# 人脸反欺诈活体检测综述

卢子谦[1], 陆哲明[1], 沈冯立[1], 王总辉[2]

[1]浙江大学航空航天学院 杭州 中国 310027

[2]浙江大学计算机系统结构与网络安全研究所 杭州 中国 310027

**摘要** 活体检测技术被广泛应用于信息安全、金融服务、文体娱乐、智慧城市、传媒社交等领域, 通常与人脸识别技术结合使用, 组成一个完整的系统部署到真实场景中。随着信息技术的推广, 电子安全问题受到越来越多的关注, 作为身份授权步骤的人脸识别的前置环节, 活体检测扮演着重要角色。本文针对近年来提出的人脸反欺诈活体检测技术进行了分类及归纳, 通过介绍、分析这些活体检测算法来进一步拓展活体检测的研究思路, 最后给出其未来可能的发展趋势及总结。

**关键词** 人脸识别; 活体检测; 信息安全

**中图法分类号** TP391.41 **DOI 号** 10.19363/J.cnki.cn10-1380/tn.2020.02.02

# A Survey of Face Anti-Spoofing

LU Ziqian[1], LU Zheming[1], SHEN Fengli[1], WANG Zonghui[2]

[1] School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China

[2] Institute of computer system architecture and network security, Zhejiang University, Hangzhou 310027, China

**Abstract** Face anti-spoofing technology is widely used in fields such as information security, financial services, sports entertainment, smart cities, and social media. It is usually deployed to real scene with face recognition system. With the development of the information technology, electronic security has attracted more and more attention. As the pre-stage of face recognition for identity authorization, face anti-spoofing technology plays in an important role of it. This paper summarizes the face anti-spoofing technology proposed in recent years and expands the orientation in this area. Finally, we give the possible development trends and summary of the face anti-spoofing technology.

**Key words** face recognition; living body detection; information security

## 1 引言

当今社会, 作为比较容易获取的生物特征之一, 人脸特征普遍被作为个人身份的认证信息。人脸特征的提取及应用技术早在 20 世纪 90 年代就成为了研究领域的热门, 早期的人脸识别技术一般采用手动提取模式特征的方法, 通过假设某种分布去直接获取低维度的表征, 如线性子空间, 稀疏表示等。2014 年, DeepFace[1]和 DeepID[2]在人脸识别数据集 LFW[3]上获得了当年的最好的成绩, 这也是人脸识别技术首次在无约束场景下超越人类眼睛的识别能力。从这以后, 研究者们就开始将研究目光转向了基于深度神经网络的方法, 随着大规模人脸数据集及 GPU 加速技术的普及, 研究人员通过不断地创新网络结构及损失函数, 不断地加深神经网络深度, 一次次刷新了人脸识别的各项记录, 从而推动了人脸识别技术在产业界的实际落地, 比如人们所熟知的金融支付、身份验证、网络购票、智能家居等。

然而在给与人们生活带来便利的同时, 人脸识别技术自身也存在着诸多的问题, 如隐私安全问题、身份有效性等问题, 而人脸特征信息作为一种极易于从生活场景(上班打卡、金融支付、账户验证等)采集获取的生物特征也很容易遭到攻击, 给个人及社会造成极大的财产损失, 因此验证人脸信息是否真实可靠对于人脸识别在现实生活中的应用十分重要。

其中, 人脸欺诈攻击是常见的针对人脸真实性的攻击手段, 因此在应用中, 完整可用的人脸识别系统必须具有反欺诈攻击的能力。反人脸欺诈检测(Face Anti-spoofing)又被称为活体检测, 是一种判断

**通讯作者:** 陆哲明, 博士, 教授, Email: zheminglu@zju.edu.cn。

本课题得到中央网络信息办公室资金资助。

收稿日期: 2019-12-30; 修改日期: 2020-2-19; 定稿日期: 2020-3-11

人脸是一张真实的实时拍摄的人脸还是伪造的人脸(如戴着人皮面具的人、电子设备屏幕中的人脸数字图像、彩色打印的人脸图片等)的技术。人脸识别技术识别的是人的身份, 那么活体检测就是用来识别是否是真人。为了防止一系列的欺诈攻击, 活体检测技术一般会与人脸识别模块相结合, 作为其前置验证模块, 验证是否是真实的人, 从而形成一套完整的真人身份验证系统, 运用在安防等领域[4]。

本文分析了近年来的人脸反欺诈活体检测算法, 对其做出归纳总结, 并且在最后提出未来可能的发展趋势及方向。主要结构如下: 第 2 章介绍了人脸识别及人脸反欺诈活体检测任务的相关知识; 第 3 章简要介绍了常用的数据集; 第 4 章介绍了传统方法下的人脸反欺诈活体检测算法以及基于深度学习的人脸反欺诈活体检测算法; 第 5 章总结了前几章的内容并对反欺诈活体检测技术的未来做出展望。

## 2 背景知识

### 2.1 人脸识别技术

相对于难提取的指纹及虹膜信息, 人脸特征信息作为一种更加常见稳定的人体生物特征, 它在当今社会扮演着重要的角色, 由于人脸在现实世界中有着高度的可变性, 自然人脸识别也一直是计算机视觉领域研究的热门。人脸识别算法始于 20 世纪 70 年代, 经过长期的研究, 它们的准确度在不断地提升, 其主要识别过程如图 1 所示。

人脸识别技术这些年已经从萌芽趋于成熟。传统机器学习方法依赖于纹理及边缘等人工设计的特征与特征分析技术(比如主成分分析、线性判别分析或支持向量机)的组合。人工设计特征识别的方法在无约束环境中很难保持检测的稳定性, 这使得过去的研究者侧重研究针对每种变化类型的专用方法, 比如能分别应对不同光照、姿势、年龄的方法等。


