

加速行业智能化白皮书

使能百模千态，赋能千行万业



本书编委会

主任委员： 王丽彪 华为企业 BG 副总裁、行业 ICT 解决方案开发管理部部长

主 编： 王 慷 华为行业 ICT 解决方案架构设计部部长

编委会委员（排名不分先后）：

魏 亮 中国信息通信研究院副院长

王爱华 中国信息通信研究院副总工程师、教授级高级工程师

邹德宝 赛迪顾问人工智能产业研究中心常务副总经理

何宝宏 中国信息通信研究院云计算与大数据研究所所长

许立言 北京大学城市与环境学院研究员、博士生导师

王静远 北京航空航天大学计算机学院教授

邓 攀 北京航空航天大学软件学院副研究员、博士生导师

边英杰 华为行业 ICT 解决方案规划与营销运作部部长

编委会秘书： 杨 虎

编写成员（排名不分先后）：

信通院： 曹 峰 李 荪 樊 威 唐甜甜

华 为： 周 倩 冯国杰

序言一

在这个数字化、智能化的时代，人工智能已经逐渐成为推动科技发展的核心驱动力。人工智能技术越来越成为面向未来、开拓创新的重要工具和手段。其中，基于知识驱动的第一代人工智能利用知识、算法和算力三个要素构建 AI；基于数据驱动的第二代人工智能利用数据、算法和算力三个要素构建 AI。由于第一、二代 AI 只是从一个侧面模拟人类的智能行为，因此存在各自的局限性，很难触及人类真正的智能。而第三代人工智能，则是对知识驱动和数据驱动人工智能的融合，利用知识、数据、算法和算力四个要素，构建了新的可解释和鲁棒的 AI 理论与方法，发展安全、可信、可靠和可扩展的 AI 技术。第三代人工智能是发展数字经济的关键，是数字经济未来发展的新灯塔和新航道。

知识和数据双轮驱动下的第三代人工智能技术正在催生人工智能产业的迭代升级，以大模型为代表的第三代人工智能，它通过文本的语义向量表示和转换器的多头注意机制，能够对文本的语义，即文本中所包含的知识进行自监督学习，从而在聊天（对话或自然语言处理）上向通用人工智能迈进一步，这是人工智能的重大突破，已经成为新一轮科技革命和产业变革的核心驱动力，助力中国经济实现高质量发展，深刻影响人民生活和社会进步。本书的典型特色，就是通过一些具体的场景应用和实践案例，把第三代人工智能赋能千行万业的作用进行了具象化，主要体现在以下几个方面：

一是第三代人工智能技术聚焦于企业和政府服务，涉及零售营销、金融、交通、医疗保健、教育、制造业、网络安全等方面。例如，在实践案例方面，中国国家气象中心与华为合作的气象大

模型，通过海量数据和算力保障充分发挥大模型算法的作用，使得中长期气象预报精度首次超过传统数值方法，速度提升 10000 倍以上。

二是第三代人工智能技术催生了很多新兴产业，这些新兴产业对于国民经济、国防、国民社会的发展至关重要。例如，在场景方面，AI 技术使能城市安全。基于 AI 技术构建的智能预测模型和智能决策模型等为城市感知体系的建设带来了新的能力。通过视觉大模型能够将之前的单场景感知增强至泛化多场景安全风险识别，面对城市安全等场景提升感知数据的通用分析能力，持续推动城市感知体系的创新和升级，为构建更智慧、更安全的城市环境注入了强大的动力。

三是第三代人工智能技术积极推动传统产业转型升级，运用大量信息技术和数字化手段，加快推进产业智能化，完善产业链数字形态，极大地提升生产效率和产品质量。例如，本书介绍了山东能源集团的智能化转型的案例。山东能源集团创新性地使用大模型和云边端分层分级架构体系，通过 AI 与业务流程深度结合、资深业务人员的深度参与，以及应用与 AI 算法协同等机制，成功解决了能源行业人工智能应用的难题。

第三代人工智能技术的发展还存在理论基础薄弱的问题，属于探索“无人区”，发展目标和发展道路不明晰，这也是人工智能产业和信息产业最大的不同。我们还欣喜的看到，本书提出了利用人工智能技术推动行业智能化的一个系统工程的理论架构。书中提到的“分层开放、体系协同、敏捷高效、安全可信”的行业智能化参考架构，对人工智能产业生态的构建会起

到重要作用，促进产业的有机更新、迭代升级，带动千行万业智能化，从而加快人工智能产业快速发展。

未来，全球经济要实现高质量发展，必须大力推动人工智能持续赋能行业智能化转型。通过本书的介绍，我们看到中国的人工智能技术和产业已经取得了长足的发展，在各个领域进行了大量的探索和实践，有诸多成功案例，可以为全球走向智能化提供一些成功的经验和实践范式。让我们期待人工智能技术的持续创新和改善，全球的人工智能产业及应用能在下一个十年蓬勃健康地发展！人工智能正在迅速发展，智能世界加速到来。

最后，我想说，人工智能的魅力就在于人工智能的研究永远在路上，需要的是坚持不懈与持之以恒。希望全社会、各行业能够更加积极的拥抱人工智能技术，共同迎接一个充满智能与创新的新时代。

中国科学院院士、清华大学教授、清华大学人工智能研究院名誉院长

A handwritten signature in black ink, appearing to be the name 'Li Xiang' (李翔), written in a cursive style.

序言二

犹如历史上蒸汽机、电力、计算机和互联网等通用技术一样，近 20 年来，人工智能正以史无前例的速度和深度改变着人类社会和经济，为释放人类创造力和促进经济增长提供了巨大的机会。人工智能是驱动新一轮科技和产业变革的重要动力源泉。人工智能的发展不但已从过去的学术牵引迅速转化为需求牵引，其基础和目标也在发生变化。人工智能技术在大数据智能、群体智能、跨媒体智能、人机混合增强智能、自主智能系统五大发展方向的重要性和影响力已系统展现。在规划及产业的推动下，这五个方向和 5G、工业互联网、区块链一起正在形成更广泛的新技术、新产品、新业态、新产业，使得制造过程更智能，供需匹配更优化，专业分工更精准，国际物流更流畅，从而引发经济结构的重大变革，带动社会生产力的整体跃升。另外，人工智能也正在赋能智慧电力、智慧交通、智能城市、智慧医疗、智慧农业等方向，带动各行各业从传统发展模式向智能化转型。总之，人工智能正不断重新定义人们的生产、生活的方方面面，同时也为我们带来了前所未有的机遇和挑战。

Chat GPT 的问世使 AI 又前进一大步。数据、算力、算法曾是人工智能发展的三大核心要素，现在开始转向大的数据、模型、知识、用户等四大要素。其中，数据是人工智能算法的“燃料”，融入知识的大模型是人工智能的基础设施，大模型的广泛使用则是人工智能系统进化的推动力量。从传统作坊式小模型向百亿千亿参数大模型的演进中，参数越多，训练的大数据越广泛、通用效果就越明显，越似乎人脑，但对算力要求也越高，杜绝偏差的难度也越大。人工智能迭代发展过程中，顶层设计要考虑到行业中数据的相容与特色、知识的建构和发展，

算力设施的同步演进，形成合力，支撑人工智能产业升级换代。

在这场人工智能的变革浪潮中，如何把握人工智能技术的发展趋势，将其应用于实际行业场景，以实现更高效率、更低成本、更广覆盖度地赋能行业智能化，已经成为社会各界关注的焦点。行业智能化转型过程中遇到的其中一个关键挑战是在各行业与 AI 之间的知识沟通，培养两栖人才。本书通过一个通用的系统工程框架比较全面地解析了该问题的解决之道，来指导行业智能化转型这个复杂的系统工程。通用的智能化系统框架像人体一样，有大脑、五官、经脉、血液、手脚、知识等，可感知，能学习，会思考，会进步。结合行业数据、知识的积累与融合，用户的体验与反馈，才能更好地支撑 AI 在行业中的发展。

人工智能将触发广泛的行业变革。未来十年，AI 的主战场正是在各行各业。我们不但要研究语言模型、图像模型、视频模型等等基础大模型，还要进一步创建行业知识与数据集，训练各行业的垂直模型，推动数据和知识双轮驱动的人工智能。数据和知识的结合将让人工智能走向更深入、更专业、更广泛。另外还需要加强安全可信、政策标准等方面的投入，以更全面、更有效的力度推进行业智能化的发展。

本书列举了人工智能技术使能不同的企业生产场景和民生场景。例如，在 AI 使能企业生产部分，本书提出了大模型支撑不同业务场景：视觉大模型应用于安监生产；语言大模型应用于经营管理；多模态大模型应用于决策指挥。在 AI 使能民生部分，通过使用人工智能技术，提高医学影像诊断的精准性、标准化和自动化。

书中对智慧煤矿、智慧钢铁、智慧港口、智慧城市交通、智慧航空、智能铁路、智慧城轨、智慧油气、智能制造、智慧电力等企业生产场景，以及城市治理、气象预报、智慧影像、精准医疗、智慧水务等民生场景的深入分析，还列举了大量生动的案例，这些案例分布在矿山、电力、制造和大企业、机场和轨道交通、公路水运口岸、油气、金融、智慧城市、政府及公共事业等众多领域。充分体现了人工智能在各行各业应用中具有巨大潜力和广阔空间。

总的来说，本书总结了华为过去几十年的技术创新研究和对行业智能化的实际应用效果，通过独特的视角和深入的思考，考察了人工智能技术在行业智能化中的关键作用，展示了如何将人工智能技术与传统产业深度融合，推动产业升级和转型，为经济智能化转型提供了有益的经验与借鉴。

当前，中国正以“万水千山只等闲”之势，生机勃勃地前进在行业智能化转型的大道上，产业界、学术界、行业用户正在一起合力构建一个万物互联的智能世界，打造无所不及的 AI。智能化的未来，是全人类共同的未来，每个国家都有权利和需求参与到智能化的发展进程中来，共同推动智能化技术的应用和创新，带动全球经济和社会走向一个高质量、高水平的快速演进期，以造福全人类。

中国工程院院士、浙江大学教授
国家新一代人工智能战略咨询委员会主任

Handwritten signature in black ink, reading '潘圣军' (Pan Shengjun).

序言三

1760年到1840年的第一次工业革命，主要技术手段是煤炭、蒸汽机，将人类带入了以机械化为特征的蒸汽机时代。19世纪末到20世纪初的第二次工业革命的主要技术手段是电力、石油，将人类带入了电气化时代。20世纪60年代开始的第三次工业革命主要技术手段是计算机、互联网，将人类带入了自动化或网络化时代。21世纪初的第四次工业革命则以大数据和人工智能技术为核心，以互联网承载的新技术融合为典型特征。相比前三次工业革命为人类社会带来的进步，第四次工业革命将人类带入了更高层次的智能化时代。人工智能技术不断演进，成为第四次工业革命的关键新兴技术，成为当前最具颠覆性的技术之一，是行业转型升级的重要驱动力。

人工智能作为一项战略性技术，已成为世界多国政府科技投入的聚焦点和产业发展的发力点。全球170多个国家相继出台人工智能相关国家战略和规划文件，加速全球人工智能产业发展落地。具体而言，美国将人工智能提到“未来产业”和“未来技术”领域的高度，不断巩固和提升美国在人工智能领域的全球竞争力；欧盟全面重塑数字时代全球影响力，其中将推动人工智能发展列为重要的工作；英国旨在使英国成为人工智能领域的全球超级大国；日本致力于推动人工智能领域的创新创造计划，全面建设数字化政府；新加坡要成为研发和部署有影响力的人工智能解决方案的先行者；基于当前复杂多变的国际形势，中国一方面要加强人工智能基础核心技术创新研究，培育自主创新生态体系，另一方面要推进人工智能与传统产业的融合，赋能我国产业数字化、智能化高质量发展。

算力、数据、算法已经构成了目前实现人工智能的三要素，并且缺一不可。人工智能算力是算力基础设施的重要组成部分，是中国新基建和“东数西算”工程的核心任务抓手。预计到2025年，中国的AI算力总量将超过1800EFlops，占总算力的比重将超过85%，2030年全球AI算力将增长500倍。中国已经在20多个城市陆续启动了人工智能计算中心的建设，以普惠算力带动当地人工智能产业快速发展。多年来华为聚焦鲲鹏、昇腾处理器技术，发展欧拉操作系统、高斯数据库、昇思AI开发框架等基础软件生态，通过软硬件协同、架构创新、系统性创新，保持算力基础设施的先进性，为行业数字化构筑安全、绿色、可持续发展算力底座。

人工智能产业的发展必然带来海量数据安全汇聚和流通的需求，大带宽、低时延的网络能力是发挥算力性能的基础。网络能力需求体现在数据中心内、数据中心间以及数据中心跟终端用户之间不同层面的需求。中国正在启动400G全光网和IPv6+网络建设以及从5G往5G-A传输网络的演进工作，旨在通过大带宽、低时延高性能网络，支撑海量数据的实时安全交互。通过全方位的网络能力建设和升级，为人工智能数据流动保驾护航。

人工智能技术的应用，是发挥基础设施价值的最后一公里。海量通用数据基础上进行预先训练形成的基础大模型，大幅提升了人工智能的泛化性、通用性、实用性。基础大模型要结合行业数据进行更有针对性的训练和优化，沉淀行业数据、知识、特征形成行业大模型，赋能千行万业智能化转型。

本书中华为联合行业全面地总结了人工智能基础设施建设以及行业智能化转型的实践经验，精选了一些 AI 使能企业生产、使能民生、加速行业智能化转型方面的典型案例进行分析，展示了图像检测、视频检索，预测决策类，自然语言处理 3 类应用场景的巨大潜力，为世界各国推动行业智能化转型落地提供了更多的思路、方法和借鉴，为全球人工智能技术发展和进步贡献更多智慧和力量。

人工智能技术将成为行业智能化的主驱动，推动各行各业实现智能化转型和发展。智能化将成为全人类共同的未来，不是个别国家的特权，不仅是因为它能够带来巨大的经济和社会效益，更因为它能够让人类的生活更加便捷、高效和舒适。全球各国可以结合各自的实际情况，相互学习和借鉴，加快 AI 算力基础设施的构建，并通过人工智能领域人才的培养、提供政策保障、制定行业标准，助推 AI 技术高质量发展，共同探索和创造更加美好的未来。

中国工程院院士、清华大学计算机科学与技术系教授

Handwritten signature in black ink, reading '郑伟' (Zheng Wei).

