化比较大的点以反映物体的运动信息。Heng 等人(2013)借助光流追踪图像特征点随时间的变化,作者还提取灰度图像梯度直方图(histogram of oriented gradient, HOG)、光流直方图(histogram of oriented optical flow, HOF)和运动边界直方图(motion boundary histogram, MBH)等作为特征来检测篡改视频。随着深度学习的发展,基于神经网络的方法显示出其优越的性能。Donahue 等人(2015)使用卷积神经网络对每一个帧提取具有判别力的特征,然后使用长短时记忆(long short-term memory, LSTM)网络分析时序信息。双流法(Simonyan 和 Zisserman,2014)则是在光流图上获取时序信息,再与图像 RGB 通道的空间信息进行联合训练来分析视频。

多帧的检测方法可以分为视频编码痕迹检测、视频内容不一致检测、视频帧的重复篡改与复制粘贴检测,简述如下:1)基于视频编码痕迹检测的方法。可以注意到,常见的视频往往是经过压缩的,这些篡改方法首先解码视频,经过篡改后再进行编码。所以篡改视频一定经过多重压缩,这种多次编码会在视频中留下痕迹。Liao 等人(2011)利用量化非零交流系数对视频的二次压缩进行检测,具体就是以这些检测结果为基础,再结合其他技术进一步分析。Stamm 等人(2012)对运动向量的统计特征进行更为深入的分析,将经过处理的特征用于机器学习算法来提高检测效果。由于实际应用中数字视频多数经过压缩编码,因此直接利用视频压缩域特征进行篡改检测的方法应用更为有效,近年来越发受到研究者重视。2)基于视频内容不一致的检测方法。当视频遭到篡改时,视频内容将不可避免地出现异常。Wang 和 Farid(2007a)利用相关系数矩阵及相位谱矩阵来分别对视频的帧重复篡改与帧内区域重复篡改进行检测。Wang 和 Farid(2007b)也研究了篡改操作对隔行扫描视频的影响,根据其去隔行效应特征对视频篡改进行检测。3)基于视频帧的重复篡改与复制粘贴检测方法。林晶等人(2016)利用了量化离散余弦变换系数以及相似度分析等方法进行检测。

1.2 针对基于 AI 生成的伪造视频的检测技术

随着深度学习技术的发展,多媒体内容篡改和网络造谣等问题日益凸显,世界各国纷纷加大了对虚假媒体取证的研究投入。

DeepFake 是一款利用深度学习将视频中的人脸替换为目标人脸的技术,其基本原理如图 3 所示。这类 AI 造假技术给虚假媒体检测带来很大的挑战。

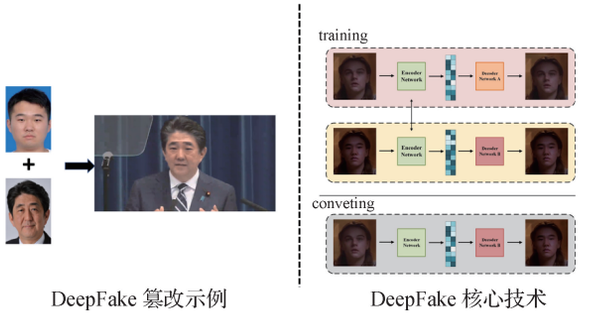


图 3 DeepFake 核心思想

Fig. 3 Core idea of DeepFake forgery

研究人员利用 AI 生成的视频仍不完美的特性,给出利用人脸图像中的对称性、牙齿和眼睛细节等特定特征检测人脸图像是否为 AI 生成的方法(Marra 等, 2018; Korshunov 和 Marcel, 2018)。Li 等人(2018a)提出了基于高维统计特征的检测方法。国外学者提出利用 DeepFake 生成的视频中,人没有眨眼、呼吸等特征,能够以很高的准确率识别出假视频(Afchar 等,2018;Güera 和 Delp,2018)。基于生物特征的检测中,Li 等人(2018b)利用视频中人物眨眼频率的生理特征,提出了长期循环神经网络(long-term recurrent convolutional network, LRCN)模型; Ciftci 等人(2020)用远程光电体积描记法(remote photoplethysmograph, rPPG)来捕捉 RGB 视频中微弱的颜色和运动变化,再结合分类器来判别伪造视频;Yang 等人(2019)通过人脸区域内部的特征点与人脸区域内部边界的特征点估计出两个头部方向,若这两个方向的夹角大于阈值,则判定为虚假图像。以 Afchar 等人(2018)的研究为代表,在基于视频语义特征来检测的方法中,考虑到视频经过压缩后,低层噪声分布特征发生改变,中高层人脸特征不能反映伪造视频制作痕迹,因此,利用神经网络的中层语义来判别伪造视频。作者在 AI 换脸(DeepFake)视频数据集与面部重现(Face2Face)视频数据集上取得了良好的分类效果。图像噪声在视频压缩的环境下严重退化,所以基于图像噪声的微观分析