**图 1 人脸识别流程图**

**Figure 1 Face recognition flowchart**

近段时间, 得益于出色的特征提取能力及大规模数据的支持, 传统的人脸识别方法已经被基于卷积神经网络(Convolutional Neural Networks, CNN)的深度学习方法所代替。其中, 卷积神经网络是人脸识别方面最常用的一类深度学习方法。Facebook 的 DeepFace[1]作为最早的用于人脸识别的卷积神经网络方法之一, 在 LFW 数据集上达到了 97.35% 的准确度, 相比之前的最好结果将错误率降低了 27%, DeepFace[1]提出了一个全新的高效 3D 人脸对齐模型, 该模型采用含有局部连接层的卷积神经网络架构对包含 440 万张人脸的数据集进行训练从而学习到人脸信息的内在特征。同时, 在 ImageNet 大规模视觉识别挑战赛(ILSVRC)中表现优异的架构也充分被研究人员借鉴, 部分研究者将 VGG[5]及 GoogleNet[6]的结构用于人脸识别任务, 在取得不错成绩的同时减少了模型参数量, 使得模型更加轻量化, 可以更方便地部署在移动设备中。

近些时间, Resnet[7]系列网络结构已经成为各类分类、识别任务中模型的热门选择。何凯明等人提出的这种带有跨层连接的网络模型可以提取更加丰富的图片特征信息, 使研究人员可以设计更加深层的网络的同时避免反向传播中的梯度消失, 从而保证了网络训练的稳定性。此外一些新的损失函数如 Triplet loss[8]、Center loss[9]也进一步提升了算法的效果。随着人脸识别技术趋于成熟, 现在研究人员开始关注更加复杂, 具有挑战性的场景, 比如多人数、多尺度、有严重遮挡的人脸识别任务。

### 2.2 人脸活体检测技术

人脸活体检测技术是一种判断捕捉到的人脸是真实的活体人脸, 还是进行处理后的人脸图片(如彩色纸张打印的人脸图, 电子设备屏幕中的人脸数字图像以及化妆、戴面具的人脸图片等)的技术, 通常情况下活体检测与人脸识别模块配合使用, 即存在“先检测, 后识别”的模式, 二者结合的统一系统被广泛应用于生活中的各个领域。

活体检测中常见的攻击方法有: 照片攻击、部分照片攻击、纸片覆盖攻击、视频回放攻击、立体面具攻击等, 因此人脸活体检测任务可以视作为二分类或者多分类任务。针对多变的攻击方式, 目前主流的检测方法主要有: 语音校验、远程心率(remote PhotoPlethysmoGraphy, rPPG)检测、颜色纹理分析、用户配合动作检测、光流法检测等, 还有一些借助外部设备, 如红外摄像头、深度摄像头的方法。此外, 由于深度学习在特征提取上的优越性, 还出现了大量基于深度学习的活体检测算法, 因此也可以将活体人脸检测算法分为传统机器学习流派,以及基于深度学习的检测流派。根据应用场景进行分类, 还可以将检测算法分为固定端活体检测算法和移动端活体检测算法。本文主要从基于传统机器学习算法和基于深度学习算法的角度对部分静默活体检测(不需要被检测人做额外动作的算法)算法进行分析和比较,

图 2 给出了几种常见的在人脸活体检测中出现的攻击方法。

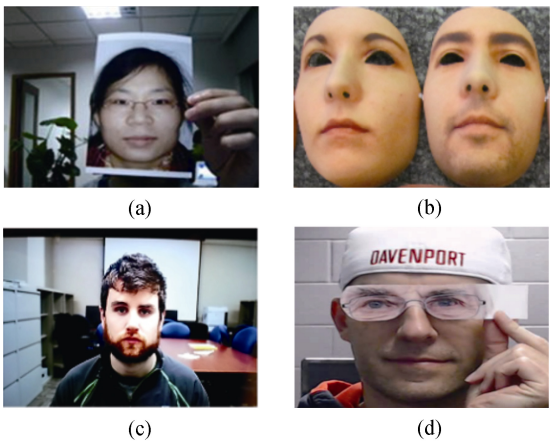



**图 2 几种常见的人脸攻击。(a)照片攻击(b) 3D 面具攻击(c) 视频回放攻击(d)纸片覆盖攻击**

**Figure 2 Several common face attacks (a) photo attacks (b) 3D mask attacks (c) video replay attacks (d) paper overlay attacks**

传统的方法一般是从图片中提取人工设计的特征, 然后再通过支持向量机(Support Vector Machine, SVM)将活体检测转变成一个二分类问题。主要使用的特征有局部二值模式(Local Binary Patterns, LBP)[10-12], 尺度不变特征变换(Scale-invariant feature transform , SIFT)[13], SURF(Speeded-Up Robust Features)[14], HOG(histogram of oriented gradients)[15]等特征。

而基于卷积神经网络的方法则一般以不同结构不同特点的卷积网络组合而成, 值得注意的是按照使用的数据类型我们可以把检测算法分为三类, 只使用单帧颜色图片的单帧检测方法[14, 16], 使用多帧颜色图片和深度图片的双模态检测算法[17-18], 以及同时使用三种或三种模态以上的图片作为算法输入的多模态检测算法[19]。

## 3 常用数据集

### 3.1 NUAA

NUAA[20]是南京航空航天大学模式识别和神经计算小组(ParNeC-NUAA)公布的用于人脸反欺诈活体检测的数据集。该数据库包含 15 个人的真实和欺诈攻击的静态图像。样本是从普通的网络摄像头(未指定型号)捕获的视频中提取的, 分辨率为 640×480 像素。该数据库是在不控制光照条件下, 在不连续的两个星期的三个不同时段中获得的。数据集的 12614 张图片中包括真实图像 5105 张以及欺诈攻击图片 7509 张, 并且不是所有受试者的图片都包括这三个时间段, 图片的数据量也不是均衡的。对于每种攻击主题, NUAA 使用网络摄像头捕获了一系列的面部图像(帧速率为 20fps, 每个主题为 500 幅图像)。在图像捕获过程中, 要求每个对象正视摄像头, 并保持普通的表情, 没有明显的动作, 例如眨眼或头部移动。换句话说, NUAA 尝试使真实的人看起来像一张照片(反之亦然)。NUAA 作为早期的图片攻击数据集为该领域的研究提供了大量的数据支撑。

### 3.2 IDIAP: The Replay-Attack Database

用于面部欺诈的视频重放数据集 IDIAP [21]包含 1300 个剪辑视频, 视频内容包含了 50 个人在不同光照场景中的人脸。具体地, 它包含了通过 320×240 分辨率的网络摄像头获取的 50 个不同个体的真实和欺骗攻击的短片(mov 格式约 10 秒)。该数据集考虑了对三种不同分辨率的攻击: (1)照片打印攻击, 对高分辨率数码照片的硬拷贝进行非法访问攻击; (2)移动攻击, 对 iPhone 4 屏幕拍摄的照片和视频进行移动攻击; (3)高清晰度攻击, 使用分辨率为 1024×768 像素的 iPad 屏幕显示照片和视频。