序言四

回望人类社会发展史，过去几千年里，社会生产力基本保持在同一水平线上。然而，自工业革命以来，这条曲线开始缓缓上升，并且变得越来越陡峭。人工智能被誉为是 21 世纪社会生产力最为重要的赋能技术，正以惊人的速度渗透进各行各业，推动一场新的生产力与创造力革命，变革未来的产业模式。凯文·凯利预测，在未来的 100 年里，人工智能将超越任何一种人工力量。变革已成为一股无法阻挡的力量，将人类引领到了一个前所未有的时代。人工智能带动数字世界和物理世界无缝融合，从生活到生产，从个人到行业，正日益广泛和深刻地影响人类社会，驱动产业转型升级。ChatGPT 和大模型的出现使得人工智能发展进一步加速，世界各国正在进入百模千态时代，人工智能与千行万业的深度融合成为热点与焦点，加速行业智能化成为未来人工智能发展的主旋律。

古人云：日就月将，学有缉熙于光明。华为始终秉持“把数字世界带入每个人、每个家庭、每个组织，构建万物互联的智能世界”愿景，基于对未来趋势的理解和把握，在 ICT（信息和通信技术）领域一直走在前沿，不断引领产业发展。在 2005 年，首先提出网络时代全面向“ALL IP”发展演进；在 2011 年，又一次提出数字化时代全面向“All Cloud”发展演进；在 2021 年，首次刊发《智能世界 2030》揭示了未来十年 ICT 技术广泛应用发展趋势。今天，我们在此提出智能时代全面向“All Intelligence”发展演进，通过人工智能领域的理论创新、架构创新、工程创新、产品创新、组合创新和商业模式创新，华为将使能百模千态、赋能千行万业，加速行业智能化发展，助力行业重塑与产业升级。据预测，2030 年全球人工智能市场规模将超过 20 万亿美元，然而在

行业的智能化落地中仍面临以下四个关键挑战：

首先，人工智能的算力挑战，大模型应用对于算力基础设施的规模提出了更高的要求，企业传统基础设施面临算力资源不足的挑战。大模型需要大算力，其训练时长与模型的参数量、训练数据量成正比。参照业界分析，能达到可接受的训练时长，需要百亿参数百卡规模，千亿参数千卡规模，万亿参数万卡规模，这对算力资源的规模提出了较高的要求。

第二，人工智能的数据挑战，各个行业都有各自长期且专业的积累，涉及物理、化学、生物、地质等多维知识表达，为了在不同行业落地应用，大模型必须结合行业知识、专有数据，完成从通用到专业的转变。获取海量高质量专有数据是一项艰巨的任务，如何智能感知、实时上传和高效存取海量生产数据，不仅需要解决设备连接的兼容性问题，还要确保实时性和高可靠性。在数据预处理、训练和推理阶段，同样面临读取性能问题、数据丢失问题以及成本效率问题等一系列挑战。行业数据是企业的核心知识资产，涉及到知识产权等问题。如何合法地获取和整合数据，并确保端到端的数据安全，满足隐私保护要求也是一项挑战。

第三，人工智能技术开发的挑战，在行业模型及应用开发的过程中，如何简化开发流程，提高开发效率，变革开发模式，高效打通数据链路，引入自动化机制，加强应用安全性和可靠性，都是大模型应用开发中面临的诸多难题。要解决这些难题，关键在于打造一个通用可靠的人工智能应用开发平台来赋能行业开发者。

最后，人工智能落地应用的挑战，由于不同规模、

不同能力的企业对大模型的建设需求不尽相同，因此需要构建不同层级的模型并提供相应的资源和部署能力，如总部层面集中建设大规模训练集群，区域层面建设规模训练平台、训推一体平台，边缘侧部署推理能力。服务于行业，除了技术问题，人工智能还需要解决人才储备、技术生态，以及法规政策等一系列挑战。

过去三年，华为成立行业军团，深入行业、深入场景，纵向缩短管理链条，更好的响应客户智能化需求；横向快速整合研发资源，全力支持千行万业的智能化转型发展。行业军团基于华为创新的智能化 ICT 基础设施和云平台，广泛联合业内解决方案伙伴，打造领先的产品和解决方案适配行业智能化场景，为行业智能化实践添砖加瓦、探索前行。比如，华为云盘古气象大模型，正被天气预报中心用来预测未来 10 天全球天气。该模型使用全球 39 年的天气数据进行训练，仅用 1.4 秒就完成了全球 24 小时天气预测，比传统的天气预报方法快 1 万倍；可以借助它进行台风路径的预报，保持了极高的精准度。山东能源集团依托盘古大模型建设人工智能训练中心，构建起全方位人工智能运行体系，探索和发掘煤矿生产领域全场景的人工智能应用，将一套可复用的算法模型流水线应用到各种作业场景，通过“人工智能大规模下矿”实现了矿山作业的本质安全和精简高效。目前行业军团已经面向金融、制造、电力、矿山、机场轨道、公路水运口岸、城市、教育医疗等 20 多个行业打造了 200 多个行业智能化解决方案，并在一系列智能化项目中产生了实际效果。

千行万业正在积极拥抱人工智能，把行业知识、创新升级与大模型能力相结合，以此改变传统

行业生产作业、组织方式。人工智能的发展与使用将成为全球各个国家行业转型升级的关键一环，助力各个国家在人工智能时代不断取得发展，华为将聚焦以下三方面，持续助力：

第一，创新引领。持续加强人工智能基础设施的创新投入，提供灵活的智能算力供给模式，高效可信的人工智能开发体系，使各层级大模型更易于部署，应用速度更快，推进 AI 应用走深向实，助力行业、企业实现场景创新。

第二，生态开放。算力开放，支持百模千态；感知开放，实现万物智联；模型开放，匹配千行万业。与各行业合作伙伴共同构建人工智能生态圈，探索更多的人工智能行业场景应用，携手企业、研究机构、学术机构等共筑安全可信的人工智能生态体系。

第三，人才培养。人才是企业发展的核心力量，支持各个行业、各个企业培养和吸引人工智能人才，打造一支高水平的人工智能研发团队，为人才提供广阔的发展空间。

结合华为行业智能化实践，以及面向智能世界 2030 的展望，我们与业界专家学者进行了万场以上座谈研讨，凝聚了各方智慧与经验，输出《加速行业智能化》白皮书。希望能够通过本书的论述和案例为行业智能化实施落地提供参考，加速拓展人工智能技术在行业中的应用。

百舸争流，奋楫者先。智能时代的大潮正奔涌而来，让我们同舟共进，引领时代，使能百模千态、赋能千行万业、加速行业智能化！

华为公司常务董事、ICT 基础设施业务管理委员会主任、企业 BG 总裁



目录

1

加速行业 智能化转型

| | |
|----------------------------|----|
| 第一章 | 3 |
| 加速行业智能化的 价值与挑战 | |
| 第二章 | 10 |
| 实现智能化转型 所需的参考架构 | |

2

加速智能化 带来新价值

| | |
|------------------|----|
| 第三章 | 33 |
| 智能化使能企业生产 | |
| 第四章 | 80 |
| 智能化使能民生 | |

3

加速智能化 转型案例

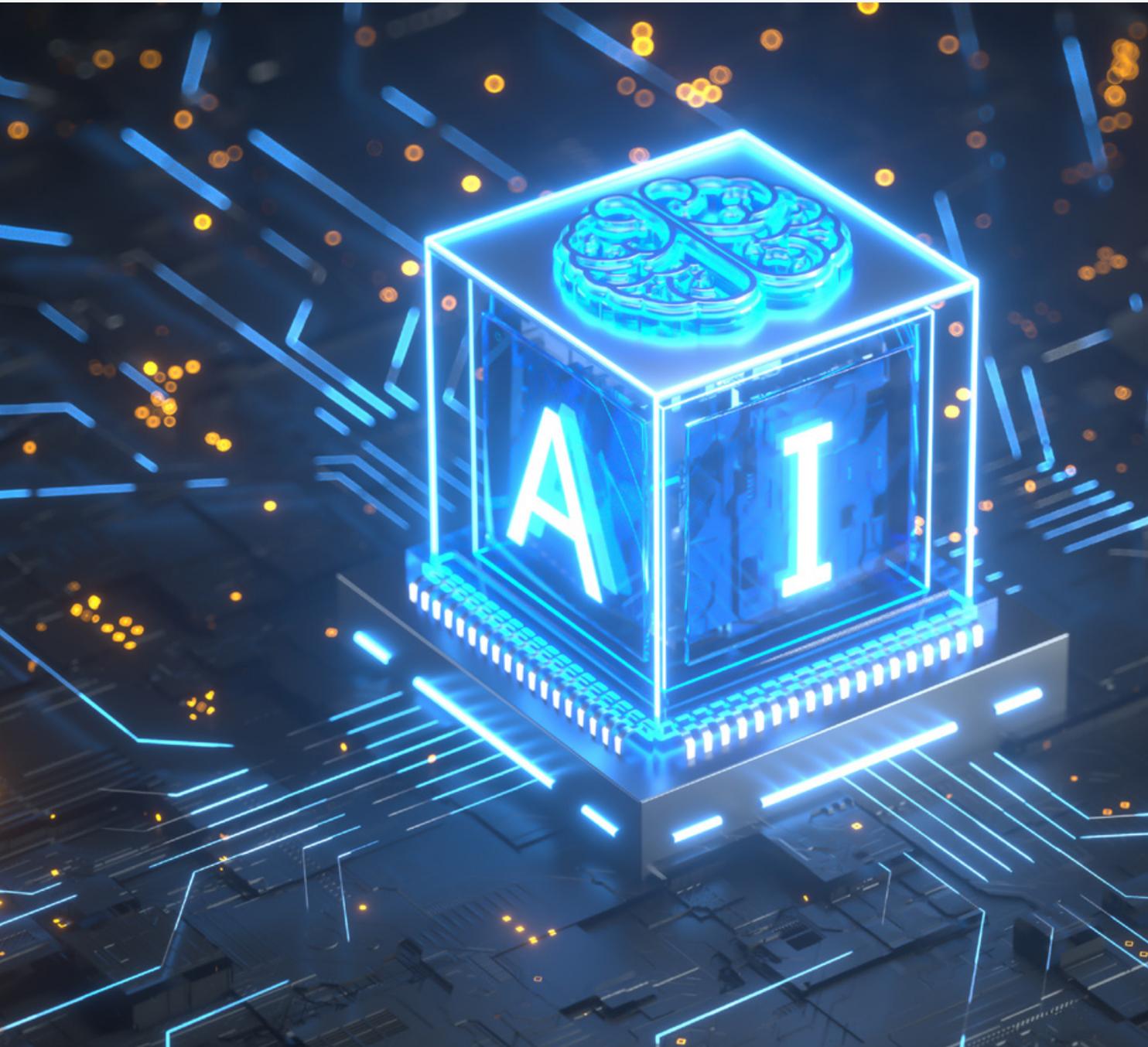
| | |
|-----------------|-----|
| 第五章 | 97 |
| 制造和大企业 | |
| 第六章 | 105 |
| 矿山及重工业制造 | |
| 第七章 | 120 |
| 金融 | |
| 第八章 | 127 |
| 电力 | |
| 第九章 | 133 |
| 油气 | |
| 第十章 | 142 |
| 机场和轨道交通 | |

| | |
|----------------|-----|
| 第十一章 | 152 |
| 公路水运口岸 | |
| 第十二章 | 169 |
| 智慧城市 | |
| 第十三章 | 174 |
| 政府及公共事业 | |