在这种情况下无效。在更高的语义层次上,人类很难分辨真假。所以采用了一种折衷的方法,即使用浅层神经网络(shallow network)进行检测。

Agarwal 和 Varshney(2019)将 GAN 生成媒体的检测转化为一个假设检验问题,并结合统计分析框架来定义真实图像与深度伪造图像之间的距离。也有基于单帧图像的两阶段模型(Hsu 等,2020),检测计算机生成人脸与自然人脸(Dang-Nguyen 等,2012;Rahmouni 等,2017),人脸形变问题(Raghavendra 等,2017)等。来自美国加州大学伯克利分校的 Farid(2019)认为,利用深度学习生成的人脸是有固定模式的,但是每个人有其独特的面部微表情,Farid 将人面部表情和动作用人脸特征点表示出来,用于表示特定个人的说话模式。尽管这些相关性在视觉上并不明显,但它们在伪造过程中常常被忽视,因此可以用于 DeepFake 伪造视频检测。说话人具有其独特的面部表情和动作,通过处理连续的视频帧,记录面部与头部特定区域的运动轨迹。处理的最小元组为 10 s 的片段,每帧提取 19 维特征,共提取 190 维特征,再使用支持向量机(support vector machine, SVM)进行分类。该方法的研究人员利用 FaceForensics ++数据库中 1 000 条视频数据与某些国家领导人的视频进行训练和测试,使用 SVM 作为分类器。通过不同特征点的组合,最终达到了 99% 的准确率。

1.3 主要伪造视频样本库

1.3.1 DFDC 数据库

Facebook 等互联网公司以及科研机构组织在 Kaggle 平台上发布了 DeepFake Detection Challenge (DFDC)AI 生成假脸的检测比赛,全球近 2 300 支团队参加比赛。

DFDC 公布的数据集包括多达 12 万个视频,单个视频时长约为 10 s,帧率范围为 15 ~ 30 帧/s,分辨率范围为 320 × 240 ~ 3 840 × 2 160 像素。训练视频中有大约 2 万个视频为真实视频,10 万个假脸视频。真实视频由 430 名演员拍摄,在此基础上,使用多种假脸生成算法生成假脸。

1.3.2 FaceForensics ++

在深度换脸领域,最流行的数据库之一是 FaceForensics ++ ,该数据集于 2019 年公开。

FaceForensics ++ 包含了 1 000 个真实视频,使用了 DeepFake、Face2Face、FaceSwap 等 3 种人脸篡改方法生成假脸。对于 FaceSwap、DeepFake 的假视频,分别使用计算机图形学和深度学习相关的方法生成,这些算法都具有良好的生成效果,并且都已在 GitHub 上开源。Face2Face (Justus 等, 2016) 是 2016 年由科研人员在学术会议 CVPR(IEEE Comference on Computer Vision and Pattern Recognition)上提出的。FaceForensics ++ 数据库中,每种假脸方法生成了 1 000 个假视频。随后,在谷歌的支持下,FaceForensics ++ 中增加了 DeepFakeDetection 数据集和 NeuralTextures 数据集。DeepFakeDetection 数据集包含了在 16 个不同场景中 28 个演员录制的 363 个视频,以及 3 000 多个换脸视频。NeuralTextures 数据集优化了 Face2Face 数据集的视觉效果。值得注意的是,FaceForensics ++ 数据库提供了 3 种不同压缩情况下的视频,为科研人员提供强大的数据支持。

1.3.3 Celeb-DF

2019 年 11 月,研究人员结合反取证技术,提出了一个具有挑战性的大规模 DeepFake 视频数据集—Celeb-DF(Li 等,2020),如图 4 所示。Celeb-DF 包含 590 个真实视频,共 225.4 k 帧,以及 5 639 个虚假视频,共 2 116.8 k 帧。

图 4 Celeb-DF 数据集

Fig. 4 Celeb-DF dataset

1.3.4 DeeperForensics-1.0

新加坡南洋理工大学的研究人员构建了一个大规模的人脸伪造检测数据集(Jiang 等, 2020),包含近 6 万个视频。合成视频具有更高的多样性,让检测算法更注重于挖掘篡改方法留下的痕迹。数据集是由一个新提出的端到端人脸交换框架生成的。用户调查显示,生成的视频质量相较于目前的数据集效果更好。

2 图像伪造检测技术

数字图像取证技术对数字图像的完整性和真实性进行验证,方法总体可以分为主动式方法和被动式方法。主动式的图像取证要在数字图像中嵌入水印或签名。而被动式的盲取证(blind forensics)方法则不受这些因素的限制,它通过检测篡改图像中的操作痕迹来鉴别图像。常见的图像伪造和篡改包括增强、润饰、区域复制和拼接合成等。总体来说,图像篡改一般要经历 4 个操作步骤:

1)获取原始图像;

2)执行篡改操作;

3)后处理;

4)重编码、压缩操作。

各种操作都会留下篡改痕迹,图像取证技术则通过检测这些痕迹判断图像是否经过篡改,以及经历过何种篡改。数字图像取证可以分为设备指纹检测、区域复制篡改检测、图像处理指纹检测和重压缩指纹检测等几种。

