一键“变脸”背后暗藏诸多风险

AI换脸技术滥用乱象调查

□ 本报记者 赵 丽
□ 本报实习生 王艺霏

“你是温峥嵘，那我是谁？”

演员温峥嵘本人浏览短视频时，意外发现“自己”正在直播。她在评论区提出质疑后，被拉黑并踢出直播间。而对方仍在用“她的脸”直播带货。

温峥嵘近期遭遇的“冒牌直播”事件，是AI换脸技术滥用的典型写照。多个直播间出现冒用温峥嵘形象的AI换脸带货行为。其团队日均举报50个假冒账号并发出律师函，可造假账号仍屡禁不止，令她陷入“很难证明我是我”的困境。

这一现象并非孤例。《法治日报》记者近日调查发现，AI换脸与声音克隆技术已形成一条涵盖教程传播与定制服务的完整灰色产业链，对个人权益、平台监管等均构成严重威胁。

换脸技术被用于违法犯罪

11月5日，温峥嵘在短视频账号进行3次以“我是真的温峥嵘”为主题的直播，讲述近期被AI伪造、换脸直播带货的经历。

此前，已有多位名人深受AI换脸困扰。主持人李梓萌的AI伪造形象出现在直播间，虚假宣传深海鱼油“能治病”；奥运冠军全红婵的声音被AI克隆，用于售卖土鸡蛋；医生张文宏被换脸，其形象被用于在直播间售卖某款蛋白棒产品，销量已过千。张文宏自述：“这些假的合成信息就像蝗灾一样，一次又一次发生……”

11月14日，中央网信办官方账号发文表示，有网络账号利用AI技术仿冒公众人物形象，在直播、短视频等环节发布营销信息，误导网民，涉嫌虚假宣传和网络侵权，严重破坏网络生态，造成不良影响。网信部门严厉处置一批违法违规网络账号。同时，督促网站平台发布治理公告，举一反三，开展集中清理整治，目前已累计清理相关违规信息8700余条，处置仿冒公众人物账号1.1万余个。

AI换脸技术的滥用甚至延伸至刑事犯罪领域。2024年10月，江苏省南京市玄武区人民法院审理的一起案件中，被告人符某非法获取195万条公民个人信息，利用AI换脸软件生成动态人脸视频，成功突破某金融支付平台的人脸识别认证系统，登录被害人的支付账户，盗刷银行卡进行消费。法院以侵犯公民个人信息罪、信用卡诈骗罪数罪并罚，判处符某有期徒刑4年6个月。

2024年，内蒙古鄂尔多斯一名消费者遭遇AI换脸诈骗，险些损失40余万元。视频通话中，不法分子利用AI换脸技术冒充该消费者朋友，以资金周转为由，诱骗其转账40余万元。好在警方介入及时，未造成实际损失。

相关教程和服务遍布网络

记者调查发现，AI换脸与声音克隆的教程及服务在网络平台流通。记者在各类社交平台以“AI换脸”“声音克隆”等为关键词搜索时，虽然平台会弹出官方提示，提醒用户使用需“合法合规、注意辨别”“切勿轻信、谨防诈骗”，但相关内容仍在网络空间大量存在。

比如，有相当数量以“AI科普”为标签的账号在各平台公开分享换脸教程。这些视频通常详细列出可访问的网站地址，逐步演示操作流程，并附有技术细节说明，如建议用户“尽量保证替换素材与原素材的人物尺寸一致、画面干净无遮挡”，以保证生成效果。

根据教程指引，记者登录相关网站验证发现，部分平台确实免费提供AI换脸或声音制作服务，支付几十元购买会员后，还能解锁去水印、高清放大等进阶功能。

除公开教程外，更为隐蔽的交易正通过网盘渠道进行。

记者观察到，有一些打包出售的“完整教程”价格低至8元，且卖家声称“没有市面主流软件那么多限制”。这类教程通常包括云端工作流地址资料和详细的操作说明，以及图片去AI化、舞蹈动作模仿、人脸精修等扩展功能与相关的提示词教学。