4

行业智能化 转型展望

| | |
|--------------|-----|
| 第十四章 | 193 |
| 总结与展望 | |





加速行业智 能化转型

第一篇



第一章 加速行业智能化的 价值与挑战

人类社会经历了从农业社会到工业社会到信息社会再到智能社会的变迁，历时几千年。蒸汽机的发明，让交通方式实现了从马车到汽车的跃迁，推动人类社会从农业时代迈入工业时代；计算机的发明，深刻地改变了生产生活的各个方面，推动人类社会从工业时代迈入信息时代；而人工智能技术的快速发展和广泛应用，正帮助人类加速跨入智能时代，这是一个波澜壮阔的史诗进程，将开启一个与大航海时代、工业革命时代、宇航时代等具有同样历史地位的新时代。

人工智能的概念提出于 1956 年，历时近 70 年时间周期的技术演进，无论技术本身还是应用均出现了日新月异的变化。据预测，截至 2030 年，人工智能的三大核心要素均将迎来创新突破：在算法方面，大模型将在应用侧持续落地、改变产业发展生态；在数据方面，人类将迎来 YB 数据时代，数据量是 2020 年的 23 倍、全球联接总数达 2000 亿；在算力层面，全

全球通用计算算力将达到 3.3ZFLOPS(FP32)，AI 计算算力将超过 105 ZFLOPS(FP16)，增长 500 倍。人工智能正在开启继互联网、物联网、大数据之后的第四次科技浪潮。

全球 170 多个国家相继出台人工智能相关战略和规划文件，将政策重点聚焦在加强技术投资和人才培养、促进开放合作以及完善监管和标准建设上，全球人工智能产业发展进入加速落地阶段。具体而言，美国将人工智能提到“未来产业”和“未来技术”领域的高度，不断巩固和提升美国在人工智能领域的全球竞争力；中国一方面要加强人工智能基础核心技术创新研究，培育创新的生态体系，另一方面要推进人工智能与传统产业的融合，赋能中国产业数字化、智能化高质量发展；欧盟全面重塑数字时代全球影响力，其中将推动人工智能发展列为重要的工作；英国旨在使英国成为人工智能领域的全球超级大国；日本致力于推动人工智能领域的创新创造计划，全面建设数字化政府；新加坡要成为研发和部署有影响力的人工智能解决方案的先行者。

华为一直致力于“把数字世界带入每个人、每个家庭、每个组织，构建万物互联的智能世界”。万物感知、万物互联和万物智能是智能世界的三大特征。要实现这三大特征，需要大量的

ICT 创新技术，打造一条 All IP、All Cloud、All Intelligence 的创新之路，并以共生共荣的生态系统作为支撑，使能整个智能世界。2005 年华为提出网络时代的“All IP”，在 2011 年华为提出数字化时代的“All Cloud”，2023 年提出智能时代的“All Intelligence”。华为持续以产品创新、系统架构创新和商业模式创新的新理念，深耕联接、计算、存储、感知、云、大模型等领域相关技术的研究创新，着眼于个人、家庭、组织在数字化转型到智能化转型过程中所遇到的现实挑战，帮助他们解决问题，实现从 All IP、All Cloud 到 All Intelligence 的产业升级。

人工智能正从感知理解走向认知智能，带动数字世界和物理世界无缝融合，从生活到生产、从个人到行业、从 C 端到 B 端，正日益广泛和深刻地影响人类社会，驱动产业转型升级。据预测，2030 年全球人工智能市场规模将超过 20 万亿美元，2030 年中国人工智能核心产业规模将超过 4 万亿美元，目前仍有较大的发展潜力和空间。人工智能无所不及，帮助人类获得超越自我的能力，成为科学家的显微镜与望远镜，让我们的认知跨越微小的夸克到广袤的宇宙，千行万业从数字化走向智能化。

让我们拥抱 AI 新时代，拥抱行业智能化转型，拥抱美好新未来！

▶ 1.1 智能社会，未来已来

1.1.1 智能世界正在加速而来

20 世纪，人工智能的发展主要经历了五个阶段（1）萌芽期（1956 年 - 20 世纪 60 年代初）：在这个阶段，以麦卡赛、明斯基、罗切斯特和申农为代表的科学家团队共同研究了机器模拟

智能的相关问题，并于 1956 年达特茅斯会议上正式提出人工智能概念。（2）启动期（20 世纪 60 年代）：上世纪 60 年代迎来人工智能的第一个黄金发展期，该阶段的人工智能主要以语言翻译、证明等研究为主，在这个阶段取得了机器定理证明、跳棋程序等一系列标准性

成果。(3) 瓶颈期(20世纪70年代初): 上世纪70年代, 经过科学家深入的研究, 科学家提出了一系列不切合实际的研发目标, 尤其是对机器模仿人类思维的错误认识, 导致人工智能发展进入低谷期; (4) 突破期(20世纪70年代初到90年代中期): 以专家系统为代表的技术突破, 推动人工智能加快应用于医疗、化学、地质等各个领域, 人工智能技术在商业领域取得了巨大的成果; (5) 平稳期(20世纪90年代以来): 上世纪90年代以来, 随着互联网技术的逐渐普及, 加速了人工智能的创新突破, 出现了深蓝计算机战胜国际象棋冠军、“智慧地球”提出等一系列标志性事件, 促进人工智能进一步与应用相结合。

进入21世纪, 人工智能正在开创下一个黄金阶段。2016年, Google AlphaGo 击败顶级围棋选手, 震惊全球, 开启了新一轮 AI 热潮。2022年11月30日, ChatGPT 横空出世, 人工智能的发展进入全新阶段。2023年3月15日, OpenAI 又推出多模态大模型 GPT-4, 在生成质量、使用性能和模型安全合规等多个领域评分均领先于现有主流模型, 被誉为“史上最强”大模型。Meta 公司于2023年2月和7月先后推出开源大模型 LLaMA 和 LLaMA2, LLaMA2 在数据质量、训练技术、能力评估、安全训练和责任的发布等方面有了显著的进步, 并首次放开商业用途许可要求。根据 GitHub 统计, 中国也已经推出超过 100 个大模型, 包括华为盘古大模型、讯飞星火认知大模型、智谱 AI-ChatGLM 大模型、百度文心大模型等。

大模型应用一般分为两个阶段: 预训练和微调, 大模型预先在海量通用数据上进行训练, 数据、知识得到了高效积累和继承, 从而大幅提升了人工智能的泛化性、通用性、实用性。在实际

处理下游任务时再通过小规模数据进行微调训练, 就能达到传统小模型的效果。大模型的出现, 减少了行业用户训练模型研发成本, 降低了 AI 落地应用的门槛, 并且上线部署过程大幅简化, 迅速驱动 AI 在千行万业的应用普及。随着大模型从“炫技”到“落地生根”, 全社会也认识到 AI 技术真正地来到了我们身边, AI 技术的“iPhone 时刻”到了, 各行业开启了从数字化到智能化的升级, 行业智能化的“iPhone 时刻”也随之到来。作为继工业社会、信息社会后更高级的社会形态, 智能社会必将迎来广泛、深刻甚至根本性的变革。

1.1.2 智能化正在改变人类的生活和工作

AI 已经成为了人们生活中不可或缺的一部分, 它正在改变着我们的生活方式和工作方式。

日常出行中, 人工智能翻译支撑跨语言、跨文化的高效沟通, 通过拍照获取景点的历史文化和背景故事。在娱乐中, 通过虚拟现实(VR)与增强现实(AR)等 AI 技术, 让用户沉浸在虚拟的现实生活中, 或者将虚拟元素融入真实环境中, 为娱乐带来全新的体验维度。

办公场景中, 未来撰文、翻译、制图、代码核查等工作一半以上可由人工智能完成。人工智能可以通过高效运算, 接管一些重复性工作, 把人类从忙碌而繁重的日常工作中解放出来, 让人类节省最宝贵的时间资源, 去做更多振奋人心、富有挑战性的工作, 如按其所长贡献创造力、策略思维等。

在教育场景中, AI 的出现和持续演进正在重构传统课堂教学, 改变学校形态、教学方式和学习方式。如随着 AI 持续改变教育的方式, 越

来越多的学生将更愿意参加线上数字课程的学习；借助人工智能技术为学习者推荐个性化的学习资源，实现学习者的个性化学习等。

另外，随着人工智能技术的不断迭代升级，未来 AI 技术将应用于更多领域，如基于多模态大模型的 AI 个人助理将极大的便利人们的生产生活，数字分身成为跨界人工智能、自媒体、科普多个领域的里程碑式数字 IP，在越来越多的垂直领域细分赛道的应用场景中出现等。

1.1.3 智能化正在持续赋能千行万业

人工智能是引领未来的战略性技术，全球主要国家及地区都把发展人工智能作为提升国家竞争力、推动国家经济增长的重大战略。通过应用牵引推动人工智能技术落地成为各国共识。美国引导人工智能技术在行业领域的创新和融合应用。2021 年 7 月，美国国家科学基金会联合多个部门和知名企业等，新成立 11 个国家人工智能研究机构，涵盖了人机交互、人工智能优化、动态系统、增强学习等方向，研究项目更是涵盖了建筑、医疗、生物、地质、电气、教育、能源等多个领域。中国“十四五”规划纲要明确大力发展人工智能产业，打造人工智能产业集群以及深入赋能传统行业成为重点。英国支持人工智能产业化，启动人工智能办公室和英国研究与创新局联合计划等，确保人工智能惠及所有行业和地区，促进人工智能的广泛应用。日本将基础设施建设和人工智能应用作为重点，重点强调了跨行业的数据传输平台，全面推动人工智能在医疗、农业、交通物流、智慧城市、制造业等各个行业开展应用。

同时世界各国也高度重视人工智能标准化工作，规范人工智能落地应用，出台战略加强标准化布局，支撑产业生态发展。ISO/IEC JTC 1 SC 42 及 IEEE 已成为国际 AI 标准的主要

供给方和参考源，影响力正向全球辐射。重点在人工智能基础共性、关键通用技术、人工智能可信及伦理方面开展标准研制工作。2020 年 3 月 18 日，中国国家标准化管理委员会批复成立全国信标委人工智能分技术委员会，主要负责人工智能基础、技术、风险管理、可信、治理、产品及应用等人工智能领域国家标准制订工作，对口国际 ISO/IEC JTC 1/SC 42，以标准化手段，分类分级分步骤推动大模型评测、算力、算法、数据和治理等领域的技术和应用，带动和引领人工智能的健康、可持续发展。

在全球各国高度关注下，人工智能从实验室加速走向应用场，推动 AI 技术应用到智能产品的开发，服务模式的创新，产业升级，赋能多行业智能化转型。

在气象预报行业，AI 让人们重新审视气象预报模型的未来，受到国内外气象组织的高度关注。《Nature》正刊发表了华为盘古气象大模型研究成果，该模型创建了首个精度超过传统数值预报方法的 AI 模型，速度相比传统数值预报提速 10000 倍以上，能够提供全球气象秒级预报。在今年的“古超”和“杜苏芮”等台风预测应用中，盘古气象大模型准确预测了台风路径，这将对气象导致的防灾减灾，保障民生安全起到非常重要的作用。

在医疗健康行业，AI 的应用逐步走向成熟，在辅助医疗、影像质控、知识问答等方面发挥着极关键的作用，尤其是在药物研发领域，AI 打破公认的“双十定律”，将研发时间由数年缩短至一个月，研发成本降低了 70%，这项突破，直接推动药物研发实现从“马拉松”到“加速跑”，开启了药物研发领域前沿技术革命的新征程。

1.1.4 社会运行方式正在悄然变化

人工智能改变了基础的生产力工具，中期来看会改变社会的生产关系，长期来看将促使整个社会生产力发生质的突破。人工智能将对消除社会数字鸿沟，实现全球包容性增长和可持续发展具有重要作用。

人与自然和谐相处是一个永恒的命题，而保护生物多样性是保护地球的一个重要手段。通过生物多样性的监测，我们可以更深入地了解身边的大自然。传统的生物多样性研究需要对数据收集的专用技术进行大量投资，并需要非常专业、经验丰富的生态保护科学家进行数据分析和洞察。相关科学文献的一项调查发现，58% 的生物多样性声学研究完全基于专家从原始数据中人工识别物种。如今，随着人工智能技术的应用，科学家能够处理的研究数量以及每项研究所能分析的数据量，都有了极大的飞跃。例如，RFCx 开发的 Arbimon 是一个开放和免费的平台，致力于为生物多样性声学监测和洞察提供端到端解决方案。该平台目前收集了 2,000 多种鸟类鸣叫样本，占全球鸟种的 20% 以上。Arbimon 团队联合华为和其他合作伙伴，基于不断增长的数据集，为新物种训练新的人工智能模型，并重新训练现有模型以提高性能。分析结果可帮助一线工作人员决策采取可行鸟类保护措施或行动。

