

第 1 章 初识人工智能

学习目标：

- 辨析数据、信息、知识、智慧、智能；
- 了解人工智能与人类智能的关系。

人工智能引领了新一代的科技革命，目前在以肉眼可见的速度渗透到我们生活的方方面面，对人类生存、生活、精神、发展产生深远的影响。党的二十大报告指出：“推动战略性新兴产业融合集群发展，构建新一代信息技术、人工智能、生物技术、新能源、新材料、高端装备、绿色环保等一批新的增长引擎。”将人工智能放在新增长引擎的前列。本章将先从数据、信息、知识、智慧、智能等概念入手，再探讨智能的类型以及人工智能的应用场景，让学习者建立起对于信息及人工智能的基本知识框架。

1.1 从数据到智能



数据是指通过特定的手段和载体，将客观事实进行逻辑归纳和记录的结果，其存在的形式多种多样，如符号、文字、数字、图像、音频、视频等。由于数据只是对客观事实的记录和描述，就其本身而言，不具有意义，只有经过加工、提炼的数据，才具有潜在的意义。从数据的存在状态来看，存储于相关载体的数据是静态数据，在系统数据流中的数据是动态数据。从数据与信息、知识的关系角度而言，数据是指构成信息和知识的原始素材，是产生信息、知识、智慧和智能的基础条件，具有无逻辑、离散等特征。

信息是加工后的数据。由于原始数据的类型多种多样，且具有无逻辑、离散等特征，人类为了更好地认识世界和改造世界，势必要通过直接或间接的方式，将原始数据经过加工改造，使之成为可以服务于人类规则的数据。

知识是系统化的信息，是对信息进行筛选、处理、综合、分析之后产生的彼此之间相互关联的数据，它不是信息的简单相加，而是多维信息的有机统一。古希腊著名哲学家柏拉图指出：一条陈述能称得上是知识必须满足三个条件，即它一定是被验证过的、正确的，而且被人们相信的。因此，知识本身有真知识和假知识之分——凡是经不起验证的、不能令人信服的系统化信息，它本身不能算是真知识，只能算是假知识。

智慧是指人类通过知识的系统化掌握，从而锻炼出发现问题、分析问题和解决问题的思维能力。智慧的形成过程就是从感性知觉到理性思维、直觉与灵感的过程。

智能是智慧和能力的合称。由智慧指导行为表达的过程，就是智能的过程。智能分为人工智能和人类智能，人工智能包括弱人工智能(weak AI)和强人工智能(strong AI)。按

照霍华德·加德纳的多元智能理论,人类智能又包括语言智能、数学逻辑智能、空间智能、身体运动智能、音乐智能、人际智能、自我认知智能、自然认知智能。

数据、信息、知识、智慧与智能之间是一种逐渐升维的关系。在特定的条件下,彼此之间能够实现相互转换和升维发展。经过加工的数据成为信息,信息之间的相互关联成为知识,知识的系统化开始产生智慧,智慧加上能力则形成智能。



辨析

数据、信息、知识、智慧与智能之间是原始、加工、应用的关系:

数据=记录的事实

信息=数据+意义

知识=信息+理解

智慧=知识的系统化

智能=智慧的应用

它们的关系如图 1-1 所示。



素养提升

在新冠肺炎疫情防控过程中,我们国家取得了举世瞩目的良好效果,精准防控、快速响应的防控手段得益于新一代信息技术,如移动互联网、人工智能技术的应用。流动人员的手机定位记录、出行记录等都是数据,通过对这些数据进行处理,就可以得到每一位相关人员的流动信息,进而可以建立疫情时期的地域人员流动模型,这样就形成了知识;通过人员流动模型,对疫情发展进行预测与研判,进而制定疫情防控的政策。这就是智慧。

有知识不一定有智慧,最有名的例子就是纸上谈兵的赵括,他虽有丰富的理论知识,但是无法在合适的场景选择适当的知识去做决策,所以他缺乏智慧。

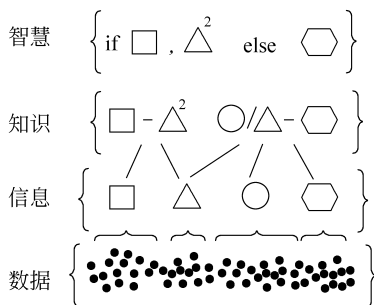


图 1-1 数据、信息、知识、智慧的关系



1.2 人工智能与人类智能

人工智能与人类智能

大家是否畅想过未来:

“机器能独立思考,并像人一样胜任任何智力性任务,就像科幻电影里的机器人一样,无所不能。”

“将来人工智能将取代人类的一切工作,人类将衣食无忧、享受生活。”

那么人工智能真有这样的能力吗?