2.1 针对局部替换图像的检测技术

局部替换图像的检测分为:1)区域复制移动篡改检测。把图像中的部分区域复制并粘贴到其他区域。在复制过程中,复制区域可能会经历多种几何变换与后期处理。2)图像处理指纹检测。简单的区域复制、拼接篡改带来的视觉差异仍较为明显,伪造者使用缩放、旋转、模糊等后处理抹除这些痕迹。3)重压缩指纹检测。篡改图像必然会经过重压缩,因此检测重压缩痕迹能有效地检测篡改图像。

2.1.1 区域复制篡改检测

Cozzolino 等人(2015)针对检测中的块匹配方法,使用 Patch Match 思想来进行相似块的匹配,提高了算法的运行效率,但是该算法并不鲁棒,无法抵抗旋转、缩放等攻击。Li 等人(2015)在关键点提取前对图像分割,降低了匹配任务的难度,大大提高了匹配的精确度。

2.1.2 图像处理指纹检测

Popescu 和 Farid(2005)通过图像相邻像素之间周期性的线性关系来鉴别篡改图像。苏文煊和方针(2019)利用 CFA(color filter array)插值特性检测篡改痕迹,使用 CFA 插值特性的变化作为特征,并根据相邻块间的 CFA 插值特征的不一致来检测篡改图像。

图像对比度增强可以调整全局亮度。增强图像局部的对比度让合成的图像更真实。Cao 等人(2014)通过从 JPEG 压缩和像素值映射产生的直方图峰值、间隙中零高间隙(zero-height gap)指纹来识别全局对比度增强。高铁杠等人(2016)使用基于超像素和游程直方图来检测图像对比度修改。中值滤波常用来平滑图像拼接的边缘。杨晓花(2018)提出一种使用图像像素间相关性的数字图像盲取证算法。该方法通过设定阈值,对像素级区域的聚类结果进行筛选,进而检测模糊操作的痕迹,进行篡改区域定位。

2.1.3 重压缩指纹检测

Galvan 等人(2014)以 DCT(discrete cosine transform)系数矩阵水平、垂直、对角和反对角方向差分的高阶马尔可夫转移概率作为分类特征实现 JPEG 重压缩图像检测。Thai 等人(2017)将量化效应和 DCT 系数统计相结合,对先前压缩并存储为无损的图像进行量化步长的估计。Tagliasacchi 等人(2013)给出了一个一般估计使用的图像变换技术和量化步长的方法。

2.2 针对伪造图像的溯源检测技术

大部分图像都是经相机拍摄采集的,相机的一般物理结构以及不同相机之间的物理差异会在拍摄的图像上留下痕迹,这些痕迹(相机指纹)在图像上表现为一系列特征。通过考察图像上所嵌有的设备指纹可以识别这个图像的获取设备,这种方法称为设备指纹检测。

由于每一次加工中出现的变化以及加工工艺无法达到百分之百完美,相机中的每一个感光单元都有着细微的差异。因此,即使在相同的光照强度下,图像中的不同像素也可能被赋予不同的值,这种不均衡性称为非一致模式噪声(photo reshponse non-uniformity, PRNU)。PRNU 是每个图像获取设备所固有的独一无二的设备指纹,因此基于 PRNU 的数字取证具有广泛的应用场景。基于 PRNU 的相机识别通常先利用图像去噪的方法估计 PRNU。由于 PRNU 可能会受到图像中其他噪声以及图像内容的干扰,如何精确估计和识别 PRNU 便是此类研究的一个重点。Chierchia 等人(2014a)利用适应性权重重新定义判断方程来提高基于 PRNU 的取证方法的准确率,该方法可有效应用于小尺寸拼接图像的设

备源识别中。另外，Chierchia 等人(2014b)利用 Markov 随机场来描述 PRNU 的统计分布，并利用 Bayesian 判决来提高识别 PRNU 的准确性。然而，PRNU 的提取不可避免地受到图像内容和其他噪声的影响，因此 Lawgaly 和 Khelifi (2017)在提取的各个阶段做出了改进，包括在滤波阶段改进了局部自适应离散余弦变换过滤器，在估计阶段提出了新技术 WA (weighted averaging)，在后处理阶段联合不同的彩色通道综合估计，取得了更好的效果。

设备指纹检测可以用来鉴别图像是否来自于合法的设备，除此之外，同一图像上指纹的不一致，也可以作为图像被拼接过的证据(吴韵清 等，2019)。由于拼接篡改可能会使用来自不同图像的区域，如果这些图像来自不同的设备，便会呈现出不同的设备指纹。然而，这类方法无法检测同一图像的区域复制，例如区域复制操作。此外，设备指纹信号通常很微弱，其检测结果很容易受到各种因素的影响。