人与自然的和谐相处也包括阻止外来物种入侵，保护当地物种免于被灭绝，防止破坏当地生活方式。在挪威贝勒沃格这个原本平静的城市正面临着外来物种入侵的威胁。驼背大马哈鱼来源于太平洋，并不是挪威的本地物种。驼背大马哈鱼比生活在大西洋的大西洋鲑更具侵略性，现在正快速占领河流，给大西洋鲑的繁殖带来了挑战。大西洋鲑属于挪威政府认定的

濒危物种，如果不采取措施，可能会遭到灭绝。2021 年 3 月，由当地猎人和渔民组成的贝勒沃格狩猎和垂钓协会（BJFF）与华为建立了合作关系，双方将共同保护斯托尔瓦河中的大西洋鲑。他们最初的目标是利用水下摄像机和人工智能来识别物种并统计鱼类的数量。该项目的第一阶段于 2021 年夏天启动。这一阶段聚焦开发算法，让计算机系统能够识别河流中的本地大西洋鲑和北极红点鲑，并记录不速之客驼背大马哈鱼。这一目标已经轻松达到。通过 6 月下旬到 9 月采集到的连续视频和数万图片，新开发的算法能够识别出 90% 以上的大西洋鲑和驼背大马哈鱼。后续可对通过的鱼进行分类，对大西洋鲑和驼背大马哈鱼分流处理。

关爱残障人士是包容社会的重要组成部分，通过数字化技术能帮助提高他们生活的舒适度和便利性。中国目前有 2700 多万听障人士，其中许多人使用手语作为其主要的交流形式。与其他手语语言一样，中国手语学习面临词汇更新慢、师资短缺、标准难统一等挑战，对此，千博信息基于华为昇思 MindSpore 和中国科学院自动化研究所紫东太初三模态大模型，带来全新的手语产品，基于 1.2 万 + 词汇、50 万 + 语法、70 万 + 语料，形成手语多模态模型，打造手语教考一体机，开创性地实现手语动作与视频、图片示意和文字说明联动，使得手语学习能够快速上手，一定程度上缓解了手语师资短缺问题。也能作为手语翻译机使用，帮助听障人士顺畅沟通、便捷生活。

▶ 1.2 行业智能化，多重挑战犹存

到 2030 年，人工智能有望颠覆制造、电力、能源、交通运输等关键产业，撬动难以限量的经济价值，若要把握这一机遇，须在多个维度进行能力建设。统筹点线面、系统推进的总体考虑下，重点围绕核心技术、基础软硬件、数据资源体系、标准规范和行业应用示范等进行部署。

1.2.1 传统算力基础设施难以匹配大模型创新需求

AI 快速发展并在多行业落地，呈现出复杂化、多元化和巨量化的趋势。不同的应用场景对算力的要求不同，要评判算力基础设施是否满足需求，需要企业根据特定的 AI 技术场景和需求，综合考虑算力基础设施的性能与灵活易用性。

大模型技术对于算力基础设施的规模提出了更高的要求，企业传统基础设施面临算力资源不足的挑战。AI 大模型需要大算力，其训练时长与模型的参数量、训练数据量成正比。根据业界论文的理论推算，端到端大模型的理论训练时间为 $8 * T * P / (n * X)$ 。其中 T 为训练数据的 token 数量，P 为模型参数量，n 为 AI 硬件卡数，X 为每块卡的有效算力。以 ChatGPT 为例，参数量为 175B（1750 亿）规模下，在预训练阶段，数据量 35000 亿，使用 8192 张卡，其训练时长为 49 天。同等条件下参数变多，计算量变大，按照业界的经验，能达到可接受的训练时长，需要百亿参数百卡规模，千亿参数千卡规模，万亿参数万卡规模。这对算力资源的规模提出了极高的要求。算力不足意味着无法处理庞大的模型和数据量，也无法有效支撑高质量的大模型技术创新。

1.2.2 基础大模型难以适应行业智能化需求

基础大模型 L0 的构建，需要顶尖人才和巨额资金的持续投入，百模千态将以行业模型的形式为主。技术门槛高，基础大模型的构建是复杂的端到端系统工程，是一个典型的复杂软件平台。资金投入大，GPT-4 训练成本约 6 千万美金，推理成本将至少是训练的 5~10 倍，达到数亿美金 / 年。

每个行业均有使用大模型的场景，行业用户及行业伙伴大多不具备从头开发大模型的能力，为了获得适配本行业的大模型，需要提供行业数据给基础大模型进行微调训练。

数据是行业用户的核心资产和竞争优势的源泉，行业用户部分关键敏感数据难以实现共享或者“出厂”，例如政务行业中涉及到城市发展、公共安全和个人隐私等方面的数据；金融行业中责权、债务关系相关的数据；制造业的资产明细、生产数据以及明确要求不可以出园区的数据等等，此时基础大模型难以适应行业智能化需求。

于是行业用户将行业非敏感数据提供给基础大模型供应商，形成行业大模型 L1，再结合场景数据在行业大模型 L1 基础上形成场景大模型 L2，以适应行业的需要。

1.2.3 数据供给难以满足大模型训练需求

数据将是构建大模型竞争力的核心要素，行业先锋都应该打造自身的数据飞轮。高价值、特定领域的工作流程，特别且必须依赖于丰富且高质量的专有数据集。

从全球来看，高质量数据都在变得更稀缺。行业用户重视数据集资产的构建和管理，但采集、存储和管理海量数据，形成优质数据集的能力仍不足。

从政府层面来看，当前缺少统筹共性数据集的建设服务，数据流通与共享机制不成熟，开放数据集“质”与“量”难保证，源头数据的治理不充分，导致数据质量不高、共享不足。

从企业内部来看，数据的实时采集受制于非数字化终端，数据的实时上传受制于低速网络，数据的实时分析受制于数据孤岛，行业数据难采、难传、难用，阻碍了行业智能化的进程。

1.2.4 法规监管难以保障智能化安全可靠需求

行业智能化后，行业信息化水平进一步提升，对信息系统的依赖加剧，减少人工干预后，被攻击入侵后损失严重，对安全可靠诉求更加强烈。行业智能化会导致信息系统更加复杂，构建安全可靠的难度剧增。

人工智能技术涉及的领域非常广泛，其潜在的风险和危害也不容忽视，比如个人隐私保护、脱敏数据使用、数据泄露等，这些都需要根据合适的法律进行规范和监管。人工智能技术的发展也与社会伦理、安全等问题存在联系，需要明确相关责任与义务等问题。因此需要更健全的人工智能法规和监管机制。

1.2.5 人才储备和生态难以支撑智能化转型需求

人才是人工智能竞争的关键，更是人工智能赋能行业智能化的有效支撑。预计到2025年，中国

ICT人才缺口将超过2000万，其中云、人工智能、大数据等新兴技术方面的人才尤其稀缺，此外既懂行业又懂技术的复合型人才也异常匮乏。人才储备不足问题成为限制行业智能化转型的瓶颈问题之一。

各国都要建立完善的人工智能产业生态，如中国人工智能发展也存在很多生态建设方面的问题，包括产业界、学术界和行业用户联合进行基础技术创新研究和应用氛围不够浓厚；行业数据未充分共享，限制了数据开发利用和数据价值深度挖掘；人工智能赋能千行万业智能化建设标准化规范缺失等。

要加速人工智能赋能行业智能化转型，就必须加大人工智能人才储备，进一步完善人工智能生态链，把企业的市场优势和数据优势，转化成人工智能的技术创新优势和应用场景优势，加速推动人工智能赋能千行万业智能化转型。



第二章

实现智能化转型所需的参考架构

▶ 2.1 行业智能化技术需求

大型集团性企业往往具有众多的所属企业或组织，在集团之下，围绕某类相似或关联性业务成立区域（或领域）组织，统筹具体业务执行单位的生产与经营。总结起来一般有总部、区域（或领域）、生产与经营单元 3 个层级。从构建企业智能化视角，生产经营单元包括边缘、终端感知。

如第一章所述，AI 大模型在加速行业智能化转型。企业获取大模型能力的途径有以下几种：一是，自行构建：企业可以自行组建研发团队，进行大模型的研发和训练。这需要企业具备强大的技术实力和资源投入，包括算力、数据、人才等方面的支持。二是，合作伙伴：企业可以与专业的 AI 公司或研究机构合作，共同开发和训练大模型。合作伙伴可以提供技术支持、算力资源、数据等方面的帮助，加速大模型的研发和应用。三是，云服务平台：许多云服务

提供商提供了强大的 AI 平台和工具，企业可以通过订阅这些平台的服务，获得大模型能力。这种方式可以节省企业的研发成本和时间，同时还能享受云服务提供商的技术支持和稳定性。四是，企业基于模式二三获得的高阶模型构建专属的低阶模型，如：L1 或者 L2 模型。模式四是当前行业客户获取大模型能力的主流。在模型构建过程中需要考虑训练和推理两个方面。

对于训练，总部通常会统一构建共享的高阶模型，需要建设大规格的 AI 算力。区域（或领域）根据自身特点基于总部的模型进行二次训练，形成更符合自身业务需求的低阶模型，需要建设较小规模的 AI 算力。

对于推理，总部推理服务于全公司共享业务，包含为全公司服务的推理业务和为全公司服务的推理资源池，算力需求较大；区域（或领域）推理服务于区域（或领域）业务，算力需求较小；边端侧有大量的实时控制业务，需要就近部署推理算法以降低网络时延，不同的现场需要适配不同的周边系统和本地化业务流程，因此需要就近部署推理算法便于集成适配、减少联动更新。边端侧推理服务范围小，算力需求最小。

综上所述，两级训练三级推理是先锋企业的典型部署范式，如下图 2-1 所示。



图 2-1 行业智能化两级训练三级推理的部署范式

基于上述的范式，各企业可以根据自己的业务需求进行适配和裁剪。我们可以从算力集群化、算力系列化、高质量数据、应用快速迭代、高质量数据通信这 5 个方面来分析在 AI 大模型加速行业智能化背景下，企业对 ICT 基础设施的需求。

大规模参数的训练呼唤大算力，需要集群来保障大规模、长期稳定、高可靠的算力供给。分层分级建设和部署大模型是各行业的基本要求，需要算力系列化适配各种业务场景。行业大模型的训练和推理都需要高质量的行业数据来支撑。需要应用快速迭代降低开发门槛保障 AI 落得实。需要高质量的数据通信保障智能化各场景业务需求，保障 AI 用得好。

2.1.1 算力集群化

因为总部和区域（或领域）需要做模型训练和推理资源池，有建设大规模 AI 算力的需求。单卡性能增长有限，想要增加性能，只有集群的模式满足大算力的需求。AI 集群是个系统工程，需要兼顾计算、网络、存储等的跨域协同及优化，助力企业构筑高效协同的算力集群。AI 集群建设有如下 3 大技术需求：

首先，需要大规模集群能力，并支持弹性扩展。总部需要部署高阶模型，不同的高阶模型对算力要求不同，百亿参数百张卡，千亿参数千张卡，万亿参数万张卡。

其次，集群需要实现低时延、大带宽、高可靠的网络。

训练集群网络的丢包率会极大影响算力效率，零丢包是训练集群网络的基本要求。万分之一丢包，算力降低 10%；千分之一丢包，算力降低 30%。

面对千亿、万亿参数规模的大模型训练，仅仅是单次计算迭代内梯度同步需要的通信量就达到了百 GB 量级，此外还有各种并行模式、加速框架引入的通信需求，使得传统低速网络的带宽远远无法支撑集群的高效计算。因此要充分发挥计算资源的强大算力，必须构建一个全

新的高性能网络底座，用高速网络的大带宽来助推整个集群计算的高效率。

在大模型的训练和推理中，集群间的多卡通讯频繁，对时延敏感。以推理为例，多卡低时延通信挑战：低时延推理需要采用多卡模型并行来分担参数和 KV 缓存，提升等效的 HBM 带宽；Decoding 阶段引入大量小数据包 allreduce 集合通信；例如 GPT-3 175B，每个 token 需要 192 次 allreduce 通信，在 BatchSize=1、FP16 计算精度下每次业务包仅为 24KB，对通信时延极其敏感。

最后，在存储方面，大模型基于算力和数据驱动，通过高频迭代学习，强化模型效果，需要实现大容量、高带宽、高 IOPS、高可靠。需求一，大容量：参数规模大、单模态走向多模态，对存储容量带动数据存储从服务器本地盘走向外置存储，且要求按需扩展。需求二：高带宽、高 IOPS：Checkpoint 模型数据集需要快速保存，需要大带宽的存储写入能力；为保障训练效率，需要加速加载海量小文件训练数据、减少 GPU 卡空载等待时间、提升算力利用率。需求三，高可靠：长时间训练过程中，存储的高可靠性直接影响模型参数的准确性和训练效率。