以下史实内容,来自沃尔特·艾萨克森的《创新者》。

在 1950 年 10 月的哲学期刊 *Mind* 上,图灵发表了论文 *Computing Machinery and Intelligence*,其中提出了一个概念——“图灵测试”(Turing Test),它为人工智能模仿人类智能提供了一个基线测试,即“如果一台机器输出的内容和人类大脑别无二致,那么我们就没有理由坚持认为这台机器不是在‘思考’”。

图灵测试也就是图灵所说的“模仿游戏”,其操作很简单,即“一位询问者将自己的问题写下来,发给处于另外一个房间之中的一个人和一台机器,然后根据他们给出的答案确定哪

个是真人。如果在相当长时间内,他无法根据这些问题判断对方是人还是计算机,那么就可以认为这个计算机具有同人相当的智力,即这台计算机是能思维的”。

图灵测试试图解决长久以来关于如何定义思考的哲学争论,他提出一个虽然主观但可操作的标准:如果一台计算机表现(act)、反应(response)和互相作用(interact)都和有意识的个体一样,那么它就应该被认为是有意意识的。

一般认为,人类智能的特点在于它是有自我意识的。目前而言,人工智能被定义为模仿与人类思维相关的认知功能的机器或计算机,其本质是对人类思维的模仿,而没有自我意识。但是,也有不同的观点认为既然人脑也是基于结构的思维机器,那么人工智能模仿人脑之后也会演化出意识。这样就产生将人工智能分为强人工智能与弱人工智能两种类型的观点。

(1) 强人工智能。强人工智能观点认为有可能制造出真正能推理和解决问题的智能机器,并且这样的机器能将被认为是有感知的,有自我意识的。强人工智能可以分为两类:类人的人工智能,即机器的思考和推理就像人的思维;非类人的人工智能,即机器产生了和人完全不一样的感知和意识,使用和人完全不一样的推理方式。

(2) 弱人工智能。弱人工智能观点认为不可能制造出真正能推理和解决问题的智能机器,这些机器只不过看起来像是智能的,但是并不真正拥有智能,也不会有自主意识。

当前,人工智能虽然在特定领域已超越人类智能。例如,在视频游戏、国际象棋、蛋白质折叠等科学问题和语言建模方面,特别是在人工智能程序 AlphaGo(阿尔法围棋)战胜人类围棋冠军之后,对很多人的认知造成很大冲击。但是人工智能在一些人类看来很简单的领域往往难以取得比较好的效果。例如,在感觉运动技能方面,人类无须经过特别训练即可掌握行走、奔跑和跳跃等技术,但人工智能还做不到;在视觉常识推理方面,也就是依据图像所呈现的场景回答问题,人工智能依然低于人类水平;在语言理解方面,人工智能在简单问题上已经超越人类基准,但在需要进行逻辑推理的复杂问题上,人工智能的效果并不理想。

2022 年 11 月 OpenAI 发布了 ChatGPT(chat generative pre-trained transformer,聊天生成预训练转换器),被认为是人工智能领域的一次重大突破,它展示了基于大型语言模型的聊天机器人的强大能力和广泛应用。ChatGPT 不仅可以用自然语言与用户进行流畅的对话,还可以根据用户的需求完成各种复杂的语言任务,如写作、编程、问答、摘要等。另外,OpenAI 不断更新并加强其能力,2023 年 3 月 OpenAI 发布了能力更强的 GPT-4(原先 ChatGPT 基于 GPT-3.5 版本),其使用更多、更丰富的训练数据,支持多模态任务,处理更长的文本输入,具备更广泛的知识 and 解决问题的能力。