2.3 针对 AI 整体生成图像的检测技术

McCloskey 和 Albright(2018)分析了 GAN 的生成过程，以检测真实和虚假图像之间的不同伪影。作者基于真实图像与虚假图像在颜色空间的差异性提取特征进行分类。随后，Yu 等人(2019)分析了 GAN 指纹的存在性和唯一性，以检测假图像。特别地，他们使用探测了不同网络结构特有的结构指纹。因此，他们学习了每个生成模型的指纹，使用图像指纹和每个模型指纹之间的相关指数进行分类。Wang 等人(2019)提出了一种通过监控网络神经元行为来识别 AI 生成假脸的方法。作者提出了一种用于逐层捕获神经元激活行为的神经元覆盖准则(mean neuron coverage, MNC)。最后，FakeSpotter 用一个简单的二值分类器区分 4 种不同类型的假脸(整体生成、属性修改、表情修改和 Deepfakes 换脸)。

与传统基于学习的方法相比，FakeSpotter 的输入不是最终层神经元的输出，而是将各层神经元的行为作为特征。研究表明，被激活的神经元能够很好地感知输入的细微特征，这些细微特征能够鉴别真实面部图像和合成面部图像。研究人员利用真实人脸样本 CelebA(CelebFaces Attributes Dataset)与 FFHQ(Flickr-Faces-High-Quality)，以及虚假人脸样本 FaceForensics++、DFDC、Celeb-DF 以及 StyleGAN2 生成的人脸数据进行训练和测试，数据划分比例为 5∶1。经过对测试集的测试，最终在 4 种篡改方法的检测结果分别为：整体生成 98.6%；属性修改 90.1%；表情篡改 100%；Deepfakes 换脸 97.8%。

来自德国弗里德里希亚历山大大学纽伦堡分校的 Matern 等人(2019)发现，现有的 Deepfakes 篡改的视频虽然看起来很真实，但总会在细节上有些许差异，许多人脸编辑算法呈现出类似于经典计算机视觉问题的伪影，这种视觉上的伪影可以用于鉴别 Deepfakes 伪造视频。

如图 5 所示，Deepfakes 伪造视频、图像存在伪影，具体表现在：双眼不一致性图 5(a)、反射缺失图 5(b)、牙齿细节图 5(c)、人脸拼接图 5(d)、光照估计和鼻子几何形状不精确图 5(e)等方面。这些伪影是计算机视觉一直存在的问题，这些问题仍然没有完全解决。通过识别这些视觉伪影，能够有效地鉴别虚假图像。研究人员经过对测试集的测试，最终在虚假人脸图像识别的最优准确率为 86.6%。

图 5 伪造伪影

Fig. 5 Forgery artifacts((a) binocular inconsistency; (b) reflex deficiency; (c) dental details; (d) face splicing; (e) inaccurate light estimation and nose geometry)

2.4 主要伪造图像样本库

2.4.1 CoMoFoD

数字图像取证的目的是确定数字图像的真实性。复制移动是最常用的伪造方法之一，通过将图像的一部分复制到同一图像的另一个位置来实现。数字图像中的复制移动伪造检测(copy-move forgery detection, CMFD)仍然是一个亟需解决的问题。虽然现有的复制移动伪造检测算法已经不少，但是可用于算法评估的基准数据库很少。用于复制移动伪造检测的 CoMoFoD 数据库包含 260 个伪造的图像集，分为两类(512×512 像素和 3 000×2 000 像素)。根据应用的操作将图像分为 5 类：平移、旋

转、缩放、组合和变形。对所有伪造和原始图像应用不同类型的后处理方法，如 JPEG 压缩、模糊、噪声添加和颜色减少等。

2.4.2 GRIP

GRIP 包含 80 幅原始图像和 80 幅逼真的复制移动篡改图像，大小都是 768 × 1 024 像素。值得注意的是，GRIP 中的一些篡改补丁非常光滑，这对基于稀疏采样（如 SIFT（scale-invariant feature transform））的复制—移动篡改检测是一个挑战。

3 音频伪造检测技术

音频伪造最初研究从文本到语音（text-to-speech，TTS）的转换。音频伪造技术主要包括：1）拼接法；2）参数法；3）混合法和 4）基于人工智能的方法。拼接式语音合成方法主要将多个语音词典中的单个词或词组按照语法拼接。参数法首先从文本中提取声码器能够识别的特征，进而使用声码器生成音频。常见的参数法 TTS 技术是基于隐马尔可夫模型来实现音频合成。混合法是拼接法和参数法的结合。现在基于人工智能的音频伪造技术逐渐成为热点，研究人员通过分析各类媒体之间的共性，结合图像、视频处理领域的经验，提出了基于人工智能的语音合成方法，包含基于生成对抗网络（GAN）、自编码器（autoencoder，AE）、自回归模型（autoregressive model，AR）等音频伪造技术。

语音转换（voice conversion）技术是指将一个人的声音变成另一个人的声音，同时保持说话内容不变。语音转换方法主要为声道谱转换方法。其中语音转换的研究主要集中在如何对声道谱进行建模和设计更有效的映射规则。目前，对声道谱转换模型的方法主要是先对语音进行统计分析，再通过参数映射的方式实现转换。声道谱转换方法包括基于码书映射的转换方法、基于高斯混合模型的转换方法、基于隐马尔可夫模型的转换方法、基于频率弯折的转换方法、基于神经网络的转换方法和基于波形生成的转换方法等 6 种转换技术。声道谱转换映射的研究突破了训练需要大量语音数据量、平行语音的限制，效率与质量也得到了提高，但是目前的转换技术仍有不足之处，所以声道谱转换是语音转换中需要重点解决的问题。