### 3.3 IDIAP: 3DMAD

3D MASK-ATTACK[22]数据集是 IDIAP 研究所提出的人脸反欺诈数据集, 可在 IDIAP 研究所网站上公开获得。它是第一个考虑面具攻击的公共数据库, 该数据集除了提供 2D 图片数据外, 还提供对应的深度图片信息。它包含 17 个不同用户的真实和攻击数据。面具攻击是由一名操作人员戴着真实人物个人的正面图片和两张其他方向的照片生成的人脸数据, 其中, 数据由两个相距两周的真实场景下采集以及第三次执行面具攻击的采集组成。在每个会话采集中, 对于每个用户, 使用 Microsoft Kinect for Xbox 360 拍摄五个 10 秒的视频。此传感器除了提供常规 2D RGB 数据(8 位), 还提供图片的深度数据(11 位), 分辨率为 640×480 像素, 每秒 30 帧。因此, 可用的数据为: 255 个 300 帧的彩色视频(170 个真实视频和 85 个面具攻击的视频), 以及具有相应深度信息的 2.5D 视频。这种数据的多样性为直接攻击的面部安全性领域的研究提供了极大的灵活性, 因为它使研究 2D 和 3D 反欺骗算法及其融合成为可能。

### 3.4 CASIA FAS DB

CASIA Face Anti-Spoofing DB [23]可从中国科学院(CASIA)生物识别和安全研究中心(CASIA-CBSR)获得。该数据库包含两种视频的短片(avi 格式大约 10 秒)。真实和欺诈的视频共有 50 位对象, 这些对象被分为训练集和测试集, 它们之间没有重叠(就用户

和样本而言)。短片是使用三种分辨率不同的设备采集的: (1)低分辨率, 使用旧的640×480 USB网络摄像头(未指定型号); (2)正常分辨率, 使用现代 480×640 USB 网络摄像头(未指定型号); (3)高分辨率, 使用1920×1080 Sony NEX-5 高清晰度摄像机。考虑了三种不同的攻击方式: (1)弯曲, 对真正用户的高分辨率数码照片在质量比常规A4打印纸更高的铜纸上使用略微弯曲的硬拷贝进行扭曲的非法访问尝试; (2)裁剪, 切掉真实用户的高分辨率数码照片的纸质副本(如前所述)中眼睛部分, 攻击者的脸放在照片后面(即, 眨眼是真实的);(3)数字视频, 使用 iPad 播放高分辨率视频。

### 3.5 MSU MFSD

MSU MFSD[24]数据集。该数据集共包含 55 个实验对象的 440 个视频(平均 12 秒)。通过两种摄像头采集这些视频, MacBook Air13 自带摄像头和 Google Nexus5 手机摄像头, 分辨率分别为 640×480 和 720×480。攻击手段一共有三种: (1)打印的照片攻击, 在 A3 纸上打印分辨为 5184×3456 的受试者图片; (2)平板电脑回放攻击, 通过 Canon 550D Single-lens reflex 摄影机拍摄受试者视频, 然后通过 iPad Air 播放视频, 分辨率为 1920×1088; (3)移动手机回放攻击, 通过 iPhone 5S 拍摄并播放 1920×1080 的视频。与其他数据集相比, 该数据集涵盖了真实和攻击的移动手机拍摄的视频, 可以用来模仿测试手机的解锁应用。同时该数据集的打印攻击采用了比其他实验数据集质量更好的, 更加大的打印纸。表 1 对比了前面介绍的 5 种常用活体检测数据集的参数。

**表 1 常用的数据集信息对比**

**Table 1 Comparison of commonly used datasets**

| 数据集 | 基本对比信息 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 整体信息 | | | 传感器信息 | | | | 光照 | |
| | 对象数 | 真实/伪造 | 种类 | 低 | 中 | 高 | 3D | 固定 | 可变 |
| NUAA | 15 | 5105/7509 | 图片 | | ✓ | | | | ✓ |
| Replay-Attack | 50 | 200/1000 | 视频 | ✓ | | | | ✓ | ✓ |
| 3DMAD | 17 | 170/85 | 视频 | | ✓ | | ✓ | ✓ | |
| CASIA FAS | 50 | 150/450 | 视频 | ✓ | ✓ | ✓ | | | ✓ |
| MSU MFSD | 55 | 440 | 视频 | ✓ | ✓ | | | ✓ | |

## 4 人脸活体检测算法

早期的人脸活体检测算法主要从人工设计的特征层面出发[10-11, 25-27], 这些算法目标很明确, 就是找到活体与非活体的之间的差异性, 然后根据这些差异来设计特征, 最后将提取的特征送到分类器去判断待测特征是否为活体特征, 后期随着卷积神经网络, 循环神经网络等特征提取器的日益强大, 主流的方法变成了通过设计一些特定的网络结构从图片中直接提取可识别信息, 然后利用训练好的模型直接区分活体与非活体[14, 16, 28], 我们将从以上这两个角度分别分析算法。

### 4.1 基于传统机器学习的算法

2014 年, Bharadwaj 等人[29]提出先对输入的多帧图像通过运动放大来增强脸部微动作, 然后, 使用两种特征提取算法(局部二进制模式和使用方向光流法直方图的运动估计)分别对纹理和运动属性进行编码, 最后对这两个正交的融合之后的特征利用 SVM 进行分类, 多种特征融合使得该算法具备有良好的鲁棒性, 结构图如图 3 所示, 输出结果为预测为活体或非活体的得分。此算法在上文所提及的 CASIA FAS 以及 Replay-Attack 数据集上进行了实验, 取得了不错的结果。此文的主要贡献有: (1)提出了可增强人的面部运动的基于运动放大率的预处理算法; (2)提出了一种多尺度 LBP 设定(称为 multiLBP), 该设定对视频的纹理进行编码, 并将其作为活体和非活体的分类特征; (3)提出了一种基于运动估计的新型活体检测算法, 该算法使用定向光流直方图(HOOF)编码的光流进行运动估计。此外, 从消融实验中我们可以看出, 多尺度 LBP 模块的效果要优于定向光流直方图模块的检测效果, 在同样经过运动增强的预处理下前者在 CASIA FAS 测试集上取得了 15.74%的等错误率 EER(Equal Error Rate), 而后者为 21.11%, 在两者融合的模型中 EER 为 14.44%, 从而证明了融合模型的有效性。同时, 相比于 2012 年的基础方法 A face antispoofing database with diverse attacks[30], 在视频攻击上 EER 降低了 4%, 在弯曲图片攻击上 EER 降低了 6%。