ChatGPT 在很多方面与人类差不多甚至已超过人类:它可以创建论文,制作幽默的帖子,回答一些困难的编程问题,生成图像,给出有用的商业建议,写出非常好的歌曲。GPT-4 在 USBAR(American bar examination,美国律师执业资格考试)里击败了 90% 的人类,在 SAT(scholastic aptitude test,美国高中毕业生学术能力水平考试)阅读考试中击败了 93% 的人类,在 SAT 数学考试里击败了 89% 的人类。

ChatGPT 引发了新一轮人工智能浪潮,与其类似的大语言模型层出不穷:国外有 Google 的 Bard、Meta 的 LLaMA、Anthropic 的 Claude 等,国内有华为的盘古、百度的文心一言、阿里巴巴的通义千问、科大讯飞的星火认知大模型、腾讯的混元大模型……

但是,这一类大语言模型依旧不能称为强人工智能。强人工智能是指能够完全模拟人

类智能的系统,具有自我意识和自主学习、推理、规划、解决问题等能力。GPT-4 虽然在某些方面达到了人类水平,但在许多实际场景中仍然不如人类。例如,它不能保证其回答的事实准确性和逻辑正确性,有时会产生错误或误导性的信息。它也不能理解编程语言的语法和语义,只是在获取代码片段之间的统计相关性,因此不能编写和调试复杂的计算机程序。总之,当前的人工智能还不具备自主学习与思考的能力、没有自主意识,依旧还是弱人工智能。



素养提升

李开复曾说:“人工智能将夺走许多单一任务、单一领域的工作。人类拥有人工智能所没有的能力,我们可以概念化、制定战略和进行创造。今天的人工智能只是一个可以接收数据并进行优化的聪明的模式识别器,但是,世界上有多少工作是可以优化的简单重复任务呢?”

1.3 应用领域

1.3.1 生活

目前人工智能渗透到我们生活的方方面面,给我们带来很多便利与乐趣,以下列举几个典型的应用场景。



生活

1. 自动驾驶

在 2021 年 10 月举行的国家“十三五”科技创新成就展上,众多自动驾驶研发及落地成果悉数亮相。其中,极狐阿尔法 S 自动驾驶车、百度“汽车机器人”、斑马智行智能座舱操作系统吸引了众多参会者的关注。尽管自动驾驶技术发展炙热,但其具有令人生畏的技术复杂性,在这个领域中,中国企业和美国企业双双走在世界前列。

自动驾驶技术的基本原理是通过感应装置感知周围环境的情况,计算机系统依据这些信息作出指令以及执行指令进行驾驶。这些感应装置包括激光雷达、毫米波雷达和摄像头,通过这些装置对周围环境进行精准识别,自主避让前方障碍物,进行自动驾驶。此外通过计算机自主学习、高精度地图定位、网络通信和激光雷达等技术,利用环境感知、自动决策和控制等技术,对各种复杂环境和突发状况采取行之有效的措施,因此自动驾驶技术是多学科交叉协同发展的。

参考美国汽车工程师学会对自动驾驶的分级定义,根据智能化程度的不同,自动驾驶被分为 L1~L5 共 5 个等级:L1 是指辅助驾驶;L2 是指部分自动驾驶;L3 是指有条件自动驾驶;L4 是指高度自动驾驶;L5 是指完全自动驾驶,即真正的无人驾驶。日渐活跃于公众视野的“无人驾驶”概念,往往是指 L3 及以上级别的自动驾驶。达到 L4 级后,自动驾驶比人类驾驶更安全。世界卫生组织发布的《2018 年全球道路安全现状报告》显示,每年全世界约有 135 万人会在交通事故当中失去生命,而 94% 的交通事故是人为原因造成的。究其原因,危险驾驶是排名第一的类型,总量占到了总刑事犯罪的 1/4 左右。AI 司机既不会醉酒

驾驶,也不会边开车边使用手机或感到疲倦,一些人为因素导致的交通事故会消除。

自动驾驶的场景是非常丰富的,在开放场景中,华为、百度的自动驾驶技术已经达到 L4 级别,如极狐阿尔法 S 华为 HI 版轿车(见图 1-2)不仅配备了华为 L4 级的自动驾驶功能,全车还装有 3 颗 96 线车规级激光雷达+13 颗高清摄像头+6 颗毫米波雷达+12 颗超声波雷达,并且它的芯片运算能力可达到 400TOPS,已经超过了特斯拉一些车型的运算能力。