传统的音频编辑方法常见的有音频片段的删除、插入、替代和拼接，此外还有上采样篡改、下采样篡改和压缩编码篡改等。

3.1 针对传统伪造语音的检测技术

在过去的几十年里，一些数字音频取证研究致力于检测各种形式的音频篡改（Zakariah 等，2018）。这些方法检查音频文件的元数据（Koenig 和 Lacey，2012），并检查来自 3 个 Olympus 记录器的 11 个音频记录的数字标头数据以进行音频篡改检测。Zhao 和 Malik（2013）提出通过将声环境特征作为检测音频伪造的重要特征验证数字音频的完整性。

3.2 针对 AI 生成语音的检测技术

AlBadawy 等人（2019）首先致力于人工智能合成的假声音的检测研究。在他们的工作中，提出了一种双谱分析方法来检测人工智能合成的假声音。他们观察到，用 DNN（deep neural network）合成的假声音中显示出特殊而不寻常的光谱相关性，其称为双谱伪影。因此，他们探索利用高阶多光谱特征对伪声音进行识别。Bishop（1994）探讨了一种方法，即利用简单的二分类器来监测基于 DNN 的说话人识别系统中神经元的行为来区分真伪声音。神经元的分层行为可以捕捉到区分真实和虚假声音时更细微的特征。

3.3 主要伪造语音样本库

3.3.1 ASVspoof 2019 数据集

ASVspoof（Todisco 等，2019）于 2019 年提出，用于第 3 次自动说话者验证欺骗和对策挑战的数据库（http://www.asvspoof.org）。SVspoof 2019 数据库包含两个部分，用于评估逻辑访问（logical access，LA）和物理访问（physical acces，PA）方案。两者均来自 VCTK（the voice cloning toolkit）基本语料库，该语料库为从 107 位说话者（男 46 位，女 61 位）中录制的语音数据。LA 和 PA 数据库本身都分为 3 个数据集，即训练集、测试集和验证集，包括来自 20 位（8 位男性，12 位女性），10 位（4 位男性，6 位女性）和 48 位（21 位男性，27 位女性）演讲者的语音数据。就说话人而言，这 3 个分区是不相交的，并且所有源数据的记录条件都相同。训练集和测试集包含使用相同算法/条件生成的欺骗攻击（称为已知攻击），而验证集还包含使用不同算法/条件生成的欺骗攻击（称为未知攻击）。

3.3.2 TIMIT 数据集

TIMIT （The DARPA TIMIT Acoustic-Phonetic

Continuous Speech Corpus)是由美国麻省理工学院(Massachusetts Institute of Technology, MIT)、斯坦福研究所(stanford Research Institute, SRI)和德州仪器公司(Texas Instruments, TI)共同发布的语音语料库。TIMIT数据集一共包含6 300段语音,由630人录制,采样频率为16 kHz,所有的句子都进行了手动分割、标记。数据集中,70%的人为男性,30%的人为女性;以白人居多。

3.3.3 RSR2015数据集

RSR2015 (The Robust Speaker Recognition 2015)(Larcher等,2012)数据库包含了超过71 h的英语使用者的讲话记录,涵盖了新加坡的各种口音。演讲者由300名参与者(143名女性和157名男性)组成,年龄从17岁到42岁。

4 结 论

可以清楚地看到,在过去的几十年中,多媒体取证的研究得到了很大的发展。特别是华人学者在该方向也取得了很大的学术成就,例如,中国科技大学俞能海和张卫明教授团队在2020年的DeepFake检测大赛中取得了第2名的好成绩,并且与第1名的检测结果相差无几。然而,许多问题仍然没有可靠的方案来解决,新的挑战每天都有。当然,深度学习的出现给媒体操纵方法和取证工具带来了巨大的推动力,开辟了新的研究领域。然而,更根本的原因是这个研究领域具有两方参与的性质。篡改方法的迭代更新往往使得旧有检测方法失效,研究者们需要提出更具鲁棒性、泛化性的方案来应对不可预见的威胁。在这个前提下,努力找出未来研究最有前景的领域是很重要的。当前多媒体取证未来可能的研究方向包括:

1)随着操作变得越来越聪明,单个工具对付各种攻击的效率将越来越低。因此,多种检测工具、多种网络、多种方法必须一起工作,而如何更好地结合所有可用的信息片段应该是一个更持久的研究目标。

2)基于深度学习方法的可解释性研究。深度学习的黑箱特性使得人们很难理解为什么会做出某个决定。深度网络可以正确地将猫的图像分类为猫,但不知道是哪些特定特征促使了这一决定。当然,对于一些取证应用程序来说,这是一个严重的问题。例如,法官几乎不会只根据统计数据作出决定。更一般地说,能够追踪深度网络的推理将有助于改进其设计和训练阶段,并对恶意攻击提供更高的鲁棒性。尽管多媒体取证有着悠久的历史,但它似乎还处于全面发展的阶段,工业和社会对多媒体取证有着很高的要求,很多应用面临着落地的问题。