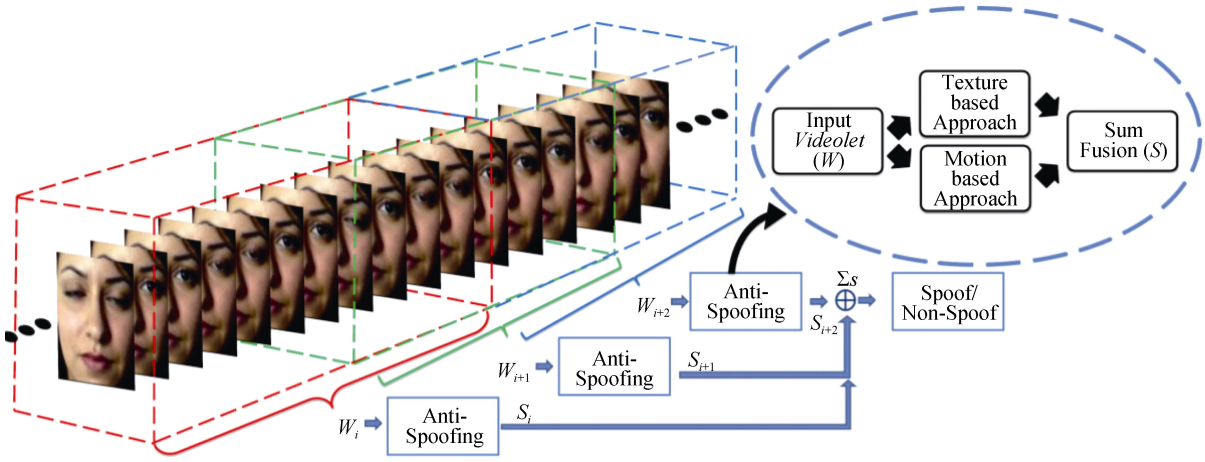



**图 3　多尺度纹理及运动特征融合结构示意图**

**Figure 3　Multi-scale texture and motion feature fusion structure diagram**

2016 年的 Boulkenafet 等人[31]提出通过颜色纹理分析的方法检测图片是否为真实的人脸图像。具体地, 从不同颜色空间中提取互补的低级特征描述, 把不同亮度和色度通道中的联合颜色纹理特征作为待检测特征进行检测分类, 本文在 CASIA FAS, Replay-Attack 以及 MSU MFSD 数据集上进行了实验分析, 在取得了当年最好结果的同时也保证了算法的稳定性。该文献认为在 RGB 空间中很难区别活体与非活体, 但在其他颜色空间里两者的纹理有明显差异, 提取 HSV 空间中的人脸多级 LBP 特征以及 YCbCr 空间的人脸 LPQ 特征线索进行 SVM 分类, 这样做极大地简化了模型复杂度, 同时使其可以轻松地部署到移动端。此文的主要贡献是: (1)不同于以往的基于灰度图像活体监测的算法, 第一次提出了使用颜色纹理分析的方法, 并证明了颜色分量在活体检测任务中的有效性; (2)分析了不同的颜色空间和描述可用于描述真实面孔和伪造面孔之间颜色纹理的内在差异的程度, 并进行了多空间的融合; (3)实验证明, 人脸颜色纹理表征具有广阔的泛化能力, 因此与灰度对应的颜色纹理相比, 颜色纹理在未知条件下可以更稳定。

以往的基于灰度图的检测算法都是利用 LBP 等手动设计的特征, 不同二进制模式的出现被收集到直方图中用以表示图像纹理信息, 对于图像中的每个像素, 二进制代码是通过圆形对称像素与中央像素的值来计算:

$$LBP_{P,R(x,y)}=\sum_{n=1}^{P}\delta\left(r_n-r_c\right)\times 2^{n-1} \tag{1}$$

其中, $r_c$ 和 $r_n$ $(1,\cdots,P)$分别表示中心像素$(x, y)$及其位于半径 $R(R>0)$的圆上的 $P$ 个邻域像素的强度值。

如图 4 所示, 尽管这种细微的差异在 RGB 图像或者灰度图像中是肉眼难以察觉的, 但在其他空间中活体与非活体的差异非常明显, 因此 Colour Texture 方法选择在容易判断这种差异性的颜色空间进行统计分析, 并利用这种差异性进行分类。此方法在 CASIA FAS 数据集的视频结果分析中取得了 3.2% EER 的出色结果, 而传统的灰度运动增强的运动放大(Motion Magnification)方法的 EER 仅为 14.4%。该方法在基于单帧的检测分析中同样有效, EER 为 2.1%, 相比于 2014 年的 Yang 等人[32] 提出的基于卷积神经网络的模型低了 5.3%。

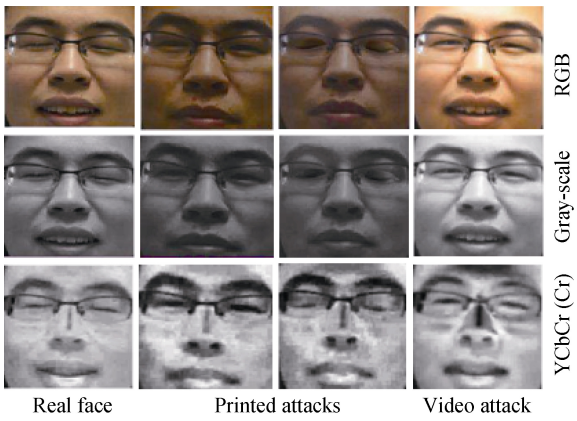



**图 4　不同空间中的活体与非活体图像示例**

**Figure 4　Exampls of Living and non-living images in different spaces**

多空间多特征的分析角度拓宽了研究人员思路, 而不只是局限于对特征的放大及对常见的 LBP 特征的计算, 同样地, 2016 年的 Li 等人 [33]从远程生物特征入手提出了解决欺诈攻击的新思路。文章从视频中的远程心率(rPPG)入手, 对输入进行一个初步判断以区分活体及照片攻击, 然后通过一个级联的纹

