



语音伪造技术对声纹司法鉴定的挑战^{*}

白雪桐 王秋云

甘肃政法大学，兰州

摘要 | 深度学习是人工智能技术的重要分支，被广泛应用于语音、视频图像处理以及无人驾驶等领域。技术的“双刃剑”属性，让深度学习技术在快速发展的同时也引起人们对其风险隐患的担忧。本文介绍了基于深度学习的语音伪造技术、声纹鉴定技术发展现状，并探讨声纹司法鉴定在人工智能时代的发展路径。

关键词 | 深度学习；语音伪造；声纹鉴定

Copyright © 2021 by author (s) and SciScan Publishing Limited

This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).

<https://creativecommons.org/licenses/by-nc/4.0/>



一、引言

人工智能经过几十年的发展，理论和技术应用已经非常成熟，各种人工智能产品已走入寻常百姓的生活，对人类社会生活造成全方位的影响和变革。深度学习技术作为当前最热门的研究方向，正引领新一轮的人工智能浪潮，以听觉领域为例，深度学习为当前声纹识别、自然语言处理、语音合成等的发展提供了有力的技术支撑。但是从科技哲学的角度来看，技术是“双刃剑”，当基于深度学习的语音合成技术被用于“语音伪造”，“deep-learning”变成“Deepfake”，便会对国家安全，社会稳定以及个人隐私构成潜在威胁。

语音是传递我们表达内容和身份特征的重要信号，深度学习技术的出现，为语音重新赋能。目前，智能语音应用非常广泛，在智能安全领域，可通过语音进行人机交互，手机、防盗门等都能够通过语

音识别解锁；在金融领域，已出现安全可靠的声纹支付技术；在游戏、影视行业，通过语音合成技术可以为虚拟角色生成特定的声音。而不法分子同样可以利用这项技术，通过合成特定对象的声音来恶意散播虚假信息、实施诈骗等，在国内外已经有多起利用伪造语音进行电信诈骗的网络犯罪活动，随着深度语音伪造技术的不断发展，犯罪分子可通过更加低成本的技术手段获取伪造语音进行犯罪，逃避侦查。深度语音伪造技术的出现给法庭科学带来巨大的挑战，声纹司法鉴定技术作为法庭科学的重要组成部分，在人工智能时代的发展是机遇与挑战并存的。

^{*} 基金项目：2020年度甘肃政法大学研究生科研创新项目（2020030）；2020年度甘肃省科技创新项目（20CX4ZA079）；2018年度兰州市科技计划项目（6100417）。

二、语音伪造技术概述

（一）深度伪造技术简介

深度伪造一词最初源于英文“deepfake”，美国在其发布的《2018年恶意伪造禁令法案》^[1]中将“deep fake”定义为“以某种方式使合理的观察者错误地将其视为个人真实言语或行为的真实记录的方式创建或更改的视听记录”。深度伪造技术的工具属性决定其本身是没有善恶的，给“深度伪造”贴上贬义标签的往往是不法分子。近年来深度伪造技术用于犯罪的情形逐渐增多，主要是音频视频图像领域的伪造，如“ZAO”AI换脸软件以及国外“FakeApp”的出现，引起了人们对如何规制深度伪造技术的激烈讨论。

（二）语音伪造技术

语音伪造技术则是以深度学习技术等AI算法为支撑，通过对特定人物音色、语言要素等的采集与训练，合成模拟出以假乱真的目标人物或特定风格的语音。例如，根据ASI数据科学公司的一项测试，通过语音生成算法，只需要借助2小时的语料并训练5天时间，就可以模拟一份足以使人相信的特朗普向俄罗斯宣战的语音。^[2]

当前我们常见的语音伪造技术主要包括语音伪装、语音合成、语音转换、重放攻击等。^[3]伪装语音是司法实践中，犯罪分子为干扰取证所采取的最为常见的一种手段，是一种非自动化的手段，如模仿特定人物说话，用特定方言说话，或者通过捂嘴、捏鼻子等手段阻挡气流的流通，使得话者的语音能量发生改变。语音合成和语音转换技术在深度学习技术不断深化的背景下得到了新的发展，是将文本内容生成特定人物的声音输出，或者将一种声音转换成另一个人的声音输出的一种方法。重放攻击是通过录音设备获取目标人物语料后，再利用设备进行播放或者将语料二次剪辑拼接后进行播放，以达到欺骗人类听觉或欺骗机器的目的。下文对语音合成技术进行专门介绍。

