中图法分类号:TP309; TP399　　**文献标识码**: A　　**文章编号**: 1006-8961(2024)02-0293-02

论文引用格式:卢伟, 钱振兴, 罗向阳, 赵险峰, 李晓龙, 张卫明, 李斌, 王员根, 王金伟, 秦川, 陈秀妍. 2024.《中国图象图形学报》数字媒体深度伪造与对抗专栏简介. 中国图象图形学报, 29(02):0293-0294[DOI:10.11834/jig.2400002]

《中国图象图形学报》数字媒体深度伪造与对抗专栏简介

卢伟[1],钱振兴[2],罗向阳[3],赵险峰[4],李晓龙[5],张卫明[6],李斌[7],王员根[8],王金伟[9],秦川[10],陈秀妍[11*]

1. 中山大学,广州　510006; 2. 复旦大学,上海　200433; 3. 中国人民解放军战略支援部队信息工程大学,郑州　450001; 4. 中国科学院信息工程研究所,北京　100195; 5. 北京交通大学,北京　100044; 6. 中国科学技术大学,合肥　230026; 7. 深圳大学,深圳　518060; 8. 广州大学,广州　510006; 9. 南京信息工程大学,南京　210044; 10. 上海理工大学,上海　200093; 11.《中国图象图形学报》编辑部,北京　100190

当前,以Deepfake为代表的媒体生成模型和以ChatGPT为代表的大语言生成模型导致以深度生成数字图像、视频、音频和文本普及化。以自媒体传播为枝蔓、紧密围绕新型互联网传播方式,加剧了深度伪造在网络空间中的滋生滥用。国际上基于深度伪造各类虚假消息充斥在包括传统主流媒体和活跃的自媒体的现代媒体传播网络中,全方位多角度地展示了伪造媒体信息对隐私、国家安全的威胁。在当前与未来的信息对抗中,网络空间开源情报与内容安全成为信息对抗的新阵地,对网络空间的分析和利用是取得信息对抗主动权的重要环节。

为了增强舆论和开源情报攻防能力,有必要加强数字媒体深度伪造和对抗的研究。《中国图象图形学报》邀请业内专家共同策划推出"数字媒体深度伪造与对抗"专栏,主要收录国内学者在相关领域具有创新性、突破性的研究成果。以期为相关领域的研究人员提供参考。

经过严格评审,"数字媒体深度伪造与对抗"专栏共收录12篇论文,包括:2篇综述、1篇数据集论文和9篇研究论文:

《深度伪造及其取证技术综述》(作者:丁峰,匡仁盛,周越,孙珑,朱小刚,朱国普*)描述了各种针对解决Deepfake相关问题的处理方法。主要参考了谷歌学术检索近5年的深度伪造论文,分为不同类别进行分析比较,并且详细介绍了深度伪造数据集的特点以及伪造方法,简述了深度伪造技术及其基本原理,介绍了近几年检测器在深度伪造技术数据集上的性能效果,分别从输入维度、浅层特征和深层特征对深度伪造检测技术进行分类,并对它们的未来发展前景进行了展望。

《人脸深度伪造主动防御技术综述》(作者:瞿左珉,殷琪林,盛紫琦,吴俊彦,张博林,余尚戎,卢伟*)对当前学术界提出的人脸深度伪造主动防御技术进行梳理、总结和讨论。首先阐述了深度伪造主动防御的提出背景和主要思想,并对现有的人脸深度伪造主动防御算法进行汇总和归类,然后对各类主动防御算法的技术原理、性能、优缺点等进行了系统性的总结,同时介绍了研究常用的数据集和评估方法,最后对深度伪造主动防御所面临的技术挑战进行了分析,对其未来的发展方向展开了思考和讨论。

《面向感知哈希的图像数据集》(作者:周元鼎,房耀东,秦川*)针对感知图像哈希任务,面向实际图像内容认证场景构建了一个新的包含116 400幅图像的数据集。通过大量实验表明,在该数据集上训练得到的模型较为稳定且具有一定的泛化能力,能够应对复杂多样的实际环境。

《联合多重对抗与通道注意力的高安全性图像隐写》(作者:马宾,李坤,徐健*,王春鹏,李健,张立伟)针对现有隐写方法很难抵御基于深度学习的隐写分析器的检测的问题,提出一种基于生成对抗图

* 通信作者

像的提升图像隐写术安全性的新方法。

《利用跨模态信息检索的鲁棒隐蔽通信》(作者:张晏铭,陈可江*,丁锦扬,张卫明,俞能海)研究了现有多媒体数据流隐蔽通信方法的不足,并提出了一种名为RoCC的隐蔽通信方法,具备高隐蔽性、高安全性和强鲁棒性。

《具有超分辨率行为伪装效果的可逆图像隐藏》(作者:贾孟霖,杨杨*,孙冬)从行为安全的角度出发,提出了一种新的基于超分辨率行为伪装的可逆图像隐藏方法,即在隐藏秘密图像的过程中,同时进行超分辨率行为伪装处理,获得具有秘密信息的超分辨率图像,从而转移未授权方的注意力,实现对秘密图像的保护,增强秘密图像的安全性。

《结合Kd-树和熵编码的密文图像可逆数据隐藏》(作者:金丹,徐达文*)提出了一种高容量密文图像可逆数据隐藏算法,该算法在图像加密前腾出空间,对比实验结果显示,能够实现更高的嵌入容量以及对原始图像的无损恢复。

《最小依赖隐藏的屏摄鲁棒水印方法》(作者:宋佳维,刘春晓*,张心怡)在重新设计噪声层的基础上,提出了一种最小依赖载体图像隐藏水印信息的屏摄鲁棒水印,将屏摄水印对于载体图像的依赖控制在最小。

《面向图像拼接检测的自适应残差算法》(作者:张玲,穆文鹏, 陈北京*)以在图像分类领域获得卓越性能的EfficientNet为骨干网络,设计了ARM模块将输入的拼接图像在网络中进行预处理以凸显未篡改区域和篡改区域图像的本质属性差异。论文设计的ARM是一个即插即用的轻量级自适应特征提取模块,可以在其他模型上进行迁移,例如Xception,ResNet等。

《边缘引导的双注意力图像拼接检测网络》(作者:吴晶辉,严彩萍*,李红,刘仁海)提出了一个边缘引导的双注意力图像拼接检测网络模型(BDA-Net),该方法可以对拼接图像的篡改区域实现像素级的定位。

《结合金字塔Transformer与浅层CNN的变电站图像篡改检测》(作者:邢建好,田秀霞*,韩奕)所提的双通道拼接篡改检测模型结合了Transformer和CNN在图像篡改检测方面的优势,提高了模型的检测精度,适用于复杂变电站场景下的篡改目标检测。

《结合图像块比较与残差图估计的人脸伪造检测》(作者:冯才博,刘春晓*,王昱烨,周其当)引入“图像块归属纯净性”和“残差图估计可靠性”的概念,提出了基于图像块比较和残差图估计的人脸伪造检测方法。

我们期待广大读者和科技人员通过“数字媒体深度伪造与对抗”专栏,能够更深入、更全面地了解该领域的最新方法和应用,吸引更多学者从事相关研究并产生具有国际影响力的优秀成果,为本领域的发展做出新的贡献。