理 LBP 特征分类器进一步区分更加相似的视频回放攻击, 在 3DMAD 以及 REAL-F 这两种 3D 面具攻击数据集中取得了出色的效果。

总的来说, 基于传统机器学习的活体检测算法注重纹理特征的设计及对图像和视频中固有属性的利用, 通过多特征融合和辅以其他生物特征作为辅助信息提升算法的性能及鲁棒性。

## 4.2 基于深度学习的活体检测算法

传统的人脸活体检测一般从人工设计的纹理、颜色等特征入手, 提取图片和视频中的统计特征, 使用 SVM 进行分类, 然而这类方法的特征提取的能力非常有限, 分类器很多时候不能很好的区分输入中的细微差别, 尤其是对视频重放攻击, 视频中的人脸和真实人脸十分相似, 即使是远程心率(rPPG)特征也很难将两者区分开, 因此研究者将关注点转向了具有更强特征提取性能的深度学习领域, 利用卷积神经网络、循环神经网络等提取的特征代替原本的手动设计的特征, 进而取得了一系列的成果, 并逐渐成为本领域研究方向的主流。图 5 展示了基于卷积神经网络的人脸活体检测的一般流程。

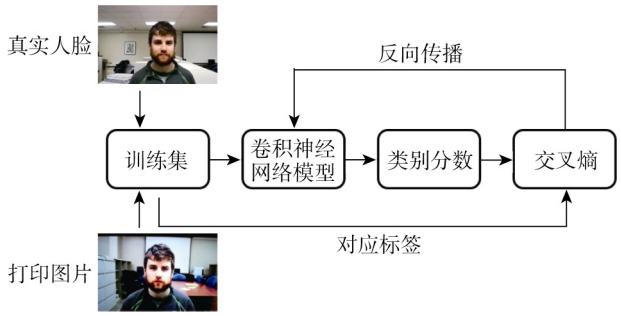



**图 5 基于卷积神经网络模型的人脸活体检测**

**Figure 5 Face anti-spoofing based on convolutional neural network mode**

2017 年, Atoum 等人[34]首次考虑把人脸深度图作为活体与非活体的差异特征作为判别依据, 他们提出的模型通过卷积神经网络提取图片特征, 最后对融合后的特征做分类。该文献创新点主要是提出了两种卷积网络用于人脸活体检测, 一个用于提取局部的人脸特征, 另一个利用深度信息来判断图片是否具有真实的人脸应有深度信息, 在这个模型中前者用于判断局部人脸特征是否具有全局图片中的空间信息, 后者用来判断待测深度信息中是否具有相匹配的深度信息。区别于之前的深度学习方法, 该方法提取的是像素级别的特征, 得到的是整张图片的得分图, 而不是每张图片提取出一个特征表示。实验结果充分证明了基于卷积神经网络的特征提取能力的优越性, 在 CASIA FAS 数据集上本方法的 CNN 特征以及全卷积(FCN)特征分别取得了 4.44% 和 2.85% 的 EER, 融合之后的结果为 2.67% 更是远远优于 2016 年的 VGG 分类方法, 此后越来越多的高性能的可插拔特征提取网络也被运用到活体检测任务中。图 6 展示了全卷积网络提取出来的深度信息的得分, 其中黄色代表为真实人脸的可能性大, 蓝色代表为伪造人脸的可能性大, 左边两列代表真实人脸, 右边两列代表照片打印人脸, 容易看出真实人脸的大部分得分都趋向于黄色, 而照片中的非活体人脸则相反。

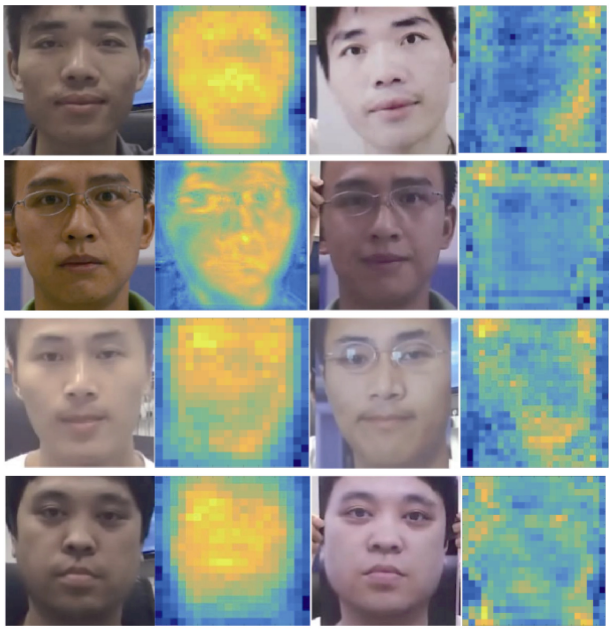


**图 6 全卷积网络分类可视化示例**

**Figure 6 Example of full convolutional network classification visualization**

尽管 CNN、FCN 等网络结构有着强大的特征提取能力, 可是在复杂多变的真实场景中, 往往面临数据量欠缺和过拟合等问题, 因此在很长一段时间内基于深度学习框架的检测算法仍然与传统算法不相上下, 而 2018 年的 Liu 等人[18]提出的利用空间和时间辅助信息的方法一举超过了传统算法。他们利用空间和时间信息作为监督的辅助信息, 这些辅助信息包括两个层面: 空间和时间, 其中空间就是图像的深度信息, 而时间信息则是使用远程心率(rPPG)特征, 两者都是通过端到端的“卷积—循环”(CNN-RNN)网络学习获得的, 此外, 针对于原有的数据集质量不高的现象, 此文使用更高精度的设备公布了一个新的数据集 Spoof in the Wild Database (SiW), 此数据集在增加了样本数量的同时还增加了不同视角的素材为以后的