（三）语音合成技术

作为当前非常热门的研究方向，语音合成技术已经非常成熟，被广泛地应用于智能客服、有声读物、AI音响、地图导航等场景。以当前非常热门的导航软件明星语音包为例，其合成语音的自然度、

流畅度与真实语音在听觉上已非常接近，足以达到以假乱真的效果。

语音合成方法主要分为三种：文本—语音转换（TTS）、语音转换（VC）、拼接式语音合成，陶建华^[4]等人进一步在此基础上将语音伪造分为对语音身份风格、语音音色与韵律的伪造。其中语音身份风格伪造主要是利用深度学习方法，通过构建模型对输入的说话人语音进行声纹特征提取，即能够模拟特定说话人声音；语音音色伪造则是通过语音拼接实现，语音拼接可以理解为我们网络上常见的“鬼畜”音频，实现前提是需要大量目标人物的语音资料。此外，陶建华团队从言语生成和感知深层机理理解，语言、口语和情感处理的深层次分析等多角度出发，提出了高效、鲁棒的自然口语语音交互技术。

我们熟知的一些头部科技公司在语音合成技术领域取得了巨大的成果。谷歌的研究团队提出原始音频生成模型WaveNet以及端到端的语音合成系统Tacotron，可以实现将语音从一种语言直接转换为另一种语言，且在翻译后的语音中保留原始说话者的声音特征，并且进一步将二者结合，使合成效果达到媲美人类说话的水平；^[5]百度在Deep Voice（实时神经文本语音转换）和Deep Voice 2（多说话人神经文本语音转换）的基础上提出了Deep Voice 3：带有卷积序列学习的尺度文本语音转换，该架构能够将字符、音素和重音等文本特征转换为各种声学特征，进而将其作

[1] Ben Sasse. S. 3805—Malicious Deep Fake Prohibition Act of 2018 [EB/OL]. [2021-11-16]. <https://www.congress.gov/bill/115th-congress/senate-bill/3805>.

[2] 袁一雪. 让奥巴马用中文演讲的“声音魔术”[J]. 中国科学报, 2016(4).

[3] 任延珍, 刘晨雨, 刘武洋, 等. 语音伪造及检测技术研究综述[J/OL]. [2021-11-16]. <http://kns.cnki.net/kcms/detail/11.2406.TN.20211015.0542.030.html>.

[4] 陶建华, 傅睿博, 易江燕, 等. 语音伪造与鉴别的发展与挑战[J]. 信息安全学报, 2020, 5(2): 28-38.

[5] Michael Chertoff, Anders Fogh Rasmussen. The Unhackable Election: What It Takes to Defend Democracy [J]. Foreign Aff, 2019(98): 160.

为声音波形合成模型的输入；^[1] Facebook AI 研发团队设计了一种新型的端到端的语音生成模型 Melnet, 生成的语音内容不仅可以重现人类的语调, 而且可以像真实的人一样说话。^[2]

三、我国声纹司法鉴定现状

(一) 声纹司法鉴定机构及鉴定人现状

我国司法鉴定机构主要分为职权部门鉴定机构、高校及科研院所鉴定机构和社会鉴定机构。三类鉴定机构中, 职权部门鉴定机构以公安机关为主, 案件类型以刑事为主, 其他两类鉴定机构则民事案件居多。

最新的《声像资料司法鉴定执业分类规定》, 将声像资料司法鉴定分为三大类: 录音鉴定、图像鉴定、电子数据鉴定, 其中录音鉴定又分为录音处理、真实性鉴定、同一性鉴定、内容分析、作品相似性鉴定五类。结合上述分类情况, 统计我国各省市自治区有录音鉴定执业资格的鉴定机构及鉴定人的具体情况见下表 1。