图 1-2 极狐阿尔法 S 华为 HI 版轿车

目前在北京、上海、广州、重庆、武汉等地已经开展了无人驾驶出租车服务。早在 2020 年 10 月百度无人驾驶出租车服务就在北京全面开放,市民可在北京经济技术开发区、海淀区、顺义区的十多个无人驾驶出租车站点直接免费试乘无人驾驶出租车。在北京的海淀公园提供了无人驾驶小巴体验服务,如图 1-3 所示。



图 1-3 无人驾驶出租车与无人驾驶小巴

在深圳街头出现无人环卫车,如图 1-4 所示,当有车辆靠近时它能及时“躲避”,走近斑马线时会“礼让”行人,到达终点后它还会自动倾倒垃圾。在北京,运行了一系列的环卫作业车,主要功能包括吸扫作业、洗地作业、垃圾收集、垃圾转运四大类 7 种型号,覆盖了传统清扫保洁作业的全流程。这些作业车拥有“眼睛”和“大脑”,集人工智能、机器视觉、图像识别、精准定位等技术于一体。“眼睛”即传感器,分别位于车顶和车身,能 360°感知周边物体;“大脑”即深度学习算法控制技术,能根据实时感知的环境信息,结合高精度地图,制订最优的路径规划,完成道路清扫保洁及垃圾清运转运任务。



图 1-4 无人环卫车

2. 自然语言处理

让全世界拥有相通的语言一直是萦绕在人们心中的梦想。当前人工智能技术实现了用机器翻译不同的语言,从最初只能翻译单词到现在可以整句或通篇翻译,甚至可以直接口译。在任何一个国家,即使看不懂文字,听不懂语言,你也能够借助机器翻译与他人进行交流和沟通,不必再为相互不能理解而困扰。

机器翻译的核心就是自然语言处理(natural language processing, NLP),简单来说,自然语言处理就是用人工智能来处理、理解以及运用人类语言,它体现了真正意义上的“人工智能”。百度机器学习专家余凯说:“听与看,说白了就是阿猫和阿狗也会,而只有语言才是人类独有的。”也就是说只有当计算机具备了处理自然语言的能力时,才算实现了真正的智能。为了让机器能与人自然交流, NLP 有两个核心的任务:一是自然语言理解,即让机器理解人们说的是什么意思;二是自然语言生成,即让机器用人类语言表达出正确的意思。

要实现机器与人的自然交流是非常困难的任务,具体来说有以下 5 个难点。

- (1) 语言是没有规律的,或者说规律是错综复杂的。
- (2) 语言是可以自由组合的,可以组合复杂的语言表达。
- (3) 语言是一个开放集合,我们可以任意地发明创造一些新的表达方式。
- (4) 语言需要联系到实践,有一定的知识依赖。
- (5) 语言的使用要基于环境和上下文。

时至今日, AI 在这些技术领域已经把识别准确率从 70% 提高到了 90% 以上,但只有当准确率提高到 99% 及以上时,才能认定自然语言处理的技术已达到人类水平,这仍然是巨大的困难和挑战。

自然语言处理技术在生活中应用广泛,其典型应用如下。

1) 上下文/情感分析

互联网上有大量的文本信息,这些信息想要表达的内容是五花八门的,通过上下文分析,可以观察人们的行为方式,了解其个性及情感表达。通过这些结果可以进行精准的广告及内容投送,并可进行舆情监测。

例如,美团在服务百万级别的餐饮商户和亿级别 C(consumer,消费者)端用户的过程中,积累了海量的用户生成内容(user generated content, UGC),包含了用户到店消费体验之后的真情实感;美团技术团队通过 NLP 技术对 UGC 进行情感分析,能够有效提取其中的关键情感极性、观点表达,辅助更多用户做出消费决策,同时也可以帮助商户收集经营状况的用户反馈信息,如图 1-5 所示。



图 1-5 美团的情感分析工具

2) 聊天机器人/人工智能客服

随着人工智能语音识别能力的大幅度提升,我们已经习惯于在微信中使用语音转文字功能、开车时直接通过语音说出目的地,而且人工智能能非常好地“理解”我们的表达。

比如,华为发布的儿童陪伴教育机器人“华为小艺精灵”能与人进行非常流畅的交流,甚至可以表达一定的“情绪”。

大家是否发现,当我们拨打客服电话或者接到销售电话时,人工智能客服的比例越来越高,而且其语义表述和沟通表达能力可达到以假乱真的地步,让客户无法分辨“真假客服”。据有关机构预测,到 2025 年,95% 的客服互动将由 AI 技术主导完成。人工智能客服具有高可用性(无须休息,24 小时在线),节省时间(高效,能快速处理数据),低成本、高效益(相对人力成本而言具有成本优势),无偏见、无情绪的优势。