参考文献(References)

Afchar D, Nozick V, Yamagishi J and Echizen I. 2018. MesoNet: a compact facial video forgery detection network//Proceedings of 2018 IEEE International Workshop on Information Forensics and Security. Hong Kong, China: IEEE: 1-7 [DOI: 10.1109/WIFS.2018.8630761]

Agarwal S, Farid H, Gu Y and He M M and Nagano K and Li H. 2018. Protecting World Leaders Against Deep Fakes//Proceedings of 2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Long Beach, USA: IEEE: 38-45

Agarwal S and Varshney L R. 2019. Limits of deepfake detection: a robust estimation viewpoint [EB/OL]. [2020-10-29]. https://arxiv.org/pdf/1905.03493v1.pdf

AlBadawy E A, Lyu S W and Farid H. 2019. Detecting Ai-synthesized speech using bispectral analysis//Proceedings of 2019 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Long Beach, USA: IEEE: 104-109

Bishop C M. 1994. Mixture Density Networks. Birmingham: Aston University

Cao G, Zhao Y, Ni R R and Li X L. 2014. Contrast enhancement-based forensics in digital images. IEEE Transactions on Information Forensics and Security, 9(3): 515-525 [DOI: 10.1109/TIFS.2014.2300937]

Chierchia G, Cozzolino D, Poggi G, Sansone C and Verdoliva L. 2014a. Guided filtering for PRNU-based localization of small-size image forgeries//Proceedings of 2014 IEEE International Conference on Acoustics, Speech and Signal Processing. Florence, Italy: IEEE: 6231-6235 [DOI: 10.1109/ICASSP.2014.6854802]

Chierchia G, Poggi G, Sansone C and Verdoliva L. 2014b. A Bayesian-MRF approach for PRNU-based image forgery detection. IEEE Transactions on Information Forensics and Security, 9(4): 554-567 [DOI: 10.1109/TIFS.2014.2302078]

Ciftci U A, Demir I and Yin L J. 2020. FakeCatcher: detection of synthetic portrait videos using biological signals. IEEE Transactions on Pattern Analysis and Machine Intelligence, (1): #3009287 [DOI: 10.1109/TPAMI.2020.3009287].

Cozzolino D, Poggi G and Verdoliva L. 2015. Efficient dense-field copy-move forgery detection. IEEE Transactions on Information Forensics and Security, 10(11): 2284-2297 [DOI: 10.1109/TIFS.2015.2455334]

Dang-Nguyen D T, Boato G and De Natale F G B. 2012. Identify computer generated characters by analysing facial expressions variation//Proceedings of 2012 IEEE International Workshop on Information

Forensics and Security. Costa Adeje, Spain: IEEE: 252-257 [DOI: 10.1109/WIFS.2012.6412658]

Donahue J, Hendricks L A, Guadarrama S, Rohrbach M, Venugopalan S, Darrell T and Saenko K. 2015. Long-term recurrent convolutional networks for visual recognition and description//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE: 2625-2634 [DOI: 10.1109/CVPR.2015.7298878]

Galvan F, Puglisi G, Bruna A R and Battiato S. 2014. First quantization matrix estimation from double compressed JPEG images. IEEE Transactions on Information Forensics and Security, 9(8): 1299-1310 [DOI: 10.1109/TIFS.2014.2330312]

Gao T G, Yang L, Xuan Y and Tong J. 2016. Contrast modification forensic algorithm based on superpixel and histogram of run length. Journal of Electronics and Information Technology, 38(11): 2787-2794 (高铁杠, 杨亮, 宣妍, 佟静. 2016. 基于超像素和游程直方图的对比度修改检测算法. 电子与信息学报, 38(11): 2787-2794) [DOI: 10.11999/JEIT160161]

Güera D and Delp E J. 2018. Deepfake video detection using recurrent neural networks//Proceedings of the 15th IEEE International Conference on Advanced Video and Signal Based Surveillance. Auckland, New Iealand: IEEE: 1-6 [DOI: 10.1109/AVSS.2018.8639163]

Heng W, Kläser A, Schmid C and Liu C L. 2013. Dense trajectories and motion boundary descriptors for action recognition. International Journal of Computer Vision, 103(1): 60-79 [DOI: 10.1007/s11263-012-0594-8]

Hsu C C, Zhuang Y X and Lee C Y. 2020. Deep fake image detection based on pairwise learning. Applied Sciences, 10(1): #370 [DOI: 10.3390/app10010370]

Jiang L M, Li R, Wu W N, Qian C and Loy C C. 2020. DeeperForensics-1.0: a large-scale dataset for real-world face forgery detection//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 2889-2898 [DOI: 10.1109/CVPR42600.2020.00296]

Justus T, Michael Z, Marc S, Christian T and Matthias N. 2016. Face2face: real-time face capture and reenactment of rgb videos//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 2387-2395 [10.1109/CVPR.2016.262]