表 1

省份	总机构数	含相关机构数	百分比	总人数	含相关人数	百分比
北京	128	27	21.09%	2182	195	8.94%
天津	28	0	0.00%	592	0	0.00%
河北	192	0	0.00%	2506	0	0.00%
内蒙古	55	5	9.09%	767	15	1.96%
山西	87	2	2.30%	1256	31	2.47%
黑龙江	98	0	0.00%	1427	0	0.00%
吉林	77	2	2.60%	914	4	0.44%
辽宁	100	7	7.00%	1356	53	3.91%
浙江	62	1	1.61%	1170	3	0.26%
江苏	131	2	1.53%	1895	8	0.42%
上海	66	5	7.58%	1108	62	5.60%
安徽	99	0	0.00%	1783	0	0.00%
福建	111	4	3.60%	1485	18	1.21%
江西	133	10	7.52%	2258	70	3.10%
山东	234	1	0.43%	2787	7	0.25%
河南	186	5	2.69%	1505	28	1.86%
湖北	114	1	0.88%	1007	5	0.50%
湖南	183	4	2.19%	1871	24	1.28%
陕西	116	4	3.45%	2083	36	1.73%
甘肃	86	3	3.49%	1206	20	1.66%
宁夏	32	3	9.38%	391	9	2.30%
青海	18	1	5.56%	237	7	2.95%

续表

省份	总机构数	含相关机构数	百分比	总人数	含相关人数	百分比
新疆	30	6	20.00%	529	13	2.46%
重庆	71	2	2.82%	1339	23	1.72%
四川	108	4	3.70%	2309	21	0.91%
贵州	45	4	8.89%	771	24	3.11%
云南	124	3	2.42%	2195	19	0.87%
西藏	3	0	0.00%	33	0	0.00%
广东	211	18	8.53%	2732	101	3.70%
海南	36	1	2.78%	557	4	0.72%
广西	51	2	3.92%	681	12	1.76%
全国	3015	127	4.21%	42932	812	1.89%

注: 数据来源于国家司法鉴定名录网及各省 2020 年度司法鉴定能力验证评价结果, 不包括职权部门鉴定机构。

分析上表可以看出, 在我国能够进行录音资料鉴定的鉴定机构占比为 4.21%, 相关鉴定人占比仅为 1.89%, 且在不同地域的分布呈现出不同的特征。华北地区以北京为主, 天津及河北无相关机构及鉴定人。山西省 86 家鉴定机构中只有两家有声像资料鉴定资格, 且只能进行录音鉴定中的录音真实性鉴定, 声像资料鉴定人 31 人中有 27 人所在机构为山西警察学院司法鉴定中心, 27 人的平均年龄在 50 岁以上, 并且因为鉴定人业务范围描述不够具体, 实际能够进行录音鉴定的鉴定人可能更少。这种在省以内以院校司法鉴定机构为录音鉴定主要力量的现象同样出现在甘肃、湖北、重庆、辽宁等省份。因为资源分布不均, 科学技术发展存在区域差异, 西北部地区相较于东部沿海地区相关鉴定机构及鉴定人数量较少, 鉴定人的能力以及机构相关技术设备相对落后, 遇到疑难录音鉴定案件可能需要将相关鉴定材料送至其他省份进行鉴定。在西藏、黑龙江、安徽、天津、河北等地无录音鉴定执业资格的鉴定机构和鉴定人。此外, 不同省份对司法鉴定人的执业类别描述存在不一致的现象, 大多省份只从法医、物证、声像资料、环境损害四个专业方向进行分类, 而鉴定人具体的执业类别没有进行详细分类。

[1] Wei P, Peng K, Gibiansky A, et al. Deep Voice 3: 2000-speaker neural text-to-speech [C]. ICLR, 2018: 568-576.

[2] Gibiansky A, Arik S, Diamos G, et al. Deep voice 2: Multi-speaker neural text-to-speech [C] //Advances in neural information processing systems, 2017: 2962-2970.