研究提供了新的可用数据。

相比于原有的远程心率(rPPG)配合 LBP 特征简单分类的方法, CNN 提取出的可辨别的深度特征以及 RNN 编码的远程心率脉冲信号更加具有区分度, 如图 7 所示, 每一个时刻的 RNN 网络和一个 CNN 网络一一对应, 形成端到端的分类模型, 从结果表现方面来看, 在常用数据集上 CNN-RNN 方法皆超过了传统算法中表现最好的 Color Texture 方法。

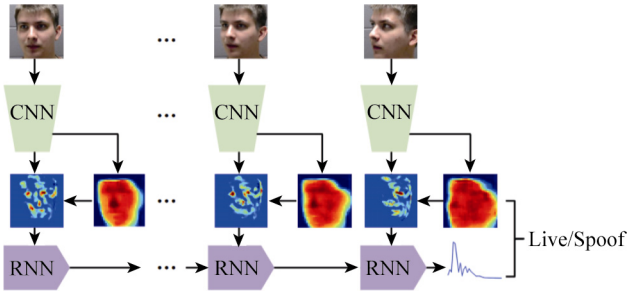



**图 7 卷积神经网络及循环神经网络级联方式**

**Figure 7 Convolutional neural network and recurrent neural network cascade method**

前面所介绍的检测方法都有相似的流程, 即第一步特征提取, 第二步特征分类, 和这些检测方法不同, 2019 年的 Muhammad 等人[35]提出利用生成方法产生数据的思路, 并以此来增加训练样本的数量, 从而提高模型的表现。这篇文献通过生成的方法来增加和提高由 Resnet 提取的图片特征的数量和质量, 然后通过 LSTM 再次增加训练样本的数量, 从而产生更好的特征。该模型在 CASIA 数据集合和 MSU 数据集下分别达到了 0.01 EER 和 0.02 EER, 在大量高质量数据的驱动下模型表现有了近百倍的提升, 同时也证明了生成算法在检测算法中可以起到不错的辅助作用。

同一年, Nikitin 等人[36] 也同样采用了生成式的方法来提高模型的表现。但与 Muhammad 等人合成特征的方法不同, 他们通过合成“非活体”的整体图片, 增加了负样本的数量。此方法理论上可以做到成倍的增加训练集, 所以该方法不需要考虑过拟合的问题, 合成的结果如图 8 所示。同时, 为了解决打印照片的攻击, 本文同时还结合了眨眼检测的方法, 用以辅助活体检测。该方法在 CASIA 数据集上也达到了 0.01 EER 的水平。进一步证实了生成算法对于检测算法的增强作用。

2019 年, George 等人 [37]通过增加人脸数据的获取手段来提高活体检测的正确率。该文献提出了一种多模态的 CNN 模型, 利用颜色, 深度, 红外和热像等多种图片作为 CNN 的输入进行活体检测, 并提

**图 8 生成的伪造图片实例**

**Figure 8 Examples of generated fake face images**

出了包含多通道图片的 Wide Multi-Channel presentation Attack (WMCA) 数据集, 四种类型的数据可以参考图 9。本文通过实验证明了, 仅仅只使用颜色信息进行活体检测得到的结果远远不如使用多重通道得到的结果, 同时使用颜色, 深度, 红外和热成像四个通道的数据进行训练时, 模型达到最好的结果。由于很多的攻击方法对图片的攻击只在某种通道中表现, 所以使用了多通道数据的本文模型在检测未出现在训练集中的攻击也有不错的效果。

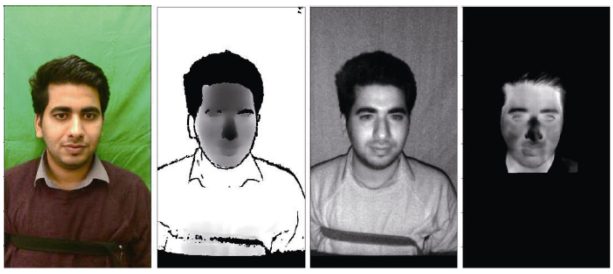



**图 9 WMCA 数据集实例(a)颜色图片(b)深度图片(c)红外图片(d)热像图片**

**Figure 9 Examples of WMCA dataset. (a) Color image (b) Depth image (c) Infrared image (d) Thermal image**

随着越来越多该领域研究的发表, 人体活体检测领域数据集的完备以及其实用的推广, 2019 年 CVPR 举办的人脸活体检测的挑战赛, 堪称一次活体检测领域的 ImageNet, 此次比赛的数据集是 CASIA-SURF, 这是中科院自动化所推出的数据集, 包含了 1000 个个体样本的 21000 段视频。采集设备是英特尔的 RealSense SR300 立体相机, 同时采集了 RGB 图, 红外图和深度图。比赛采用 TPR@10e-4 FPR 即当假正比率(FPR)为 10e-4 时候的真正比率

(TPR)的数值作为衡量指标。

最终 Top3 的队伍都取得了十分卓越的成绩, 来自俄罗斯的 VisionLabs 冠军代表队更是取得了99.88%的 TPR, 他们的 Multi-modal Face Spoofing 从网络结构入手, 充分利用了在分类任务中有效的 SE[38]模块, 该体系结构利用了多模态的图像数据并在多个网络层上聚合了通道内特征, 同时作者受人脸识别方法的启发, 把人脸识别网络的预训练数据运用在活体检测任务中, 并融合了 res-18, res-34 以及 res-50 等不同结构的特征。可以看出, 该文献从多模态多通道特征入手, 利用简单的 SE 模块增加通道间信息, 最终对池化后的特征做 softmax 分类。简单的模式也从侧面证明了当前对人脸活体检测领域的研究与已经取得出色成果的人脸识别具有相通性, 大量的数据支持及特征提取是解决问题的关键。

此外在算法日益强大过程中, 研究人员也思考如何能在保证精度的同时尽可能地减少模型复杂度以及参数量, 这样基于深度学习的人脸活体检测算法才更具有实用性也更容易与轻便的传统算法作比较, 挑战赛第三名的 FeatherNets [39]给出了轻量化的解决方案, 该文献最主要的创新点就是提出了流模块(Streaming Model)来代替全局平均池化层, 在人脸任务中我们的关注点主要在人脸中心部分, 因此各区域均衡权值的全局平均池化层(GAP)并不适用于人脸任务而是比较适合物体检测问题, 因此作者选用深度卷积层来解决问题, 在提升了精度的同时也减少了参数量, 其中所提出的 FeatherNet 在 i7-CPU 上推理时间仅为 1.87ms, 参数量也仅为 0.35M 是传统模型的一半甚至是几十分之一, 同时保证了 99.76%的 TPR@10e-3 FPR。