2.1.2 算力系列化

基于前述的部署范式，企业总部需要构建共享的高阶模型训练和共享推理资源池，根据不同的业务需求需要使用 AI 算力集群、训练服务器、推理服务器、训推一体机来建设训练中心和推理中心。比如自行构建万亿参数的大模型，需要万卡级规模的 AI 算力集群；构建千亿参数的大模型，需要千卡级规模的 AI 算力集群；基于预训练的大模型构建百亿参数的模型，需要百卡规模的 AI 算力集群。

区域需要建设低阶模型训练和较小的推理算力，不同企业区域业务规模、业务种类差异大，对算力需求差异大。需要使用 AI 算力集群、训练服务器、推理服务器、训推一体机来建设训练中心和推理中心。比如总部构建的模型已覆盖大部分业务，区域只需要针对差异化的业务基于总部预训练的模型做差异化的训练，训推一体机即可满足。

边缘端侧算力需求小，同时有环境限制，需要使用不同规格的推理服务器、推理模块、终端内嵌的推理算力几种方式来满足工业生产现场、野外、本安、工地等差异化的业务场景诉求。边缘设备需要具备边缘自治能力，实现边缘侧业务自闭环。比如野外环境恶劣，在网络中断的情况下，需要提供可靠的本地 AI 推理能力、快速切换故障节点，保障业务不中断。

2.1.3 高质量数据

数据将是构建大模型竞争力的核心要素，高质量的行业数据尤为稀缺，从企业外部采购数据成本高，行业先锋都需要打造自己的数据飞轮，将海量数据采集好、管理好，支撑行业智能化。

行业大模型的训练和推理都需要高质量的行业数据来支撑。训练、推理数据不单有 IT 数据，还有 OT 数据，OT 数据是指由操作技术（Operational Technology）产生的数据，包括来自工业控制系统、物联网设备、传感器和其他 OT 设备的数据。OT 数据的实时采集和管理为大模型提供高质量的数据集，驱动“数据 - 模型 - 应用”飞轮，让模型更好用。

高质量的数据需要统筹规划感知、存储、网络、数据治理、数据安全，需要做好整个系统的顶层设计和各个子系统的协同。有 5 大技术需求：

首先，需要感知层全面采集 OT 数据和接收任务命令。在 OT 数据的采集和上传上，行业的传感器、仪表和设备数字化和智能化参差不齐，普遍存在智能化和智能化改造的诉求：一些仪表还是机械式、模拟式仪表，无法提供智能化输出；OT 现场设备联网的挑战是“有线联网难、无线联网慢”，导致 OT 设备联网率很低，即使是一些已经联网的仪表、传感器，也仅仅支持本地互联，感知数据仅用于有限场景，在本地经过处理的结果向上级系统提供，原始的感知数据不上传；一些设备（仪表、传感器）对外连接七国八制，集成复杂，可用性低，导致数据无法及时、全量上传，数据采集时间无法对齐，数据无法支撑智能 AI 训练；需要智能化的终端，通过智联操作系统，将协议复杂、系统孤立的终端有机协同起来，实现对同一感知对象的联动感知能力，做到“一碰传、自动报”，以满足更加复杂、高精度、高速度、智能化和协同的作业要求。

其次，需要大容量、支持访问协议多样化的存储。OT 数据量大，文件格式多样，训练前需要从跨域多系统将大量的 OT 样本数据上传到模型训练中心，对存储要求高。需要存储提供大容量、支持从跨域多数据源拷贝 PB 级原始数据以及访问协议多样化。

再次，构建 IT 与 OT 融合的扁平化工业网络对于工业企业而言意义非凡，将真正实现“数据上得来、算力下得去、上下游贯通”。数据上得来：数据“逐层上送”的过程中，每一层的数据传递过程都存在较多的数据信息损失，因此工业企业需要打破层级壁垒，解决跨系统数据流转难问题，实现数据直采。算力下得去：算力下沉至边缘 / 端侧，让工业数据与 AI、算力相融合，帮助工业企业实现 AI 质检、智能排产、工艺优化等创新应用。上下游贯通：通

过生产现场进度与客户实时共享、产品设计方案与合作伙伴共享等上下游的连接和协同，帮助工业企业与客户、合作伙伴之间建立更紧密、更坚固的纽带。

第四，工业企业需要有效的数据治理，提供高质量的数据，支撑大模型训练。IT 数据需要从源头治理，制定数据标准，从应用规范、应用产生开始治理，保障数据“优生”。传统数据治理主要用于大数据业务，智能化时代数据治理将为模型的训练和推理服务，因此数据治理需要与 AI 打通。数据还需要进一步外溢和延展，比如：企业的数​​据从局限于自身内部转变为企业间数据协同、流通、共享与交易，便于在上下游和产业间形成协同、在企业间和产业内形成数据交易、数据经济和数据市场。

最后，数据安全作为数据价值化的基石，对于保障商业秘密、企业数据主权，确保数据使用合规至关重要。为此工业企业需要构建事前预防、事中预警、事后追溯的全套数据安全能力，让数据使用更安全。

2.1.4 应用快速迭代

随着 AI 技术的发展、智能化应用的深入，应用场景变得更多元、更复杂。每个行业的应用场景都有成百上千，每个子场景对 AI 模型的泛化性要求不同，在特定的场景，需要对 AI 模型进一步优化和重构，以适应生产环境。比如：在工业场景下，有工业质检、安全巡检等应用，质检子场景中生产线要适配不同的订单，生产的零部件千差万别，安全巡检子场景中巡检区域可能地形复杂多样并且自然环境恶劣。在交通场景下，包含若干细分的子场景，以公路子场景交通拥堵治理来说，每一个区域（车辆数、拥堵点、停车场、周边路网承载力），

每一个路口（全时、全域、全要素状态、事件、行为），每一辆车（出行、时间、地点、路线）情况和需求不尽相同，复杂多变。

这些多元复杂的场景，都提出了对 AI 的需求，在将 AI 应用于这些场景过程中，即使有自动化程度很高的工具，在传统的 AI 模型开发模式下，也不得不逐个定制开发，如同作坊式开发，投入人力多，开发周期长，再加上 AI 人才供给不足，导致 AI 算法开发产能不能满足行业智能化的需要。大模型为解决这类问题提供了很好的方案。有了预先训练好的大模型，每个场景化 AI 开发，都不必再从 0 开始，而是基于大模型做增强训练，并自动化抽取适合该场景部署的小模型，开发周期从月级缩短为天级，相对于以前的作坊式开发，AI 工业化开发效率可以提升 10~100 倍，实现了 AI 模型从作坊式开发到工业化开发的转变。但这样的开发效率需要高效的开发工具链支撑，需要做到针对不同业务场景需求，快速响应、动态按需适配，实现从需求到智能化应用的快速迭代、敏捷应用短闭环。

具体来说，智能化应用的开发工具链需要满足以下 4 个关键需求：

首先需要大模型增强训练敏捷化，减少开发的中间环节，让应用的使用者参与到模型的构建中来，实现全民参与，贡献行业经验，加速 AI 算法的创新和孵化。模型开发工具链要能够支撑从数据标注、模型选择、模型训练、模型评估、模型部署到结果反馈全流程高效作业，降低开发门槛。

其次需要应用开发敏捷化，从传统的“瀑布式”开发走向敏捷式开发，支持低代码、零代码开发，让业务人员可以直接参与开发。打通应用

和 AI 开发工具，实现应用对 AI 算法的调用，使智能化应用的上线周期从月级跨越到天级，让应用开发更加友好、简单。

再次需要数据与 AI 融合，打通数据库、湖、仓和 AI 开发平台之间的互访，让数据在多个引擎间能够高效流转，实现一站式的 DataOps 与 MLOps，最终通过将数据和 AI 的能力整合，加速模型开发。

最后，通过将 AI 引入开发过程，提供代码和测试用例的智能生成、代码解释和代码翻译的智能交互问答、需求获取和代码提交的智能协同，大幅提升开发效率。

2.1.5 高质量数据通信

海量数据是基础大模型训练、行业场景模型二次训练、模型持续迭代进化的“养料”，也是无所不在的推理交互的输入输出。随着 AI 训练、推理走向“工业化开发、规模化应用”，训练集群内部需要高效数据通信，两级训练三级推理部署范式之间的持续迭代需要高效数据通信，数据采集与推理服务器之间也需要高效数据通信。

从总部到区域 / 领域、从区域 / 领域到边端，海量训练样本数据上传、模型推送与训练迭代、无所不在的推理交互，都需要高质量的广域数据通信网络。以自动驾驶模型训练为例，每天有 1TB 的样本数据需要上传到提供训练服务的数据中心，突发带宽高、潮汐效应明显，需要 10GE 接入、400GE 汇聚的高速数据通信；训练生成的模型文件下推，需要广覆盖、大带宽、任务式的弹性数据通信，灵活建立联接；千万应用场景的实时推理交互，需要低时延、高可靠、高并发的低时延数据通信；为了提升投资效率，我们需要在一张数据通信网络同时支持多种应用，既要满足传统应用“万物智联”的要求，也要满足 AI 时代“万智互联、万数智算”的要求。为此，我们需要引入超融合以太 (HCE)、IPv6+、自动驾驶网络 (ADN) 等创新技术，打造新一代数据通信网络，面向行业智能化提供弹性超宽、智能无损、自智自驭的数据通信基础设施。行业智能化基础设施的可持续发展，还需要考虑数据中心的异地容灾、东数西算、协同计算，这就需要引入网络切片、OXC、网络数字地图等技术，打造超大容量极低时延、应用感知体验保障、算网融合智能无损的 IP 骨干网络、OTN 骨干网络，实现跨地域的实时算力调度、极速数据运送、高效协同计算。

与此同时，数据通信的安全至关重要。行业智能化时代，数据流量持续增长，加密威胁持续增加，我们需要高性能、高弹性、高效拦截加密威胁的网络安全解决方案，为行业智能化保驾护航。

行业智能化时代，新一代高运力数据通信网络从“万物智联”走向“万智互联、万数智算”，成为万物智联、弹性超宽、智能无损、自智自驭的关键基础设施。

▶ 2.2 行业智能化参考架构

企业智能化转型是一个长期的、循序渐进的过程，如何选择转型道路、如何分层分级建设智能化 ICT 基础设施，将成为智能化转型的关键，需要有一个明确的指导思想来引领转型过程，在不同的阶段做出匹配的选择，避免走弯路、走错路，提升转型的效率。基于在城市、金融、交通、制造等 20 多个行业智能化实践过程中的总结，华为提出具备分层开放、体系协同、敏捷高效、安全可信等特征的，全行业通用的行业智能化参考架构，联合行业伙伴共同构筑行业智能化的基础设施，使能百模千态的 AI 大模型，加速千行万业走向智能化。

2.2.1 行业智能化参考架构



图 2-2 行业智能化参考架构

行业智能化参考架构是系统化的架构，它包含智能感知、智能联接、智能底座、智能平台、AI 大模型、千行万业等六层。这六层之间不是独立的，而是相互协同的，就像人体一样，能感知、会思考、可进化、有温度，共同服务于千行万业的智能化发展。行业智能化参考架构是一个面向全行业的、能够服务不同智能化阶段的参考架构，通过分层分级建设，选取合适的技术能力和产品，提升企业的智能化水平。它有四个特点：协同、开放、敏捷、可信。

协同：大模型时代，智能化产业的上下游产业多，产品能力复杂，需要各从业企业基于行业智能化参考架构来构建产品和能力，相互之间协同以形成合力，共同完成智能化体系的建设。各个行业的企业之间也需要协同，共同构筑有竞争力的基础大模型、行业大模型，服务于行业的智能化发展。企业内在智能化过程中通过云、管、边、端的协同，业务信息实时同步，提升业务的处理效率；并通过应用、数据、AI 的协同，打通组织鸿沟，使能业务场景全面智能化。

开放：行业智能化发展是一个庞大的工程，需要众多的企业共同参与，以开放的架构助力行业智能化发展：通过算力开放，以丰富的框架能力支持各类大模型的开发，形成百模千态；通过感知开放，接入并打通品类丰富的感知设备，实现万物智联；通过模型开放，匹配千行万业的应用场景，实现行业智能。

敏捷：企业在智能化的过程中，可按照业务需要灵活匹配合适的 ICT 资源，并通过丰富、成熟的开发工具和框架架构智能化业务，让业务人员直接参与智能化业务的开发，快速上线智慧应用。