素养提升

在搜索引擎上搜索科大讯飞语音合成(或百度语音合成、华为语音合成),体验中国一线人工智能厂商的产品。可输入任意文本,让它使用极具表现力和类似人类的声音朗读出来,且支持多种朗读风格,包括新闻广播、客户服务、呼喊、耳语以及高兴、悲伤等,甚至支持粤语、东北话、四川话等方言。

3) 机器翻译

机器翻译技术在近年来取得了长足进步,并且逐步逼近平行对译的境界。在西方诸语种之间,机器翻译的准确率已经稳居 90% 以上。中国在这个领域的领先企业是科大讯飞,早在 2018 年时,其就于业内率先提出“听得清、听得懂、译得准、发音美”的 AI 翻译四大标准。当时科大讯飞中英文机器翻译的效果就已经达到英语六级的水平。在 2022 年 8 月发布讯飞翻译机 4.0 之际,提出增加的 AI 翻译机新标准:够自然,旨在让跨语言交流更加顺畅、自然、高效。其最新的讯飞翻译机 4.0 通过前后端一体化的语音识别技术,充分地利用前端由多个麦克风组成的麦克风阵列和后端复杂模型的精细建模能力,大幅提升复杂场景下的语音识别率,通过“基于语言特征强化的多语种机器翻译统一建模技术”并积累超过 120 万的行业术语、6000 万句对的定制语料,使其能够支持 83 种语言在线翻译,16 种语言离线翻译,16 大领域行业翻译,32 种语言拍照翻译。

在各类国际会议中,同声传译对于人类来说是一项很有挑战性的任务,因为人类同传译员需要同时做到听、理解、翻译并说出翻译内容。人类高质量同传一般只能持续 15 分钟,且译出率一般在 60% 左右,漏翻、错翻等情况时常发生。作为 AI 同传技术的领军者,百度 AI

同传首次做到了直接从语音到文字的同传翻译,该模型达到了更低的时延和更准确的翻译,彻底摒除了 ASR(automatic speech recognition,自动语音识别)错误对翻译模型的影响,进一步简化了整体同传框架。百度 AI 同传能达到汉译英准确率为 85.71%,英译汉准确率为 86.36%,并且 PK 3 位经验丰富的人类同传译员,最终结果极具竞争力,在评估所用的 BLEU (bilingual evaluation understudy,双语评估替补)和人工评价双重评价中,百度 AI 同传均达到与人类同传译员媲美的水平。百度 AI 同传已经服务了百度开发者大会、2021 年中国国际服务贸易交易会、中国国际进口博览会等大型会议,此项成果获得国家科技进步二等奖。

在 2022 年北京冬奥会的视频转播中,一位担任手语解说的女士出现在画面下方,见证了中国队夺金的各场比赛。与过往那些经验丰富、身经百战的手语解说员不同,这是她首次进行手语解说工作,却有着完全不逊于前者的准确度和反应速度。这位女士并不是真人,而是由腾讯打造的 3D 手语数智人“聆语”。“聆语”基于《国家通用手语词典》的标准手语和深度的机器学习训练,以及针对体育、艺术等专业领域的优化补充,目前共掌握约 160 万个词汇和语句。在解说比赛时,会先通过机器翻译将比赛解说的语言低延迟转换为高准确率的手语语言表征,再运用腾讯多模态端到端生成模型,进行联合建模及预测生成高准确率的动作、表情、唇动等序列,实现自然专业且易懂度高的手语效果。并且“聆语”具备快速学习补充新词、热词,根据业务场景快速学习专业用语的能力。为应对本次赛事,手语解说针对体育赛事方面的用语做了定向优化,覆盖了超过 15000 个相关词汇。