Koenig B E and Lacey D S. 2012. Forensic authenticity analyses of the header data in re-encoded WMA files from small Olympus audio recorders. Journal of the Audio Engineering Society, 60(4): 255-265

Korshunov P and Marcel S. 2018. DeepFakes: a new threat to face recognition? Assessment and detection [EB/OL]. [2021-02-08] https://arxiv.org/pdf/1812.08685.pdf

Laptev I. 2005. On space-time interest points. International Journal of Computer Vision, 64(2/3): 107-123 [DOI: 10.1007/s11263-005-1838-7]

Larcher A, Lee K A, Ma B and Li H Z. 2012. RSR2015: Database for text-dependent speaker verification using multiple pass-phrases//Proceedings of the 13th Annual Conference of the International Speech Communication Association. Portland, USA: IEEE: 1580-1583.

Lawgaly A and Khelifi F. 2017. Sensor pattern noise estimation based on improved locally adaptive DCT filtering and weighted averaging for source camera identification and verification. IEEE Transactions on Information Forensics and Security, 12(2): 392-404 [DOI: 10.1109/TIFS.2016.2620280]

Li H D, Chen H, Li B and Tan S Q. 2018a. Can forensic detectors identify GAN generated images?//Proceedings of 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. Honolulu, USA: IEEE: 722-727 [DOI: 10.23919/APSIPA.2018.8659461]

Li J, Li X L, Yang B and Sun X M. 2015. Segmentation-based image copy-move forgery detection scheme. IEEE Transactions on Information Forensics and Security, 10(3): 507-518 [DOI: 10.1109/TIFS.2014.2381872]

Li Y Z, Chang M C and Lyu S W. 2018b. In ictu oculi: exposing AI created fake videos by detecting eye blinking//Proceedings of 2018 IEEE International Workshop on Information Forensics and Security. Hong Kong, China: IEEE: 1-7 [DOI: 10.1109/WIFS.2018.8630787]

Li Y Z, Yang X, Sun P and Qi H G and Lyu S W. 2020. Celeb-df: a large-scale challenging dataset for deepfake forensics//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Virtual: IEEE: 3207-3216 [10.1109/CVPR42600.2020.00327]

Liao D D, Yang R, Liu H M, Li J and Huang J W. 2011. Double H.264/AVC compression detection using quantized nonzero AC coefficients//Proceedings of SPIE 7880, Media Watermarking, Security, and Forensics III. San Francisco Airport, USA: SPIE: 78800Q

Lin J, Huang T Q, Lai Y C and Lu H N. 2016. Detection of continuously and repeated copy-move forgery to single frame in videos by quantized DCT coefficients. Journal of Computer Applications, 36(5): 1356-1361 (林晶, 黄添强, 赖玥聪, 卢贺楠. 2016. 采用量化离散余弦变换系数检测视频单帧连续多次复制—粘贴篡改. 计算机应用, 36(5): 1356-1361) [DOI: 10.11772/j.issn.1001-9081.2016.05.1356]

Marra F, Gragnaniello D, Cozzolino D and Verdoliva L. 2018. Detection of GAN-generated fake images over social networks//Proceedings of 2018 IEEE Conference on Multimedia Information Processing and Retrieval. Miami, USA: IEEE: 384-389 [DOI: 10.1109/MIPR.2018.00084]

Matern F, Riess C and Stamminger M. 2019. Exploiting visual artifacts to expose deepfakes and face manipulations//Proceedings of 2019 IEEE Winter Applications of Computer Vision Workshops (WACVW). Waikoloa, USA: IEEE: 83-92. [DOI: 10.1109/WACVW.2019.00020]

McCloskey S and Albright M. 2018. Detecting GAN-generated imagery using color cues [EB/OL]. [2021-02-08]. https://arxiv.org/pdf/1812.08247.pdf

Popescu A C and Farid H. 2005. Exposing digital forgeries by detecting traces of resampling. IEEE Transactions on Signal Processing, 53(2): 758-767 [DOI: 10.1109/TSP.2004.839932]

Raghavendra R, Raja K B, Venkatesh S and Busch C. 2017. Transfera-

ble Deep-CNN features for detecting digital and print-scanned morphed face images//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, USA: IEEE: 1822-1830 [DOI: 10.1109/CVPRW.2017.228]

Rahmouni N, Nozick V, Yamagishi J and Echizen I. 2017. Distinguishing computer graphics from natural images using convolution neural networks//Proceedings of 2017 IEEE Workshop on Information Forensics and Security. Rennes, France: IEEE: 1-6 [DOI: 10.1109/WIFS.2017.8267647]

Simonyan K and Zisserman A. 2014. Two-stream convolutional networks for action recognition in videos//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: ACM: 568-576

Stamm M C, Lin W S and Liu K J R. 2012. Temporal forensics and anti-forensics for motion compensated video. IEEE Transactions on Information Forensics and Security, 7(4): 1315-1329 [DOI: 10.1109/TIFS.2012.2205568]