基于深度学习的活体检测方法注重对数据的信息利用, 通常融合了图像的 RGB, 深度以及红外[40]等有用信息, 通过训练高性能的特征提取网络来区分活体与非活体, 同时相比于传统的算法, 深度学习方法在提高性能的同时增加了许多参数量这使得计算复杂度以及成本也成为了落地过程中需要考虑的对象。

## 5 总结与展望

从一开始的通过简单的纹理特征放大到多特征多颜色空间的融合分析再到多模态的卷积神经网络特征提取, 人脸活体检测任务始终关注活体人脸与非活体攻击之间的差异性, 尽管提取角度与思路不同, 究其根本, 提取出具有辨识度的可分特征是解决问题的关键, 同时在深度学习特征提取能力日益强大的背景下利用相似任务先验知识配合多特征融合分类成为了活体检测的有效的策略, 传统与深度, 单帧与多帧, 灰度与色彩的结合也成为了更多研究者的选择。

着眼未来, 人脸活体检测任务大多还处于二分类任务的模式下, 随着大规模高质量数据集的建立, 多分类, 多人场景下的活体检测任务会成为研究人员关注的热点。同时, 在深度学习框架下, 如何更加高效实时地处理也至关重要。此外, 在监督学习取得成功之后, 如何提升模型的泛化能力使其适用于更加符合真实场景的半监督, 无监督学习也带来了新的机遇与挑战。

## 参考文献


[1] O.M. Parkhi, A. Vedaldi , A. Zisserman, Deep face recognition[C]. *British Machine Vision Conference (BMVC),* 2015: 1891-1898.

[2] Y. Sun, X. Wang , X. Tang, Deep learning face representation from predicting 10,000 classes[C]. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2014: 1891-1898.

[3] G.B. Huang, M. Mattar, T. Berg , E. Learned-Miller. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, 2007.

[4] N. Chen, A Survey on Face Anti-Spoofing[J]. *China Computer and Communication*, 2019, 426(8):114-115+118.
(陈宁珏,人脸活体检测综述[J]. *信息与电脑*, 2019, 426(8): 114-115+118。)

[5] K. Simonyan , A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. *Computer Science*, 2014, 34(2):98-102.

[6] C. Szegedy, W. Liu, Y. Jia, et al. Going deeper with convolutions[C]. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2015: 1-9.

[7] He K M, Zhang X Y, Ren S Q, et al. Deep Residual Learning for Image Recognition[C]. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 27-30, 2016. Las Vegas, NV, USA. Piscataway, NJ: IEEE, 2016: 770-778.

[8] Schroff F, Kalenichenko D, Philbin J. FaceNet: A Unified Embedding for Face Recognition and Clustering[C]. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 7-12, 2015. Boston, MA, USA. Piscataway, NJ: IEEE, 2015: 815-823.

[9] Wen Y D, Zhang K P, Li Z F, et al. A Discriminative Feature Learning Approach for Deep Face Recognition[M]. Computer Vision – ECCV 2016. Cham: Springer International Publishing, 2016: 499-515.

[10] Maatta J, Hadid A, Pietikainen M. Face Spoofing Detection from Single Images Using Micro-texture Analysis[C]. *2011 International Joint Conference on Biometrics (IJCB)*, October 11-13, 2011. Washington, DC, USA. Piscataway, NJ: IEEE, 2011: 1-7.

[11] de Freitas Pereira T, Anjos A, de Martino J M, et al. Can Face

Anti-spoofing Countermeasures Work in a Real World Scenario?[C]. *2013 International Conference on Biometrics (ICB)*, June 4-7, 2013. Madrid, Spain. Piscataway, NJ: IEEE, 2013: 1-8.

[12] de Freitas Pereira T, Anjos A, de Martino J M, et al. LBP–TOP Based Countermeasure Against Face Spoofing Attacks[M]. Computer Vision-ACCV 2012 Workshops. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013: 121-132.

[13] Patel K, Han H, Jain A K. Secure Face Unlock: Spoof Detection on Smartphones[J]. *IEEE Transactions on Information Forensics and Security*, 2016, 11(10): 2268-2283.

[14] Li L, Feng X Y, Boulkenafet Z, et al. An Original Face Anti-spoofing Approach Using Partial Convolutional Neural Network[C]. *2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA)*, December 12-15, 2016. Oulu, Finland. Piscataway, NJ: IEEE, 2016: 1-6.

[15] J. Yang, Z. Lei, S. Liao and S.Z. Li, Face liveness detection with component dependent descriptor, *2013 International Conference on Biometrics (ICB),* 2013: 1-6.

[16] Yang J W, Lei Z, Liao S C, et al. Face Liveness Detection with Component Dependent Descriptor[C]. *2013 International Conference on Biometrics (ICB)*, June 4-7, 2013. Madrid, Spain. Piscataway, NJ: IEEE, 2013: 1-6.

[17] Patel K, Han H, Jain A K. Cross-Database Face Antispoofing with Robust Feature Representation[M]. Biometric Recognition. Cham: Springer International Publishing, 2016: 611-619.

[18] Hernandez-Ortega J, Fierrez J, Morales A, et al. Time Analysis of Pulse-Based Face Anti-Spoofing in Visible and NIR[C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 18-22, 2018. Salt Lake City, UT. Piscataway, NJ: IEEE, 2018: 544-552.

[19] Liu Y J, Jourabloo A, Liu X M. Learning Deep Models for Face Anti-Spoofing: Binary or Auxiliary Supervision[C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018. Salt Lake City, UT, USA. Piscataway, NJ: IEEE, 2018: 389-398.

[20] Zhang S F, Wang X B, Liu A J, et al. A Dataset and Benchmark for Large-Scale Multi-Modal Face Anti-Spoofing[C]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 15-20, 2019. Long Beach, CA, USA. Piscataway, NJ: IEEE, 2019: 919-928.