可信：企业的智能化系统必须是可信的，在系统安全性、韧性、隐私性、人身和环境安全性、可靠性、可用性等方面全面构筑可信赖的能力，并

从文化、流程、技术三个层面确保在各场景中落地；企业智能化应用的运行过程必须是可信的，可追溯、防篡改，避免受到外部的恶意破坏。

2.2.2 行业智能化参考架构分层解析

2.2.2.1 智能感知

智能感知是物理世界与数字世界的纽带，它基于品类丰富、泛在部署的终端设备，对传统的感知能力进行智能化升级，构建一个无处不在的感知系统。

智能感知具备多维泛在、开放互联、智能交互、易用智维等特点。

- 多维泛在

智能化时代，需要对事物进行全方位的感知，才能获取到完整、全面的信息，支撑后续的智能化业务处理。在感知时，通过雷达、视频、温度传感、气压传感、光纤感知等多种类型的感知设备从不同的维度获取数据，进而汇总成为更全面的信息，支撑后续的智能分析和处置；同时，为了保证能够获取到准确且实时的信息，感知设备还需要贴近被感知的对象，并保持实时在线，充分获取感知数据，实时上传至处理节点，形成无处不在的感知。

- 开放互联

行业里各类感知终端种类繁多，协议七国八制导致数据难互通，难以支撑复杂的业务场景。因此，需要开放终端生态，通过鸿蒙或其它智联操作系统，将协议复杂、系统孤立的终端有机协同起来，实现对同一感知对象的联动感知能力，做到“一碰传、自动报”。开放应用生态，ICT 技术与场景化深度融合，实现精细化治理。

· 智能交互

随着各类智能终端的广泛应用，人与人之间、人与设备之间的协同也越来越广泛，视频会议、远程协作等交互场景在行业应用中得到了很大的推广。通过云边协同、AI 大模型等技术的应用，极大地提升设备认知与理解能力，实现软件、数据和 AI 算法在云边端自由流动，并通过包含智联操作系统的终端设备，基于对感知数据的处理结果，在物理世界中进行响应处理，实现智能的交互能力。

· 易用智维

行业的业务场景复杂，对感知的要求也有很大差异，感知设备有相当大的比例安装在不易于部署维护的地点，如荒野、山顶、铁路周界、建筑外围等，其中一部分设备在获取电力、网络资源也存在一定的困难。因此，需要感知设备具备网算电一体集成、边缘网关融合接入等能力，实现感知设备智简部署、即插即用，智简运维平台和工具数字化、智能化，实现无人化、自动化的可视可管可维。

智能感知层的关键技术和部件包括鸿蒙感知、多维感知、通感一体等。鸿蒙感知是以鸿蒙智联操作系统为核心的智能终端系统，具备接入简单，一插入网、一跳入云，安全性强等特点；多维感知是通过雷视拟合、光视联动等技术融合创新，提高全场景感知精准性；通感一体，通过有线和无线组合，实现无处不在、无时不在的感知。

2.2.2.2 智能联接

行业智能化的场景复杂多样，智能联接用于智能终端和数据中心的联接、数据中心之间的联接、数据中心内部的联接等，解决数据上传、数据分发、模型训练等问题。各种场景对联接

都有不同的要求。比如某个工业园区场景中智能终端和数据中心的联接，AOI 机器视觉质检要求实时推理交互，软件包下载要求高峰值带宽，视频会议要求稳定带宽，需要借助网络切片保障不同流量的互不干扰。在数据中心中，AI 训练集群网络丢包率会极大影响算力效率，万分之一的丢包率会导致算力降低 10%，而千分之一的丢包率会导致算力降低 30%。因此，行业智能化需要万物智联、弹性超宽、智能无损、自智自驭的智能联接。

· 万物智联

在行业智能化时代，种类繁多的感知终端（如：雷达、行业感知、光纤感知、温度感知、气压感知等）都需要通过网络自动上传感知数据，以支撑各种类型的业务系统。数据上传需要实时、准确，不能有丢失。智能联接综合采用 5G-A、F5G Advanced、Wi-Fi 7、超融合以太（HCE）、IPv6+ 等多种网络技术，推进全场景、全触点、无缝覆盖的泛在联接，支撑数据采集汇聚，推进智能应用普及，为智能化参考架构的持续进化构筑万物智联的基础。

· 弹性超宽

随着行业智能化不断发展，感知能力不断丰富与增强，生成的业务信息量也在极速增长，支撑大模型的训练数据更加丰富完善，训练出的模型更加精准；训练出的模型也要迅速下发，推动业务处理更加智能。面向 PB 级样本训练数据上传、TB 级大模型文件分发的突发性、周期性、超宽带联接需求，需要建设大带宽、低时延、智能调度的网络；基于时延地图和带宽地图动态选择最优路径，实现极速推理和实时交互，为行业智能化参考架构打通“数据上得来、智能下得去”的持续进化循环。

· 智能无损

面向超大规模 AI 集群互联需求，以 400GE/800GE 超融合以太、网络级负载均衡等技术实现大规模、高吞吐、零丢包、高可靠的智能无损计算互联；智算数据中心网络升级，以网强算，通过算、网、存深度协同优化，支撑万亿级参数的模型训练，让智能化参考架构越来越智能。面向海量智能感知终端连云入算、AI 助理以云助端等场景，基于网络大模型 (NLM) 实现智能感知应用类型、智能优化联接体验、智能保障网络质量，为极速推理、协同工作、音视频会议等各种应用提供智能无损的高品质联接，让智能持续进化，服务更多的生产、生活场景。

· 自智自驭

基于网络大模型识别应用与终端类型，准确生成配置与仿真验证、准确预测故障与安全风险，并实现网络零中断（智能预测网络拥塞准确率 99.9%）、安全零事故（智能预测未知威胁）、体验零卡顿（智能识别应用类型并保障体验），加速网络自动驾驶向 L4 级迈进，实现网络的自智自驭，提升智能化参考架构的整体运转效率。

智能联接的主要涉及到接入网络、广域网络、数据中心网络，其技术特征有：

接入网络：承担着感知设备的接入及汇聚到数据中心网络或广域网络的职责。接入网络通过 5G-A、F5G Advanced、Wi-Fi 7、超融合以太 (HCE)、IPv6+ 等技术，实现稳定、可靠、低时延的感知设备接入；同时，接入网络还承载着多种业务类型，比如实时业务处理、训练数据采集与上传、推理模型下发至边缘计算节点等，需要接入网络能够根据业务类型分别设置网络资源，为不同的业务数据设置不同的资源优先级。

广域网络：具备多分支机构的大型企业存在大量的数据跨分支机构互传的场景，如训练数据上传、算法模型下发、业务应用下发、业务数据传输等，相应的需要在分支机构之间提供稳定、大带宽的广域网络。企业可根据自身的实际情况，选择租用运营商网络或自建广域网络的方式，获取稳定、可靠、高带宽的多分支机构间的网络联接能力。

数据中心网络：随着 AI 大模型的兴起，大模型训练成为数据中心的一个重要职责，其超大规模的数据分析对数据中心的网络也带来了新的挑战，传统的基于计算机总线的数据中心网络技术已无法满足大模型训练的要求。因此，数据中心网络需要新的网络架构，能够打通各协议间的壁垒，“内存访问”直达存储和设备；并统一芯片侧高速接口，打破“带宽墙”，使能端口复用。数据中心的业务类型也是多样的，例如在大模型训练时就存在参数面、业务面、存储面等网络平面，需要能够按照业务类型建立网络平面，并相互隔离。

2.2.2.3 智能底座

智能底座提供大规模 AI 算力、海量存储及并行计算框架，支撑大模型训练，提升训练效率，提供高性能的存算网协同。根据场景需求不同，提供系列化的算力能力。适应不同场景，提供系列化、分层、友好的开放能力。另外，智能底座层还包含品类多样的边缘计算设备，支撑边缘推理和数据分析等业务场景。

智能底座层具备算能高效、开放繁荣、长稳可靠、算网协同等特点，以更好地支撑行业智能化。

· 算能高效

随着大模型训练的参数规模不断增长、训练数

数据集不断增大，大模型训练过程中需要的硬件资源越来越多、时间也越来越长。需要通过硬件调度、软件编译优化等方式，实现最优的能力封装，为大模型的训练加速，提升算能的利用率。同时，针对基础大模型、行业大模型、场景大模型的训练算力需求，以及中心推理、边缘推理的算力需求，提供系列化的训练及推理算力基础设施配置，根据业务场景按需选择，确保资源价值得到最大化的利用。

在数据存储方面，闪存技术具备高速读写能力和低延迟特性，并伴随着其堆叠层数与颗粒类型方面突破，带来成本的持续走低，使其成为处理 AI 大模型的理想选择。通过全局的数据可视、跨域跨系统的数据按需调度，实现业务无感、业务性能无损的数据最优排布，满足来自多个源头的价值数据快速归集和流动，以提升海量复杂数据的管理效率，直接减少 AI 训练端到端周期。

· 开放繁荣

不同场景、不同类型的大模型，根据大模型的参数规模、数据量规模，需要的算力有着很大的差异；在推理场景，中心推理和边缘推理对算力的要求也不一样。行业用户可以根据实际业务场景选择不同的模组、板卡、整机、集群，获取匹配的算力；并可在品类丰富的开源操作系统、数据库、框架、开发工具等软件中进行选择，屏蔽不同硬件体系产品的差异，帮助用户在繁荣的生态中选择合适的产品和能力，共同形成行业智能化的底座。

· 长稳可靠

大模型业务场景下，一次模型训练往往要耗费数天甚至数月的时间，如果中间出现异常，将会有大量的工作成果被浪费，耗费宝贵的时间和计算资源。为减少异常导致的训练中断、资源

浪费，要保证训练集群长期稳定，提升集群的稳定性；同时，在出现极端情况时，可以使用过程数据恢复训练，降低因外部因素带来的影响。

· 算网协同

随着大模型的参数数量、训练数据规模不断增长，模型训练所消耗的时间也不断增加，逐渐变得不可接受。传统的计算机总线 + 网络的数据传输方式已成为瓶颈，难以继续提升效率。因此，我们需要算网协同的传输架构，提升数据的传输效率和模型训练速度。同时，网络需要参与计算，减少计算节点交互次数，提升 AI 训练性能。

同样，在大模型训练过程中，数据在存储、内存、CPU 间移动，占用大量的计算和网络资源。为减少这些资源占用，需要存算协同架构，通过近存计算、以存强算的能力，让数据在存储侧完成部分处理，将算力卸载下沉进存储实现随路计算，减少对 AI 计算能力的占用。

智能底座的主要技术特征有：

计算能力：简称算力，实现的核心是 GPU/NPU、CPU、FPGA、ASIC 等各类计算芯片，以及对应的计算架构。AI 算力主要以 GPU/NPU 服务器为主。算力由计算机、服务器、高性能计算集群和各类智能终端等承载。算力需要支持系列化部署，训练需要支撑不同规格（万卡、千卡、百卡等）的训练集群、边缘训练服务器；推理需要支持云上推理、边缘推理、高性价比板卡、模组和套件。并行计算架构需要北向支持业界主流 AI 框架，南向支持系列化芯片的硬件差异，通过软硬协同，充分释放硬件的澎湃算力。

数据存储：复杂多样的业务场景，带来了复杂多样的数据类型。数据存储需对不同类型的数

据，通过全闪存存储、全对称分布式架构等技术手段，为不同的业务场景提供海量、稳定高性能和极低时延的数据存储服务；为特定业务场景提供专属数据访问能力，如直通 GPU/NPU 缩短训练数据加载时间至 ms 级；并具备数据的备份恢复机制，以及防勒索机制等安全能力，确保数据的安全、可用。

操作系统：操作系统对上层应用，要屏蔽不同硬件的差异，提供统一的接口，要完成不同硬件的兼容适配，提供良好的兼容性，为应用程序的部署提供尽可能的便利；针对不同的硬件的特征，操作系统需要针对性的优化，确保能充分发挥硬件的能力；在多 CPU、CPU 和 GPU、NPU 协同的情景下，操作系统如何协调调度，也是一个关键的能力。

数据库：海量、格式多样的数据，追求极致的业务性能，对数据库也带来了新的挑战。为了适应业务的变化，数据库需要高性能，海量数据管理，并提供大规模并发访问能力；高可扩展性、高可靠性、高可用性、高安全性、极速备份与恢复能力，都是对数据库的基本要求。

云基础服务：智能底座上运行的各种应用、服务，在不同的时间段对应的业务量是有差异的，为了合理利用智能底座的硬件资源，智能底座通过虚拟化、容器化、弹性伸缩、SDN 等技术，对外提供云基础服务能力，提升资源的利用效率。

2.2.2.4 智能平台

在海量的数据从感知层生成、经过联接层的运输，汇聚到智能平台，通过数据治理与开发、模型开发与训练，积累行业经验，最终服务智能应用的构建。

智能平台具备智简创新、敏捷高效、极致体验等特点，理解数据、驱动 AI，支撑基于 AI 大模型

的智慧应用的快速开发和部署，使能行业智能化。

- 智简创新

围绕软件、数据治理、模型、数字内容等生产线能力，提供一系列的开发使能工具，并通过数据、AI、应用的协同，让智慧应用的构建更高效、更便捷；让行业应用的创新更简单、更智能。