（二）声纹鉴定技术现状

“声纹”的概念最早由 bell 实验室在 20 世纪 40 年代提出，他们通过肉眼首次完成语谱图的匹配，并提出通过此方法进行声纹识别的可能性；之后信号处理技术用于声纹识别，声纹的研究进入新的阶段；70 年代以后，开始利用共振峰及基频轮廓进行声纹识别；80 年代以后，人工神经网络法等技术开始被运用于语音识别研究；进入 21 世纪，人工智能技术飞速发展，深度学习技术逐步成为声纹识别研究的主要方法。

在司法领域，语音在 17 世纪就被用于司法身份认证。声纹同指纹、虹膜、基因等都是生物识别的重要参考特征，声纹识别技术在我国司法领域的应用起步较晚，20 世纪 80 年代，我国开始对声纹鉴定技术的研究，从国外引进到自研，取得一系列的成果，除了传统的一些技术方法，和常见的计算机辅助软件的应用，现在已经出现多种智能鉴定系统，为声纹司法鉴定注入新的活力。

目前我国的声纹鉴定主要分为两种模式，一种是在侦查中，形成鉴定意见或结论提供给侦查部门，用于指导侦查方向，另一种是在庭审过程中作为证据使用。声纹司法鉴定在操作过程中主要依靠鉴定人对音频材料的听辨，利用语言学及语音学进行特征分析，对可供比对的语音单元进行声学分析，对共振峰、基频、音强及其他声学特征进行定量统计，从而的得出可靠的鉴定结论。由于声纹司法鉴定特殊性，当前在鉴定过程中主要依赖鉴定人的经验性判断，和一些常用软件如 AU、praat 等的计算机辅助操作，但在语音技术日新月异的今天，声纹鉴定技术也应当及时更新，以应对新的挑战。

（三）语音鉴伪技术现状

深度学习技术的出现，使人们对语音识别技术的研究不断深入，当前对于语音伪造与鉴伪技术的研究，技术相对成熟的一般是一些大型的科技公司，尤其是语音识别技术在我国金融、健康、娱乐等领域的应用非常广泛。而在司法领域，大多数司法鉴定机构通常只注重鉴定实务，由于声纹鉴定案件数量相对较少，且鉴定机构缺少研究能力，针对语音伪造的研究则相对落后。这一现状无疑会对声纹司法鉴定的发展带来不利影响。

四、声纹司法鉴定面临的挑战

国内外已经开始出现 Deepfake 诈骗案例，2019 年，英国某能源公司高管在一个小时之内通过匈牙利供应商向一诈骗犯转账 24 万美元，仅因诈骗犯通过 AI 语音合成技术仿造了其母公司 CEO 的声音。美国网络安全公司 Symantec 研究人员表示，他们发现至少有三起公司高管的声音被人模仿以用于电信诈骗。其中一起案件的损失总计达数百万美元。

在《中国声纹识别产业发展白皮书》1.0 及 2.0 中都提到，随着语音合成技术水平的不断提高，利用伪造语音进行假冒攻击的威胁性将会越来越大。当前声纹识别系统已应用于公共安全、银行、个人隐私等领域，而如何阻止伪造语音通过声纹识别系统，建立有效的防御体系，仍是当前非常重要的研究方向。

声纹司法鉴定在实务中最主要的鉴定类别为话者同一认定，通过对不同语料的综合分析和特征比对，确定是否为同一人所说的语音。深度语音伪造技术不同于常见的通过模仿的伪装语音或是重放提前录制的语音，以深度学习技术为支撑，当合成语音的流畅度、可懂度与真实语音越接近，合成语音的质量越高，鉴定人在听觉检验的过程中，就越容易被欺骗，做出错误的判断。同样对于频谱检验，鉴定人通常借助一些常见软件获取频谱数据进行比对分析，当用于深度学习的某个特定人物的语料越多，合成语音与自然语音的频谱数据差异性就越小，鉴定人在比对的过程中就越容易做出错误的判断，形成错误的同一性鉴定意见。