声音克隆类的教程售价低至9.9元，并在宣传中强调“支持任何形象和声音”。记者询问商家了解到，只要购买此安装包，无需注册会员即可免费破解使用对口型、克隆音色等功能。据商家描述，只需提供一段音频文件，系统便可以在无身份认证的情况下进行克隆，单次支持30秒音频，且不限次数。

为验证此类教程的实际应用效果，记者购买了一份标价8元的换脸教程。该教程包含一段5分钟的视频教程和两份文字材料，明确提供了平台访问链接、新用户福利信息和操作步骤，并特别注明“对敏感词限制极少”“生成内容无水印”。

记者发现，该平台提供多种功能模板，内容涵盖替换视频人物、给人物指定部位更换衣物、对图片中人物身体部位进行放大或动态处理等不同功能。每个模板均配备相应的提示词教学和详细操作步骤，用户只需上传参考图片、视频或待处理素材，等待一段时间即可生成内容。

记者实测使用一段时长5秒的视频进行换装操作，约3分钟后成功生成视频。画面中，人物动作与神态还原度很高，被遮挡部位的服装贴合效果也较为自然、精准。

记者注意到，该平台内容生成功能采用运行时长计费模式，每秒收费0.2虚拟币。教程中附有专属邀请码，使用者可凭此免费领取1000虚拟币。平台还设置了虚拟币奖励机制：每天登录赠送100虚拟币，邀请新用户可获500虚拟币。

更换表述可绕过审核机制

电商平台上，有商家为AI换脸服务披上伪装售卖。

在某购物平台，记者直接搜索“AI换脸”，显示无任何结果。但记者观察到，有部分商家会在商品详情页用表情符号隐喻其能提供换脸“服务”；也有商家以“啥都能做”“不限视频内容，懂的都懂”为宣传语，暗示私聊客服了解商品详情。

记者了解到，10秒换脸视频报价从15元至300元不等。例如，有商家提出，20秒内视频（包括仅生成音频的）统一收费80元；部分商家还推出298元200分钟生成时长、490元年卡等套餐，承诺“一对一教学指导”，宣传语中不乏“让朋友客户难辨真伪”等诱导性表述。多数商家仅核查素材内容，对用途与来源未作甄别。

为深入了解AI工具自身的内容审核机制，记者对用户使用较多的10款文生图、文生视频AI工具进行实测，以对明星“一键变脸”等指令测试其技术边界与审核能力。

结果显示，仅一款AI工具明确拒绝生成真实公众人物虚拟图像。部分应用在简单操作后，即可生成高度逼真的明星虚假内容。

例如，“即×”起初会以“涉及肖像权”为由拒绝明星换脸、变装指令，但记者选择替代选项后仍可生成；“豆×”“文×”等直接执行指令且无显著风险提示；“通×”“天×”等仅在生成内容角落位置标注不清晰的“AI生成”字样。

值得注意的是，部分AI工具在首次接收到涉及真实人物的指令时会予以拒绝，但通过选择平台的替代选项或更换表述，用户仍可绕过初始限制，实现图片或视频生成。

有平台方公开回应称，AI内容侵权识别属于行业性技术难题，恶意仿冒账号在持续与平台进行技术对抗，未来会持续加大技术投入，积极应对挑战，维护创作者、商家及消费者的合法权益。

漫画/李晓军

公司未经授权『换脸』制作模板被判侵权

□ 本报记者 张雪泓

近日，北京互联网法院对外发布一起网络侵权责任纠纷案。法院认为，被告未经授权对他人视频进行AI换脸处理，构成对他人个人信息权益的侵害，遂判决被告某科技文化公司向原告廖某书面致歉，并赔偿其精神损失及精神抚慰金。