- 敏捷高效

智能化的开发生产线能力，为业务人员提供了多样化的业务开发方式选项；强大的 DevOps 能力让业务迭代开发过程更敏捷，一键发布能力让业务上线速度更快，效率更高。

- 极致体验

具备简单易用的低代码、零代码业务配置能力，开发门槛低，业务人员可以直接参与到模型开发、数据治理、应用开发中；为不同的用户提供个性化的操作界面，提升使用者的体验。

智能平台层的主要技术特征包括数据治理生产线、AI 开发生产线、软件开发生产线以及数字内容生产线。智能平台支持 AI 模型在不同框架以及不同技术领域的开发和大规模训练。

数据治理生产线：核心是从数据的集成、开发、治理到数据应用消费的全生命周期智能管理。一站式实现从数据入湖、数据准备、数据质量、到数据应用等全流程的数据治理，同时融合智能化治理能力，帮助数据开发者大幅提升效率。

模型开发生产线：是 AI 开发的一站式平台。提供从算力资源调度、AI 业务编排、AI 资产管理以及 AI 应用部署，提供数据处理、算法开发、

模型训练、模型管理、模型部署等 AI 应用开发全流程技术能力。同时，AI 应用开发框架，屏蔽掉底层软硬件差异，实现 AI 应用一次开发、全场景部署，缩短跨平台开发适配周期，并提升推理性能。

软件开发生产线：提供一站式开发运维能力，面向应用全生命周期，打通需求、开发、测试、部署等全流程。提供全代码、低代码和零代码等各种开发模式。面向各类业务场景提供一体化开发体验。

数字内容生产线：提供 2D、3D 数字内容开发，应用开发和实时互动框架。根据用户需求，生成服务，如数字人等。使用者无需专业设备，即可使用的内容生产工具。

2.2.2.5 AI 大模型

AI 大模型分为三层，即基础大模型、行业大模型、场景模型。基础大模型（L0），提供通用基础能力，主要在海量数据上抽取知识学习通用表达，一般由业界的 L0 大模型供应商提供；行业大模型（L1），是基于 L0 基础大模型，结合行业知识构建，利用特定行业数据，面向具体行业的预训练大模型，无监督自主学习了该行业的海量知识，一般由行业头部企业构建；场景大模型（L2）指面向更加细分场景的推理模型，是实际场景部署模型，是通过 L1 模型生产出来的满足部署的各种模型。

AI 大模型在发展过程中呈现出了行业重塑、持续演进、开放共建等特点。

· 行业重塑

AI 大模型叠加行业场景，赋予行业场景更智能的处理能力，提升业务效率，降低企业成本，

促进行业创新，为行业的发展注入新的生命力，重塑行业的智能化进程。

· 持续演进

行业场景使用大模型提升业务效率的同时，也会产生大量的业务数据，这些数据再对大模型进行训练，让大模型的能力越来越强大，推理越来越准确，成为行业智能化的有力支撑。

· 开放共建

行业客户与大模型供应商共同打造多样化多层次的大模型，构筑满足各类场景各种需要的大模型，为不同行业场景提供多样化的选择，服务行业智能化发展。

大模型聚焦行业，从 L0、L1 到 L2，遵从由“通”到“专”的分层级模式，可实现从 L0 通用模型到 L1 行业模型再到 L2 专用模型的快速开发流程。

在建设大模型体系时，要依照企业的规模、能力、组织结构和需求因地制宜，层层落实，要充分考虑云网边端协同、网算存的协同，让 AI 上行下达。大模型可以分层分级建设，从 L0 到 L1，再到 L2，不断的有行业数据加入来提升模型的训练效果，同时也需要模型压缩来节约推理资源。模型压缩是实现大模型小型化的关键技术，大模型通过压缩技术可以达到 10~20x 参数量级压缩，使千亿模型单卡推理成为可能，节省推理成本；同时，模型压缩降低计算复杂度，提升推理性能。

在实际应用中，需要结合业务场景变化，迭代演进 AI 大模型能力，边学边用，越用越好。对于 NLP 大模型，可以结合自监督训练方式，进行二次训练，不断补充行业知识；在具体任务场景下，可以使用有监督训练方法进行微调，

快速获得需要达到的效果；进一步的，可以基于自有训练后的模型，进行强化学习，获得更出色的模型。对于 CV 大模型，企业 / 行业用户，可以结合自有行业数据，进行二次训练，迭代获得适配与自身行业的 L1 预训练大模型；同时，在具体细分场景上，可以提供小样本，基于行业预训练的 L1 模型进行微调，快速获得适配自身业务的迭代模型，小样本量，迭代也更快速。

大模型的三级模型之间可以交互优化。L0 模型可以为 L1 模型提供初始化加速收敛，L1 可以通过模型抽取蒸馏产生更强的 L2 模型，L2 也能够在实际问题中通过积累难例数据或者行业经验反哺 L1。

2.2.2.6 智赋万业

千行万业的智慧应用是行业智能化参考架构的价值呈现，每个个体所能感受到的个性化、主动化服务体验都来自应用。智慧应用的发展关键是探索可落地场景，对准其痛点，通过 ICT 技术和行业 / 场景 AI 大模型的结合，快速创造价值。所有这些场景汇聚起来，便能涓滴成河，逐步完成全场景智慧的宏伟蓝图。

▶ 2.3 华为行业智能化技术实践

行业智能化参考架构为行业的智能化发展提供了一条有效的道路，华为依托行业智能化参考架构，叠加多年的行业智能化实践，打造了一系列的智能感知、智能联接、智能底座等硬件能力，一系列的模型使能工具，并发布了盘古大模型，以开放的心态拥抱行业伙伴，共同构筑智能化的未来。

2.3.1 智能感知

在智能感知领域，华为推出了雷达、光纤传感以及软件定义摄像机等感知产品，并结合不同感知设备特点进行组合，实现了光视联动智能感知和雷视拟合智能感知。

2.3.1.1 光视联动智能感知

华为通过组合光纤传感、软件定义摄像机两大感知产品，实现光视联动智能感知，具有全覆盖，低漏报，全天候，少误报的优势。

分布式光纤传感振动探测产品 OptiXsense EF3000 采用华为独有的低噪声相干接收系统和高性能 oDSP 算法，对微弱信号引起的光纤细微拉伸有极高的检测灵敏度，检测形变量达到头发直径的三十万分之一；同时结合了华为独有的超高分辨率采样技术和大尺度线性探测技术，无论是强信号还是弱信号都能真实还原，动态范围大。

在铁路、机场、油气管线等重要行业中，关键场地、设施等防范是安全的中中之重，天然存在周界防护的刚性要求。针对相关场景，华为基于光视联动智能感知能力打造周界防护站方案，通过对抗模型的细节特征提取，对入侵行为和干扰行为的细节差异进行识别，并基于环境感知网络的全局性特征提取网络，对环境全局特征进行提取。将环境特征与细节特征融合，再通过全域态势判决将信号区分为入侵和干扰。将告警准确率提升 90%，普通单事件检测算法误报率为 10 次 / 公里 · 天，IDF-AD 支

持误报率降低至 1 次/公里·天，提供“全覆盖、全天候、全智能”的防护监测能力，实现无人化巡检防护。

2.3.1.2 雷视拟合智能感知

华为通过组合雷达、软件定义摄像机两大感知产品，实现智能感知雷视拟合。华为 ASN850 毫米波感知雷达，采用华为先进的 5G massive MIMO 大规模天线阵列和超分辨率算法；自研高清摄像机结合昇腾计算卡、墨子镜头、烛龙传感器等，支持 4K 小目标检测。

以智慧高速建设场景为例，通过超远探测和雷视拟合技术，使得感知距离扩大到 1000 米，针对弯道等特殊场景采用视场动态增强技术，整体同比业界方案可减少 25% 杆站建设成本。

2.3.1.3 鸿蒙感知

鸿蒙系统，是基于微内核的全场景分布式 OS，可按需扩展，实现更广泛的系统安全。鸿蒙系统拥有分布架构、内核安全、生态共享、运行流畅四大优势。

鸿蒙系统可实现全场景，统一内核，相比于传统“物联感知终端操作系统”碎片化，系统割裂，鸿蒙系统基于一套操作系统灵活组装，实现设备系统归一，无论设备大小，只需要一个操作系统；鸿蒙系统采用分布式软总线技术实现不同领域感知设备之间近场感知、自发现、自组网，完成无感连接，多个感知设备自动协同宛如一个物理设备，可以提供任务在多个感知设备上的一致体验感。

鸿蒙已经在城市、矿业、电力、公路等行业应用。构建形成统筹规范、泛在有序、标准统一、互联互通的感知体系，促进感知终端的共建共享，

提升感知数据汇通共用水平。

华为于 2021 年将鸿蒙系统 L0 到 L2 层面的代码全部捐献给开放原子开源基金会。目前开源鸿蒙项目群拥有多位捐赠人，生态快速成熟壮大。

2.3.2 智能联接

华为提供万兆园区网络、数据中心智算网络、IP 广域网络、网络安全解决方案，IntelligentRAN 网络以及 F5G 智简全光网，推进智联万物，加速网络升级，实现以网强算，促进开放互联，保障网络安全。

2.3.2.1 高品质万兆园区网络

华为高品质万兆园区网络助力企业建设以体验为中心的园区。Wi-Fi 7 和多速率交换机实现终端万兆极速接入，音视频体验保障方案实现万人园区视频会议 0 卡顿，融合 AI、大数据和华为数十年网络运维经验的 NCE 园区网络数字地图，网络开局效率提升 10 倍，故障定位效率提升 40 倍。华为已为比亚迪、苏黎世联邦理工学院等全球众多知名客户打造极致体验的园区。在 2022 年 Gartner 企业有线无线网络基础设施的魔力象限中，华为成功入选领导者象限。

在企业分支场景，华为 SD-WAN 方案助力企业构建极简组网，极简运维的高品质分支网络。华为超融合网关，1 台顶 6 台，实现分支极简组网，提升部署运维效率 30%；企业业务上云加速，华为超融合云原生网关，NetEngine AR6700V 能够实现弹性扩容最大到 200G，性能升级业务不中断，助力企业高品质上云。华为 SD-WAN 方案在中国石化等全球众多知名客户部署，良好的交付品质及口碑帮助华为连续 3 年赢得 Gartner “客户之选”。

2.3.2.2 华为星河 AI 智算网络

AI 大模型依赖大规模集群互联，需要数据中心提供高性能、高可靠、可运维、大规模的开放网络作为支撑。华为星河 AI 智算网络基于超融合以太网技术，实现网络智能无损 0 丢包；同时，通过网络级负载均衡 NSLB 算法，以网强算，实现 AI 训练网络吞吐大幅提升至 98%；武汉人工智能计算中心是首个面向产业的人工智能计算中心，选用华为网络、计算、云、存储端到端解决方案，首期建设了 100P AI FLOPS 算力的智算网络；该项目对比测试中，在 MPI、Benchmark 等典型应用测试中，华为网络性能与 IB 网络整体基本持平，完全满足业务的高性能需求。

2.3.2.3 华为智能 IP 广域网络

面向海量训练数据和模型文件高效传递，华为提供智能 IP 广域网络解决方案，以 IPv6+ 为技术基础，实现流、应用、带宽、时长等多维感知和弹性调度，大模型训练数据和模型数据快速高效传输，效率提升百倍，解决企业训推数据传递低带宽等不起，高带宽用不起的业务痛点。通过全业务路由器大缓存、抗突发的能力，满足企业到各级训练中心、推理中心的灵活接入、边云协同的联邦训练的极致高吞吐数据传递，带宽利用率最高达到 95%，逼近物理极限。作为智能煤矿建设的先行者，山西晋能控股集团寺河煤矿基于华为矿山融合 IP 工业网解决方案，建成全国领先的智能煤矿，极大提高了调配管理水平、生产效率及矿井安全性。

2.3.2.4 华为网络安全解决方案

人工智能融合到行业过程中，需要高效模型传输和海量数据训练，华为基于可信网络安全解决方案，在数据传输到 AI 算力网络前进行安

全防护，提供算力基础设施的“非侵入”式的安全防护，实现高效算力和安全防护的平衡。其次华为安全提供训练数据的脱敏、数据防泄漏等数据安全防护能力，实现敏感数据的安全防护。最后对于防护智能应用的终端和用户，提供零信任安全防护能力，保证终端、人员可信接入，最小权限访问业务资源，最大程度保障 AI 关键基础设施的整体安全。“非侵入式”防御架构在华为南方工厂等行业智能化先锋企业进行了部署，已连续多年保障业务持续稳定运行，为业界树立了标杆。