五、人工智能时代声纹鉴定的发展路径探析

随着语音合成技术的成熟，滥用语音伪造的情形会越来越多，声纹鉴定的难度也会不断增加。因此当务之急是在“大数据 + 人工智能”背景下，推进声纹司法鉴定的技术创新和应用，提升司法鉴定能力。

（一）加强新技术的研究与开发应用

在科学技术飞速发展的今天，涌现出一大批新型科学技术，大数据技术、网络安全、云计算技术、生物技术、人工智能技术等世界科技前沿领域成为主导人类进步的新力量。司法鉴定是一门综合多种

专业、领域的交叉学科,作为法庭科学的重要部分,需要极高的专业水平和技术能力。司法鉴定同样应当结合当前新技术的发展,及时将这些新技术、新手段、新方法吸收转化,加强司法鉴定专业技术的更新迭代,有效应对社会发展过程中出现的新问题。

当前人工智能、大数据等技术已广泛应用于计算机视觉、声音研究等领域,人脸识别、声纹识别的工业应用也非常成熟,但大量的新技术却不能及时转化为鉴定的技术手段。通过深度学习技术合成的伪造语音已经被用于犯罪,而鉴定人员仍在通过传统的一些软件以及个人鉴定经验进行分析,因此,司法鉴定机构应当关注当前新型技术的发展,实现新技术在司法鉴定中的具体运用,提升司法鉴定能力与司法鉴定意见的证明力,推进司法鉴定技术的转型升级与更新换代。

此外,由于我国区域经济发展的不平衡,使得各地区科学技术的发展也存在极大差异,司法鉴定技术的发展也不例外。东部沿海地区经济水平高,科技公司以及高水平科研院所密集,这些地区的司法鉴定机构的鉴定能力,以及鉴定人员的经验水平与经济欠发达地区相比,存在较大差别。我国存在大量的社会鉴定机构,这些机构要考虑成本和盈利,缺乏科研能力和科研资金的投入;同时还有大量的院校司法鉴定机构,高校重视理论与教学,科研成果不能有效转化为鉴定能力。因此,应当重视司法鉴定的跨区域技术交流与合作,加强高校、科研机构以及科技公司的合作,共同推进司法鉴定能力的提升。

声纹鉴定的发展应重视理论技术的研究,尤其是提高对深度伪造语音的鉴定能力,形成一套能够有效应对伪造语音威胁的鉴伪方法。声纹作为重要的生物识别特征,应重视不同个体言语特征的分析,从个体的差异性寻找有效鉴别的方法;探讨能够更加快速有效法学自然语音与合成后的机器语音的差别;其次,还应当重视相关机器学习算法的开发,提升计算机自动化鉴伪的能力。

(二) 加强声纹鉴定人才培养

除了新技术的开发与应用,鉴定人员的培养同样应当重视。当前司法鉴定主要有物证类、法医类、视听资料等方向,司法鉴定是一门交叉学科,需要多种综合专业知识,以本文讨论的声纹司法

鉴定为例,声纹鉴定专家需要具备计算机、语言学、语音学以及法学等方面的知识。当前不少高校在本科、硕士以及博士阶段直接或间接开设与鉴定科学相关的专业,但对于专门的司法鉴定人才的培养还处在摸索阶段。如司法鉴定硕士专业的设置,大多学校的学科设置主要集中在司法鉴定的理论制度、管理体制等方面;对于鉴定实务则是注重鉴定方法和经验的传授。因此,我国法庭科学教育目前定位较为模糊,并不能很好地培养高素质、有能力的鉴定人才。

对于鉴定人才培养,要以鉴定业务的需求为人才培养的导向。以声纹鉴定为例,以现有高等教育为平台,优化学科建设,本科阶段重视鉴定实务能力的培养,专业知识与法学知识并重;硕博阶段则以理论研究、技术研发为主,建立起声纹鉴定的技术队伍,形成符合行业发展的人才梯队,不断为司法鉴定行业输送新鲜血液。同时优化司法鉴定人员准入制度、鉴定资质认定制度,吸收具备专业能力高精尖专家进入鉴定队伍,以应对鉴定过程中出现的疑难问题。