[21] Anjos A, Chingovska I, Marcel S. Anti-spoofing: Face Databases[M]. Encyclopedia of Biometrics. Boston, MA: Springer US, 2014: 1-13.

[22] I. Chingovska, A. Anjos , S. Marcel, On the effectiveness of local binary patterns in face anti-spoofing[C]. *International Conference Biometrics Special Interest Group (BIOSIG),* 2012: 1-7.

[23] Erdogmus N, Marcel S. Spoofing in 2D Face Recognition with 3D Masks and Anti-spoofing with Kinect[C]. *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, September 29-October 2, 2013. Arlington, VA, USA. Piscataway, NJ: IEEE, 2013: 23-30.

[24] Komulainen J, Hadid A, Pietikäinen M. Face Spoofing Detection Using Dynamic Texture[M]. Computer Vision - ACCV 2012 Workshops. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013: 146-157.

[25] Wen D, Han H, Jain A K. Face Spoof Detection with Image Distortion Analysis[J]. *IEEE Transactions on Information Forensics and Security*, 2015, 10(4): 746-761.

[26] Komulainen J, Hadid A, Pietikainen M. Context Based Face Anti-spoofing[C]. *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, September 29-October 2, 2013. Arlington, VA, USA. Piscataway, NJ: IEEE, 2013: 1-8.

[27] Boulkenafet Z, Komulainen J, Hadid A. Face Anti-Spoofing Using Speeded-Up Robust Features and Fisher Vector Encoding[J]. *IEEE Signal Processing Letters*, 2016, 24(2):141-145.

[28] Marcel S, Nixon M S, Li S Z. Handbook of Biometric Anti-Spoofing[M]. London: Springer London, 2014.

[29] Feng L T, Po L M, Li Y M, et al. Integration of Image Quality and Motion Cues for Face Anti-spoofing: A Neural Network Approach[J]. *Journal of Visual Communication and Image Representation*, 2016, 38: 451-460.

[30] S. Bharadwaj, T. Dhamecha, M. Vatsa , R. Singh, Face anti-spoofing via motion magnification and multifeature videolet aggregation[C]. *International Conference on Pattern Recognition (ICPR)*, 2014: 1-14.

[31] Zhang Z W, Yan J J, Liu S F, et al. A Face Antispoofing Database with Diverse Attacks[C]. *2012 5th IAPR International Conference on Biometrics (ICB)*, March 29-April 1, 2012. New Delhi, India. Piscataway, NJ: IEEE, 2012: 26-31.

[32] Boulkenafet Z, Komulainen J, Hadid A. Face Spoofing Detection Using Colour Texture Analysis[J]. *IEEE Transactions on Information Forensics and Security*, 2016, 11(8): 1818-1830.

[33] J. Yang, Z. Lei , S.Z. Li, Learn convolutional neural network for face anti-spoofing[OB.IE]. 2014: arXiv preprint arXiv:1408.5601.

[34] Li X B, Komulainen J, Zhao G Y, et al. Generalized Face Anti-spoofing by Detecting Pulse from Face Videos[C]. *2016 23rd International Conference on Pattern Recognition (ICPR)*, December 4-8, 2016. Cancun. Piscataway, NJ: IEEE, 2016: 4244-4299.

[35] Atoum Y, Liu Y J, Jourabloo A, et al. Face Anti-spoofing Using Patch and Depth-based CNNs[C]. *2017 IEEE International Joint Conference on Biometrics (IJCB)*, October 1-4, 2017. Denver, CO. Piscataway, NJ: IEEE, 2017: 319-328.

[36] U. Muhammad, T. Holmberg, W.C. de Melo et al. Face Anti-Spoofing via Sample Learning Based Recurrent Neural Network (RNN)[C]. *The British Machine Vision Conference (BMVC),* 2019: 23-31.

[37] M.Y. Nikitin, V.S. Konushin , A.S. Konushin, Face anti-spoofing with joint spoofing medium detection and eye blinking analysis[J]. *Компьютерная оптика*, 2019, 43(4):23-31.

[38] George A, Mostaani Z, Geissenbuhler D, et al. Biometric Face Presentation Attack Detection with Multi-Channel Convolutional Neural Network[J]. *IEEE Transactions on Information Forensics and Security*, 2020, 15: 42-55.

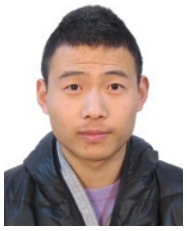
[39] Hu J, Shen L, Sun G. Squeeze-and-Excitation Networks[C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018. Salt Lake City, UT. Piscataway, NJ: IEEE,

2018: 7132-7141.

[40] P. Zhang, F. Zou, Z. Wu, et al. FeatherNets: Convolutional Neural Networks as Light as Feather for Face Anti-spoofing[C]. *The IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW),* 2019:567-571.

[41] S. Bhattacharjee, A. Mohammadi, A. Anjos, et al. Recent advances in face presentation attack detection[M]. Handbook of Biometric Anti-Spoofing, Cham, Switzerland:Springer, 2019: 207-228.

**卢子谦** 于2018年在青岛理工大学通信工程专业获得学士学位。现在浙江大学空天信息技术专业攻读博士学位。研究领域为零样本学习。研究兴趣包括: 深度伪造及检测、迁移学习、生成对抗网络。Email: ziqianlu@zju.edu.cn



**沈冯立** 于2014年在浙江大学飞行器设计与工程专业获得学士学位。现在浙江大学空天信息技术专业攻读博士学位。研究兴趣包括：计算机视觉、物体检测、语义分割。Email：216424013@zju.edu.cn

**陆哲明** 于 2001 年在哈尔滨工业大学测试计量技术及仪器专业获得博士学位。现任浙江大学教授。研究领域为人工智能、信息安全。研究兴趣包括：深度学习及其应用、多媒体信息隐藏。Email：zheminglu@zju.edu.cn

**王总辉** 于2018年在浙江大学计算机科学与技术专业获得博士学位。现任浙江大学计算机系统结构与网络安全研究所高级工程师。研究领域为计算机系统软件和信息安全。研究兴趣包括：系统安全、内容安全、云平台及虚拟化和区块链技术等。Email：zhwang@zju.edu.cn