3. 医疗服务

人工智能在医疗领域主要有以下应用场景。

1) 医疗机器人

医疗机器人是在医院、诊所进行治疗或辅助治疗工作的机器人。医疗机器人的分类方式较多,按照其用途不同,主要可以分为临床医疗用机器人、护理机器人、医用教学机器人、药物配送机器人、其他医疗服务机器人等。在机器人技术的赋能下,中国机器人辅助腔镜手术数量持续攀升。2021 年手术数量在 8.75 万台左右,较 2020 年增长近 85%。而当前最先进的腹腔镜手术机器人为达·芬奇外科手术系统,其由控制平台、4 只机械手臂、1 个三维摄像系统组成,手术视野放大倍数可达到 10 倍以上,能为主刀医生呈现患者体腔内三维立体高清影像,细小的血管和淋巴结也能一目了然。同时,4 个机械臂可模拟人手腕的灵活操作,滤除颤动,超越了人手的精准度,即使在人手不能触及的狭小空间也能精准操作,超越了人手的局限性。达·芬奇外科手术系统还有一项更大的优势,就是创口小。许多肿瘤切除手术需开腹,创口在 15~30cm 内,而达·芬奇外科手术系统手术的创口不到 1cm,仅用创可贴便可覆盖。手术创口小、患者恢复快,能大幅缩短病人的术后住院时间。但是,达·芬奇外科手术系统需要从美国进口,价格昂贵,单台接近 2500 万元人民币,年维护费用也高达 120 万元人民币。我国在此领域奋起直追,自主研发了“术锐”“妙手”等手术机器人,而且利用我国 5G 网络的优势,率先实现了远程手术商用化。

2) 医学影像识别

医学影像是医疗数据最密集的领域,超过 80% 的医疗数据来源于医学影像,人工智能技术已经应用在医疗行业多个领域,而医学影像是应用非常成熟的领域。深度学习算法模型的训练需要海量数据支撑,医学影像由于其数据密集的特性,让以深度学习为代表的人工

智能技术有了广阔的发挥空间,而其中又以 X 光、CT 等类型影像的识别分析最为成熟。在皮肤癌、乳腺癌、前列腺癌、肺结节等病种上,人工智能识别准确率均超过 90% 甚至达到 99%,在识别速度、准确率上均超过医生诊断。

3) 辅助诊断

诊断是医疗中的一个核心环节,诊断依赖患者体征、患者描述与检查数据,其中体征与检查数据是主要判断依据,在这个过程中,医生获取、解读信息的时间比较长,同时也存在误诊的可能性。AI 辅助诊断工具本质上就是为了解决上述问题。

AI 实现辅助诊断分为两大类:一类是由专家基于个人经验搭建知识库。这种相对比较精确,但是覆盖面小,只适合特定领域的辅助诊断。另一类是使用机器学习算法,从医院和网上抽取数据,构造一个更为全面的知识图谱。但是受限于没有数据规划,以及当前技术水平,在精确度上有所不足。

例如,在互联网医疗的背景下我国基于医疗知识图谱、自然语言理解技术研发了问诊机器人,可以基于用户简单的输入,一步一步地追问详细的症状信息,如发烧的时间、温度,发烧的温度变化;如果是疼痛,则追问疼痛的具体部位和其他相关症状发生的顺序,包括家庭、个人的病史。把相关信息结构化展现给医生,让病人能够选择更合适的科室,让医生在互联网环境下一目了然地看到病人所有的相关信息。

4. 休闲娱乐

人工智能技术已经进入娱乐圈与电竞界了,这里介绍几个典型的应用技术。

1) 深度合成

深度合成技术,其实就是借助可以自主学习的深度学习算法模型来实现的,其主要使用的两项技术是自动编码器和生成式对抗网络(generative adversarial network, GAN)。前者用于训练数据的合成,后者由生成器和鉴别器组成。一个用来进行新数据的生成,另一个用来对其进行鉴别,经过二者无数次的对抗,最终生成“以假乱真”的合成数据。视频换脸、语音合成、影像修复、虚拟数字人等越来越频繁地出现在社交娱乐、影视制作、教育、广告营销等领域,发展出多元化的商业应用。这些应用的背后,是深度合成技术的广泛应用,如图 1-6 所示。