除了鉴定人才的培养,司法鉴定人同样应当重视鉴定人的培训工作,注重鉴定人能力的提升。声纹鉴定中,各种深度学习技术的出现,使得专家经验知识在鉴定中的可靠性下降,对于不断革新的技术方法,鉴定人要通过定期培训学习,定期组织行业的学术交流,及时更新鉴定技能,才能应对鉴定中出现的新问题。

(三) 加快公安部“声纹库”建设

与指纹、DNA类似,声纹也是个人身份确认的生物特征。电信诈骗案件呈现井喷趋势,声纹特征在公共安全领域的应用价值日益凸显。为有效防止此类案件给人民财产安全带来的危害,公安部将声纹识别技术纳入犯罪预防和侦破的方案中,各地公共安全领域相关部门也在加大声纹采集力度。与此同时,声纹数据库建设工作和建库规范也开始提上日程。

我国“声纹库”的建设,公安部与省级建设并行,通过省部两级数据库的联通,实现业务的联动。与指纹数据和DNA数据库类似,以声纹数据的存储、管理为基础,建立个人声纹档案,对重点人员进行定位跟踪,在相关案件发生后,通过获取的关键线

索,利用大数据分析比对,能够有效提高案件侦破效率。同时声纹数据库的建设,能够提升声纹司法鉴定的效率以及声纹司法鉴定意见的准确度。

(四) 完善司法鉴定标准和规范建设

建设司法鉴定标准体系,包括鉴定的技术标准、鉴定程序标准、鉴定管理标准,鉴定设备软件标准,鉴定意见表述标准等,对涉及专业技术领域的问题予以规范、整合,以建立统一、科学的鉴定技术标准和程序体系。不仅能够为鉴定机构实施鉴定活动提供统一的程序规范以及需遵循的技术标准,对实施鉴定过程是否违反程序,违反程序的程度以及违反的法律后果予以明确,还能够明确法官对鉴定意见证据能力以及证明力审查的标准,为鉴定意见的采信提供依据,实现司法公正。^[1]

司法鉴定标准是为保证鉴定方法科学性和有效性而制定的基本原则和方法规范,对于提高鉴定意见证明力,增强司法鉴定公信力有积极的推进作用。而目前我国尚无专门的司法鉴定相关法律,司法实践中,不同的鉴定方向有不同的鉴定标准和技术规范,在不同的地区和部门也有各自出台的标准和规范,出现国家标准、地方标准、行业规范等不同的标准共存,导致鉴定标准和规范出现重复矛盾现象,不利于司法鉴定的发展。我国现行法庭科学声纹鉴定标准规范以公安部、司法部制定为主,公安部出台有《法庭科学录音的真实性检验技术规范(GA/T 1430-2017)》等6个标准,司法部出台有《录音资料鉴定(SF/Z JD0301 001-2010)》等6个标准,最高人民检察院也有《语音同一认定(SPPD-A-1-2015)》等5个执行标准文件。从上述标准的内容来看,同一部门出台的不同规范有重复内容,同一个鉴定内容在两个部门的规范中,规定的内容不尽相同。以录音的真实性检验为例,公安部规定的鉴定意见有检材经过剪辑、未经过剪辑、无法判断三种,而司法部的规定除以上三种外,还有倾向认为录音经过剪辑和倾向认为录音未经过剪辑两种;而对于语音同一认定技术规范,公安部规定不同鉴定意见所依据的检材和样本语音特征总体符合率和差异率要达到要求的百分比,而司法部规定的判断依据只有“存在足够的符合特征”“差异或变化能得到合理解释”等主观性表述。因此,制定统一的、规范的、权威的声纹司法鉴定技术标准

在当前非常重要。

科学技术的发展与司法鉴定的发展相辅相成,不断出现的新技术推动鉴定方法的革新,这个过程中出现的问题即新方法在旧标准体系下能否应用,新方法的有效性、科学性达到什么样的标准,依据新的技术方法作出的鉴定意见才能够被采纳。因此,我国司法鉴定发展的过程中,首先要关注技术标准的革新和新方法的运用,同时形成一套引入高新技术方法与淘汰落后技术的机制,始终能保持标准的先进性。这需要司法鉴定行业与相关领域的专业人士的共同努力,才能保证鉴定所依据的方法和技术标准是科学有效的。