廖某是一名古风短视频博主，在全网拥有较多粉丝。他发现，某科技文化有限公司在未经自己授权同意的情况下，使用自己出镜的系列视频制作换脸模板，并上传至某软件中供用户使用以此牟利。廖某认为，该公司的行为侵犯了自己的肖像权与个人信息权益，故诉至法院，要求对方书面赔礼道歉、赔偿经济损失与精神损失。

庭审中，被告某科技文化有限公司认为，平台发布的视频均有合法来源，并且面部特征并非原告，公司并未侵害原告肖像权。此外，涉案软件所使用的“换脸技术”实际由第三方提供，公司并未处理原告的个人信息，未侵害原告的个人信息权益。

法院认为，被告的行为不构成对原告肖像权的侵害。涉案换脸模板视频与原告创作的系列视频的妆容、发型、服饰、动作、灯光及镜头切换呈现一致特征，但出镜人的面部特征均不相同且并非原告。涉案软件通过第三方公司的服务实现换脸功能，用户交纳会员费可以解锁所有换脸功能。判断是否侵犯肖像权的关键在于是否具备可识别性，可识别性强调肖像的本质在于指向特定的人，而肖像的范围以面部为核心，也可能涉及独特的身体部位、声音、识别性较高的特定动作等能够与特定自然人对应的部分。本案中，被告虽然使用原告的视频制作视频模板，但并未利用原告的肖像。模板中所保留的妆容、发型、服饰、灯光、镜头切换等要素并非与特定自然人不可分割，与自然人与生俱来的人格要素存在区别。同时，被告将视频模板提供给用户使用的行为并未丑化、污损、伪造原告肖像，因此，不构成对原告肖像权的侵害。

法院认为，被告的行为构成对原告个人信息权益的侵害。被告收集了包含原告人脸信息的出镜视频，将该视频中的原告面部替换为自己提供的照片的面部，该合成过程需要将新的静态图片中的特征与原视频部分面部特征、表情等通过算法进行融合。上述过程涉及对原告个人信息的收集、使用、分析等，属于对原告个人信息的处理。被告处理该信息未经过原告同意，因此构成对原告个人信息权益的侵害。对于被告擅自使用他人制作的视频侵害他人创造性劳动成果的，应由相应权利人主张权利。

最终，法院判决被告向原告书面致歉，赔偿原告精神损失及精神抚慰金。该判决目前已生效。

北京互联网法院副院长孙铭溪认为，近年来，AI合成技术的滥用问题屡见不鲜，其中不乏利用AI换脸、合成声音等进行恶搞、制作虚假信息的行为，这些行为可能涉及民事责任、行政责任乃至刑事责任。具体来说，未经授权使用他人肖像、声音合成内容，直接侵害人格权，行为人需承担停止侵害、赔礼道歉及赔偿损失等民事责任。未对合成内容进行显著标识违反网络信息管理规定，将受到警告、罚款乃至停业整顿等行政处罚。非法获取、泄露生物识别信息可构成侵犯公民个人信息罪；利用伪造视频实施诈骗、敲诈勒索则分别构成相应财产犯罪；制作传播淫秽合成内容还可能涉嫌制作、传播淫秽物品罪。

专家：强化数字水印与平台责任

□ 本报记者 赵丽
□ 本报见习记者 丁一

近年来，人工智能技术迅猛发展，AI换脸、拟声技术逐渐从专业领域走向大众应用。然而，随着生成内容逼真度的提升和使用门槛的降低，此类技术已被滥用于制作虚假视频、侵犯个人隐私乃至实施诈骗等违法违规活动。

AI换脸换声技术背后潜藏哪些法律风险？该如何有效治理？带着问题，《法治日报》记者采访了中央财经大学法学院教授王叶刚，北京航空航天大学法学院副教授王天凡和京都律师事务所律师常莎。