图 1-6 深度合成技术在影视音乐方面的应用

电影修复是将年代久远的胶片电影重新复制到数字载体上,通过修复、降噪、补光、调色

等技术处理,还原和优化影片原貌的过程。在过去,老电影主要采用胶片作为拍摄和存储的介质。早期胶片的材料是一种比纸更易燃的硝酸片基,后来由醋酸片基、涤纶片基取而代之,但不管是哪种片基,在常温状态下都很难保存,温度、湿度以及搬运移动、使用、播映等外部因素很容易造成胶片损伤。正常情况下,一位熟练的修复师一天最多可以修复 200 帧画面,但如果胶片保存不佳,脏污、裂痕、变色等问题严重,修复师一天可能只能修复 1s(24 帧)。再者,如果画面涉及夜戏、雨戏、烟雾戏、特效戏等复杂场景,修复周期更是成倍增加。电影修复最大的难题是如何保留影片艺术风格和美感,4K 修复版电影曾一度因为颠覆胶片的美感而引发争议。一部老电影,经过漫长的岁月侵蚀,早已变得暗淡模糊,通过 AI 技术把它修复得光鲜漂亮并不难,难的是如何还原老电影的感觉。

2021 年 10 月 20 日,字节跳动公司的火山引擎联合西瓜视频共同发布了“经典中视频 4K 修复计划”,表示在未来 1 年内,将与央视动漫集团和上海美术电影制片厂合作,共同利用 4K 技术修复《舒克和贝塔》《西游记》等 100 部家喻户晓的经典动画。本次 4K 修复使用火山引擎智能处理产品中的部分技术能力,通过超分辨率、智能插帧、智能降噪、色彩增强等算法增强视频画质。其修复过程大致为:先利用视频降噪算法做前期处理,再进一步通过几类不同的超分辨率增强算法将画幅扩大到 4K 分辨率,并生成更精细的细节,最后通过插帧算法和 HDR 重制算法将一个原本充斥着噪声和压缩损伤问题的视频增强为一个主观画质舒适的 4K 60 帧 HDR(high dynamic range,高动态范围)节目。

2022 年 4 月 1 日,除了是愚人节,还有“哥哥”张国荣让人缅怀。当天晚上 8 点,一场尘封了 21 年的张国荣《热·情》演唱会在朋友圈刷屏,截至当晚 9 点,视频号播放量已过千万。因为是 21 年前的演唱会,腾讯云多媒体实验室团队对视频质量做了修复,从母带大概只有 720 像素×480 像素的分辨率,修复为接近 4K 的分辨率,分辨率是原来的 6 倍。AI 修复演唱会的时间并不长,只用了两天就完成了,但团队为达到“修旧如旧”的效果,在前期艺术算法微调上花了很长时间。

2022 年 5 月 8 日是一代歌后邓丽君逝世的 27 周年,当日,酷狗音乐阿波罗实验室用黑科技“复活”邓丽君的声音,将她的 16 首经典歌名串联成歌,打造出首支 AI 演唱单曲《没有寄出的信》。

数字领域(Digital Domain)公司利用数字人技术将邓丽君栩栩如生地呈现在舞台上,在江苏卫视 2022 跨年演唱会上,邓丽君“复活”并与周深合唱《大鱼》,让各位网友深受感动。

2) 游戏人工智能

人机游戏有着悠久的历史,已经成为主流的验证人工智能的关键技术。研究人员设计各类 AI 来挑战人类职业玩家。游戏分为 4 种典型类型:围棋棋盘游戏、纸牌游戏(如德州扑克、斗地主和麻将)、第一人称射击游戏、实时战略游戏(如星际争霸、Dota2 和王者荣耀)。下面按类型进行介绍。

在围棋棋盘游戏方面,1994 年国际跳棋程序 Chinook 打败了美国西洋跳棋棋王 Marion Tinsley,IBM 的“深蓝”(Deep Blue)在 1997 年击败国际象棋大师 Garry Kasparov,更为著名的是由谷歌(Google)旗下 DeepMind 公司开发的 AlphaGo 在 2016 年 3 月与围棋世界冠军、职业九段棋手李世石进行围棋人机大战,以 4:1 的总比分获胜。不仅如此,2017 年推出的进化版本 AlphaGo Zero 更是强大,经过短短 3 天的自我训练,AlphaGo Zero 就强势打败了此前战胜李世石的旧版 AlphaGo,战绩是 100:0。经过 40 天的自我训练,AlphaGo