Su W X and Fang Z. 2019. Identifying image authenticity based on CFA inconsistency of interpolation characteristics. Journal of Applied Sciences, 37(1): 33-40 (苏文煊, 方针. 2019. 基于 CFA 插值特性不一致的图像真伪鉴别. 应用科学学报, 37(1): 33-40) [DOI: 10.3969/j.issn.0255-8297.2019.01.004]

Tagliasacchi M, Visentini-Scarzanella M, Dragotti P L and Tubaro S. 2013. Transform coder identification//Proceedings of 2013 IEEE International Conference on Acoustics, Speech, and Signal Processing. Vancouver, Canada: IEEE: 5785-5789 [DOI: 10.1109/ICASSP.2013.6638773]

Todisco M, Wang X, Vestman V, Sahidullah M, Delgado H, Nautsch A, Yamagishi J, Evans N, Kinnunen T and Lee K A. 2019. Asvspoof 2019: future horizons in spoofed and fake audio detection. [EB/OL]. [2021-02-08]. https://arxiv.org/pdf/1904.05441.pdf

Thai T H, Cogranne R, Retraint F and Doan T N C. 2017. JPEG quantization step estimation and its applications to digital image forensics. IEEE Transactions on Information Forensics and Security, 12(1): 123-133 [DOI: 10.1109/TIFS.2016.2604208]

Wang R, Juefei-Xu F, Ma L, Xie X F, Huang Y H, Wang J and Liu Y. 2019. FakeSpotter: a simple yet robust baseline for spotting AI-synthesized fake faces [EB/OL]. [2020-10-29]. https://arxiv.org/pdf/1909.06122.pdf

Wang W H and Farid H. 2007a. Exposing digital forgeries in video by detecting duplication//Proceedings of the 9th Workshop on Multimedia and Security. Dallas, USA: ACM: 35-42 [DOI: 10.1145/1288869.1288876]

Wang W H and Farid H. 2007b. Exposing digital forgeries in interlaced and deinterlaced video. IEEE Transactions on Information Forensics and Security, 2(3): 438-449 [DOI: 10.1109/TIFS.2007.902661]

Wu Y Q, Wu P, Chen B J, Ju X W and Gao Y. 2019. Image splicing localization method based on fully convolutional residual networks. Journal of Applied Sciences, 37(5): 651-662 (吴韵清, 吴鹏, 陈北京, 鞠兴旺, 高野. 2019. 基于残差全卷积网络的图像拼接定位算法. 应用科学学报, 37(5): 651-662) [DOI: 10.3969/j.issn.0255-8297.2019.05.007]

Yang X, Li Y Z and Lyu S W. 2019. Exposing deep fakes using inconsistent head poses//Proceedings of 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019). Brighton, UK: IEEE: 8261-8265 [DOI: 10.1109/ICASSP.2019.8683164]

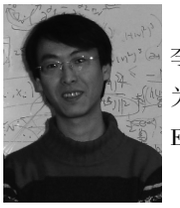
Yang X H. 2018. Blind digital image forensics based on correlation detection algorithm. Microelectronics and Computer, 35(4): 114-118 (杨晓花. 2018. 基于相关性检测的数字图像盲取证算法仿真. 微电子学与计算机, 35(4): 114-118) [DOI: 10.19304/j.cnki.issn1000-7180.2018.04.023]

Yu N, Davis L and Fritz M. 2019. Attributing fake images to GANs: learning and analyzing GAN fingerprints//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea(South): IEEE: 7555-7565 [DOI: 10.1109/ICCV.2019.00765]

Zakariah M, Khan M K and Malik H. 2018. Digital multimedia audio forensics: past, present and future. Multimedia Tools and Applications, 77(1): 1009-1040 [DOI: 10.1007/s11042-016-4277-2]

Zhao H and Malik H. 2013. Audio recording location identification using acoustic environment signature. IEEE Transactions on Information Forensics and Security, 8(11): 1746-1759 [DOI: 10.1109/TIFS.2013.2278843]

作者简介



李晓龙,1976 年生,男,教授,主要研究方向为多媒体内容安全和图像处理。
E-mail:lixl@bjtu.edu.cn

俞能海,男,教授,主要研究方向为多媒体信息检索、图像处理与视频通信、数字媒体内容安全。
E-mail:ynh@ustc.edu.cn

张新鹏,男,教授,主要研究方向为多媒体信息安全、AI 安全、图像处理。E-mail:zhangxinpeng@fudan.edu.cn

张卫明,男,教授,主要研究方向为信息隐藏、数据隐私保护和人工智能安全。E-mail:zhangwm@ustc.edu.cn

李斌,男,教授,主要研究方向为多媒体内容安全、智能信息处理、机器学习。E-mail:libin@szu.edu.cn

卢伟,男,教授,主要研究方向为多媒体内容安全、多媒体信号处理、模式识别和机器学习。
E-mail:luwei3@mail.sysu.edu.cn

王伟,男,副研究员,主要研究方向为多媒体内容安全、多模态内容分析与理解、人工智能安全。
E-mail:wei.wong@ia.ac.cn

刘晓龙,男,博士研究生,主要研究方向为多媒体内容安全和图像处理。E-mail:18112008@bjtu.edu.cn