记者：近年来，利用知名人士形象进行AI换脸换声的视频层出不穷，一些有心人士将其用于直播带货甚至诈骗。这种行为可能触犯哪些法律法规？

常莎：通过AI技术生成明星抑或普通民众的人脸形象，前提是要采集、提取其面部特征生物识别信息。这些信息属于法定的敏感个人信息，侵权方未经授权使用，构成对公民个人信息权益的侵害；还可能侵犯他人肖像权、名誉权等。

若平台“明知”侵权方利用AI换脸实施诈骗、销售假冒商品等犯罪，仍为其提供帮助，则可能构成帮助信息网络犯罪活动罪。

AI换脸直播事件中，侵权方的直播带货属商业广告行为，其通过AI伪造明星肖像与声音，营造名人本人推荐的虚假表象，使消费者因信任明星作出购买决策，违反诚信原则，符合虚假广告认定标准。同时，AI换脸行为还可能涉嫌不正当竞争。

此外，伪造特定名人形象、虚构事实以获取非法收入，主观上具有非法占有目的，客观上实施了欺骗行为，符合诈骗罪的基本构成要件，该行为涉嫌构成诈骗罪。

记者：目前针对AI换脸、换声类视频，我国有哪些相关规定？此类乱象屡禁不止，是否反映出相关法规尚存不足？

王叶刚：对于此类视频，我国目前已有《生成式人工智能服务管理暂行办法》《互联网信息服务深度合成管理规定》《人工智能生成合成内容标识办法》等规定。其中，《互联网信息服务深度合成管理规定》要求，任何组织和个人不得利用深度合成服务从事危害国家安全和利益、损害国家形象、侵害社会公共利益、扰乱经济和社会秩序、侵犯他人合法权益等法律、行政法规禁止的活动。深度合成服务提供者和技术支持者提供人脸、人声等生物识别信息编辑功能的，应当提示深度合成服务使用者依法告知被编辑的个人，并取得其单独同意。

《人工智能生成合成内容标识办法》要求相关行为人必须对AI生成图片或者视频作出标识。但仅依靠内容标识制度，还无法彻底遏制此种行为。标识只能解决AI生成内容的识别问题，无法解决相关的责任认定和承担问题。同时，从责任认定层面来看，标识行为本身在法律上具有何种意义，特别是其在民事责任认定方面具有何种义务，缺乏明确规定。

记者：对于AI换脸换声技术背后潜藏的法律风险，该如何有效治理？

王天凡：解决问题的关键路径，在于《人工智能生成合成内容标识办法》所鼓励的“数字水印”。对数字水印的技术要求和准确性、稳定性等要进一步提升，例如确保水印从生成源头即被嵌入。为此，相关部门应尽快建立明确的技术标准，同时推进构建第三方投诉反馈机制。

王叶刚：应进一步明确删除、篡改、伪造、隐匿标识行为的法律责任。现有相关规定只是规定行为人不得实施恶意删除、篡改、隐匿标识的行为，但行为人实施相关行为需要承担何种法律责任，相关规定亟待进一步明确。

对于通过技术手段能够发现的AI换脸直播带货行为，平台应当尽到必要的审核义务，在发现相关行为后，可以考虑采取禁止或者限制直播、下架产品或者服务等措施。

实现多部门联合执法，网信部门、市场监管部门等在行使本部门职权的基础上，与其他部门进行沟通、配合，实现对违法AI换脸换声等行为的全过程监管和治理。

常莎：在技术方面，防范与创新并重，支持研发反深度合成技术，提升鉴别能力，为监管部门和个人提供可靠的技术工具。

个人应提高对肖像权和隐私权的保护意识，谨慎发布包含个人清晰肖像、独特声线的影像资料，避免给不法分子可乘之机。同时，如果侵权行为发生，则应积极通过法律途径维护自身合法权益，及时收集并保留侵权内容截图、传播记录等相关证据，以便为后续维权提供有力支持。