(五) 完善专家辅助人制度

美国国家科学院发布的《美国法庭科学的加强之路》中认为“法庭科学难以摆脱被人为操控的危险”,“对于不完美的检测和分析所产生的证据和证言而言,赋予其过度的证明力存在潜在风险。而且,不准确或者夸大的专家证言在有时也会促成对错误的或者误导性的证据的采纳”。^[2]

我国的法庭科学领域同样存在上述危机,以深度伪造语音对我国声纹司法鉴定的挑战为例,当鉴定人能力不足,专业知识欠缺,所依据的技术方法落后的情况下,作出的鉴定结论必然存在瑕疵,而法官并不能对这一专业领域的知识有准确把握,很容易造成不采纳或错误采纳的结果。因此专家辅助人在鉴定意见采纳过程的作用凸显,专家辅助人在庭审过程中主要是帮助当事人对鉴定意见进行质证,对鉴定意见发表专业意见;其次能够帮助法官全面认识和理解鉴定意见内容,对能否采纳作出判断;还能及时纠正鉴定意见可能出现的错误。^[3]

完善专家辅助人制度,应当拓宽专家辅助人选拔的渠道,严格选拔的要求,吸纳行业权威,具备专业知识的人才进入专家辅助人队伍,建立规范的

[1] 何晓丹, 吴何坚. 再论司法鉴定标准体系的建设[J]. 中国司法鉴定, 2020(1): 87-92.

[2] 美国国家科学院国家研究委员会. 美国法庭科学的加强之路[M]. 王进喜, 译. 北京: 中国人民大学出版社, 2012: 5.

[3] 霍宪丹, 郭华. 中国司法鉴定制度改革与发展范式研究[M]. 北京: 法律出版社, 2011.

专家辅助人人才库。让专家辅助人从科学原理、经验考察等方面对鉴定意见的科学性、准确性、可靠性进行审查,帮助法官对鉴定意见的采纳作出正确的判断。

(六) 加强相关法律法规的制定

深度语音伪造技术的发展依靠大量的语音素材供研究人员建立数据集进行研究,犯罪分子在合成实施犯罪所用语音的时,为了获取大量的语音资料,目前已经出现买卖语音资料的黑色产业链,这种通过非法交易获取他人语音的行为,对于个人隐私和国家安全都有潜在的危险。

2021年11月1日起,我国首部专门针对个人信息保护的法律《中华人民共和国个人信息保护法》正式实施。声纹作为重要的生物特征,在生物识别技术日新月异的今天,极易被犯罪分子利用。《个

人信息保护法》的出台,将有效保护公民的隐私信息,也将进一步规范数据的使用。同时,应当在国家层面制定法律、行业内部制定规则,进一步加强对深度伪造技术的规制,使技术更好地为人类社会服务,而不被犯罪分子滥用。

六、结语

近年来人工智能技术正飞速发展,基于深度学习的深度伪造内容越来越成熟,除了语音伪造,AI换脸、计算机视觉伪造等同样值得关注。司法鉴定是保障诉讼活动顺利进行,维护公平正义的重要手段,应当直面新技术带来的挑战,加强新技术的运用,不断变革完善,提升疑难案件的应对能力,更好地为诉讼服务。

(责任编辑:邵灿灵)

Challenges of Deep Fake Voice to Voice Forensic Identification

Bai Xuetong Wang Qiuyun

Gansu University of Political Science and Law, Lanzhou

Abstract: Deep learning is an important branch of artificial intelligence technology, which is widely used in the fields of voice, video and image processing and unmanned driving. The "double-edged sword" attribute of technology makes deep learning technology develop rapidly, but also causes people to worry about its risks and hidden dangers. This paper introduces the development status of voice forgery technology and voice print identification technology based on deep learning, and discusses the development path of voice print forensic identification in the era of deep learning.

Key words: Deep learning; Deep fake Voice; Voice forensic identification