2.3.2.5 华为 IntelligentRAN 网络

面向新业务的多样化发展以及新站型新频的引入带来的无线网络结构性挑战，华为提供 IntelligentRAN 解决方案，以数智感知、数字孪生、意图开放、智能空口为技术基础，实现对无线环境信息高精度感知、多维分析决策、意图驱动人机交互以及高效空口信道测量的能力；无线大模型通过海量数据训练特征学习，在闭环流程中形成认知与决策，有望将无线网络真正推向意图自适应的网络级智能化高阶水平。通过多频协同、宏微协同、网络级节能、网络故障预防预测、直播业务体验保障等热点场景引入智能化技术，满足无线网络极致节能体验按需的网络智优，故障预测预防主动排障的运维智简，应用级确定性体验保障的业务智营，真正帮助运营商实现提质增收、降本提效。

2.3.2.6 F5G 智简全光网

F5G 是第五代固定通信网络，由欧洲电信标准协会 (ETSI) 定义，致力于从光纤到户迈向光联万物。F5G 基于光纤通信技术，在增强超宽带 (eFBB)、全光联接 (FFC) 和可保障品质体验 (GRE) 三大场景已经取得巨大成功，并向绿色敏捷全光网 (GAO)、实时韧性联接

（RRL）和光感知与可视化（OSV）三大全新应用场景演进。华为推出面向 AI 时代的 F5G 智简全光网，包括行业生产网络、园区网络、工业网络、地产前装、周界防护等场景的创新产品和解决方案，为 AI 全面赋能行业打造坚实底座。

在智慧办公，F5G 全光园区为企业打造绿色极简的上云联接，并实现高效的异地办公与远程沟通。在智能制造，全产业链信息高效流转，与精密制造的自动化精准及远程操控同样依靠 F5G 技术所赋能的无损工业网络。在智慧电力，F5G 技术为电力行业构建安全稳定可靠易演进运维的全光通信底座，赋能新型电力系统。在智慧医院，F5G 一举摆脱了传统网络在部署、运维和网络性能上的劣势，实现院区网络多网合一，通过硬隔离方案满足医疗信息系统安全等级保护要求，在确保传输网络可靠性的同时，为诊室、医学影像、远程诊疗等场景提供高效的数据传输和信息互联。

2.3.3 算力底座

2.3.3.1 昇腾系列（AI 计算）

华为致力于打造领先的、坚实的 AI 算力底座，以系统级架构创新，突破规模算力瓶颈，使能百模千态的繁荣发展。华为昇腾 AI 产业生态，如下图 2-3 所示。



图 2-3 华为昇腾 AI 产业生态

华为昇腾 AI 产业生态，围绕昇腾 AI 基础软硬件平台（包括 Atlas 系列硬件、异构计算架构 CANN、全场景 AI 框架昇思 MindSpore、昇腾应用使能 MindX 以及一站式开发平台 ModelArts 等）持续创新，释放昇腾 AI 澎湃算力，性能保持业界领先。

为向世界提供用得起、用得好、用得放心的普惠 AI，我们研发了昇腾系列板卡，作为华为人工智能解决方案的核心部分和关键支撑。基于昇腾系列，华为推出了 AI 训练集群 Atlas 900、AI 训练服务器 Atlas 800、智能小站

Atlas 500、AI 推理与训练卡 Atlas 300 和 AI 加速模块 Atlas 200，完成了 Atlas 全系列产品布局，覆盖云、边、端全场景，面向训练和推理提供强劲算力。

华为的 AI 算力集群，支持万亿参数大模型训练，具备超高稳定性，拥有创新超节点架构，超强性能，大模型训练效率高的算力集群。

首先，依托华为在网络领域的长期能力积累，集群互联采用了全新的华为星河 AI 智算交换机 CloudEngine XH16800，借助高密的 400GE 端口能力，即可实现大规模 NPU 无收敛互联。

其次，超级集群 3 倍提升断点续训性能，10 倍提升稳定性。采用全互联 AA 数据保护架构，独创故障现场保存技术，和盘控协同的超高带宽技术。针对万亿参数大模型训练所特有的，PB 级的 CheckPoint 数据读写，分钟级即可完成。断点续训恢复时间也缩短到分钟级，恢复速度是业界的 3 倍。同时，通过将华为在通信领域积累的高可靠性系统工程能力引入到 AI 集群，实现了器件级的故障预测、节点级精准液冷，万卡集群的稳定性从天级提升到月级，提升 10 倍。

再次，集群采用创新超节点架构，多卡通过总线互联成一个 NPU。通过总线和 openEuler 操作系统的协同，为大模型训练提供充足内存，规避因内存不足所带来的算力空耗现象。

最后，超强性能，大模型训练效率高。通过华为独有的网络级负载均衡 NSLB 算法，在网络硬件不变的情况下，集群内网络吞吐率从 50% 提升到 98%，性能提升 1 倍。

华为坚持技术生态与商业生态双轮驱动，发展

昇腾 AI 产业。我们以开发者为核心，繁荣技术生态，以伙伴与客户为核心，做厚商业生态。坚定践行开放、开源的策略，进一步开放异构计算架构 CANN，全面兼容业界的算子、AI 框架、加速库和主流大模型，以更加灵活的算子开发使能开发者业务创新，以共建、共享、共治，持续贡献昇思开源社区，把昇思 MindSpore 打造成支持大模型和科学智能等 AI 创新的框架。

2.3.3.2 OceanStor 存储

华为推出的 OceanStor 系列是可大规模横向扩展、弹性伸缩的数据中心级智能分布式存储产品，通过系统软件将存储节点的本地存储资源组织起来构建分布式存储池，可向上层应用提供分布式文件存储服务、分布式对象存储服务、分布式大数据存储服务与分布式块存储服务，以及丰富的业务功能和增值特性。借助华为数控分离架构的高性能存储 OceanStor A800，实现数据流 I/O 直通，减免了 CPU 计算耗时，数据加载效率高。

OceanStor 系列存储采用闪存技术加速大模型训练：具备高速读写能力和低延迟特性，并伴随着其堆叠层数与颗粒类型方面突破，带来成本的持续走低，使其成为处理 AI 大模型的理想选择。数据读写性能的大幅提升，将减少计算、网络等资源等待，加速大模型的上市与应用。据华为测算，以 GPT-3 采用 100PFlops 算力下，当存储的读写性能提升 30%，将优化计算侧 30% 的利用率，整体训练时间缩短 32%。

OceanStor 系列存储通过数据编织方案提升 AI 大模型协同分析效率：通过全局的数据可视、跨域跨系统的数据按需调度，实现业务无感、

业务性能无损的数据最优排布，满足来自多个源头的价值数据快速归集和流动，以提升海量复杂数据的管理效率，直接减少 AI 训练端到端周期。

华为存储支持灵活开放：OceanStor 存储软件之上，可以集成第三方的 GPU 服务器、网络节点以及 AI 的平台软件。方便集成商自由、灵活的装配他们的超融合节点。

2.3.3.3 多种算力开放模式



图 2-4 多种算力供给模式

华为致力于打造领先的、坚实的 AI 算力底座，提供多种算力供给模式，来满足行业客户的差异化需求。使能“百模千态”，加速千行万业走向智能化。

(1) 裸算力模式，我们提供算力底座，包括智能感知、智能联接和智能底座。通过直接提供领先昇腾 AI 算力，使能客户和伙伴灵活打造差异化的算力平台和 AI 服务。

(2) 多租户模式，我们提供智能感知、智能联接、智能底座和基础云平台，通过基础算力

+HCS/HCSO 基础云平台能力，方便客户面向多租户提供 AI 算力。

(3) 云算力模式，为不同租户提供叠加的 DataArts 数据治理，ModelArts 一站式 AI 开发平台，使能客户快速进行大模型开发。

(4) MaaS (Model as a Service) 模式，面向千行万业的中小企业，提供开箱即用模型即服务，即 MaaS 模式，加速业务应用上线，快人一步。

2.3.4 盘古大模型



图 2-5 盘古大模型架构

2.3.4.1 盘古大模型

盘古大模型达到千亿级参数，相对于以前的作坊式开发，AI 工业化开发效率可以大幅提升，同时 AI 模型具备更佳的性能。

盘古大模型分为三层，L0 基础大模型，L1 行业大模型，L2 场景大模型。L0 层是盘古的基础大模型，包括自然语言大模型（NLP）、视觉大模型（CV）、多模态大模型、预测大模型、科学计算大模型，它们提供满足行业场景的多种技能。

L1 层是行业大模型，既可以提供使用行业公开数据训练的行业通用大模型，包括政务，金融，制造，矿山，气象等；也可以基于行业用户的自有数据，在盘古的 L0 和 L1 上，为用户训练自己的专有大模型。

L2 层是为企业提供更多细化场景的模型，更加专注于某个具体的应用场景或特定业务，为用户提供开箱即用的模型服务。

盘古 NLP 大模型可用于内容生成、内容理解等方面，并首次使用 Encoder-Decoder 架构，兼顾 NLP 大模型的理解能力和生成能力，保证了模型在不同系统中的嵌入灵活性。在下游应用中，仅需少量样本和可学习参数即可完成千亿规模大模型的快速微调 and 下游适配。2019 年权威的中文语言理解评测基准 CLUE 榜单中，盘古 NLP 大模型在总排行榜及分类、阅读理解单项均排名第一，刷新三项榜单世界历史纪录；总排行榜得分 83.046，多项子任务得分业界领先，是目前最接近人类理解水平（85.61）的预训练模型。

盘古 CV 大模型可用于分类、分割、检测方面，

也是实现模型按需抽取的业界最大 CV 大模型，实现兼顾判别与生成能力。基于模型大小和运行速度需求，自适应抽取不同规模模型，AI 应用开发快速落地。使用层次化语义对齐和语义调整算法，在浅层特征上获得了更好的可分离性，使小样本学习的能力获得了显著提升，达到业界第一。

盘古大模型赋能千行万业，在矿山、气象、药物、分子、电力、海浪、金融等行业发布了 L1 行业大模型。

视觉大模型在煤矿运输场景中，会有大块煤、锚杆、钢筋等异物堵塞皮带。通过煤矿皮带异物识别 API，可快速发现运输皮带上的异物，识别精度超过 98%，检测效率提升 10~100 倍，安全事故降低 90%，避免了停工停产带来的巨大损失；在工业质检上，可以在偏光片流水线上进行质检；可以在铁路 TFDS 的机车进行缺陷检测；还可以做电力巡检，无人机每天会拍摄海量的照片，快速筛选有缺陷的样本，过去用几十个小模型来做这一步，现在用一个统一的大模型能够加快筛选过程，提高筛选质检质量。

盘古气象大模型，是全球首个精度超过传统预报方式的 AI 模型。相关成果在国际顶级学术期刊《Nature》发表。可以在秒级时间内完成全球未来 1 个小时到 7 天的天气预报，精度超过传统数字分析方法的同时预测速度提升 1 万倍。盘古药物分子大模型可以提高小分子合成物筛选速度，使过去数年的传统药物研发周期缩短至一个月以内，大幅提高研发效率。

盘古制造大模型，实现大模型 + 求解器，助力数实融合，过去单产线制定器件分配计划，往往要花费 3 个小时以上才能做齐 1 天的生产计划。盘古制造大模型学习了华为产线上各种器件数据、业务流程及规则以后，能够对业务需

求进行准确的意图理解，并调用天筹 AI 求解器插件，1 分钟即可做出未来 3 天的生产计划。

盘古大模型采用完全的分层解耦设计，可以快速适配、快速满足行业的多变需求。企业既可以为自己的大模型加载独立的数据集，也可以单独升级模型。

2.3.4.2 开放丰富的“开发生产线”

华为提供了丰富的“生产线”，一系列的模型使能工具，使能伙伴共同开发，繁荣生态。

华为数据治理生产线 DataArts 提供一站式数智融合的开发与治理，支持 40+ 数据源接入，10 万 + 并发作业调度，全流程托拽式开发，全生命周期数据治理，帮助企业快速构建从数据接入到数据分析的端到端智能数据系统，消除数据孤岛。

模型开发生产线 ModelArts，是 AI 开发的一站式平台。ModelArts 持续构建大模型训练、推理加速能力，从算力资源调度、AI 业务编排、AI 资产管理到 AI 应用部署，提供数据处理、算法开发、模型训练、模型管理、模型部署等 AI 应用开发全流程能力支撑。ModelArts 提供开放架构，联合伙伴以第三方工具集成的方式，打造全流程工具链赋能各类 AI 开发场景。IDC 报告显示，华为云 ModelArts 蝉联中国公有云机器学习市场 TOP1。

软件开发生产线 CodeArts，提供一站式 DevSecOps 开发运维能力，面向应用全生命周期，打通需求、开发、测试、部署等全流程。同时，提供全代码、轻代码和低代码等各种开发模式，支持多种主流编程语言和开发框架，内置代码检查规则，支持千万 TPS 的自动化测试并发请求。

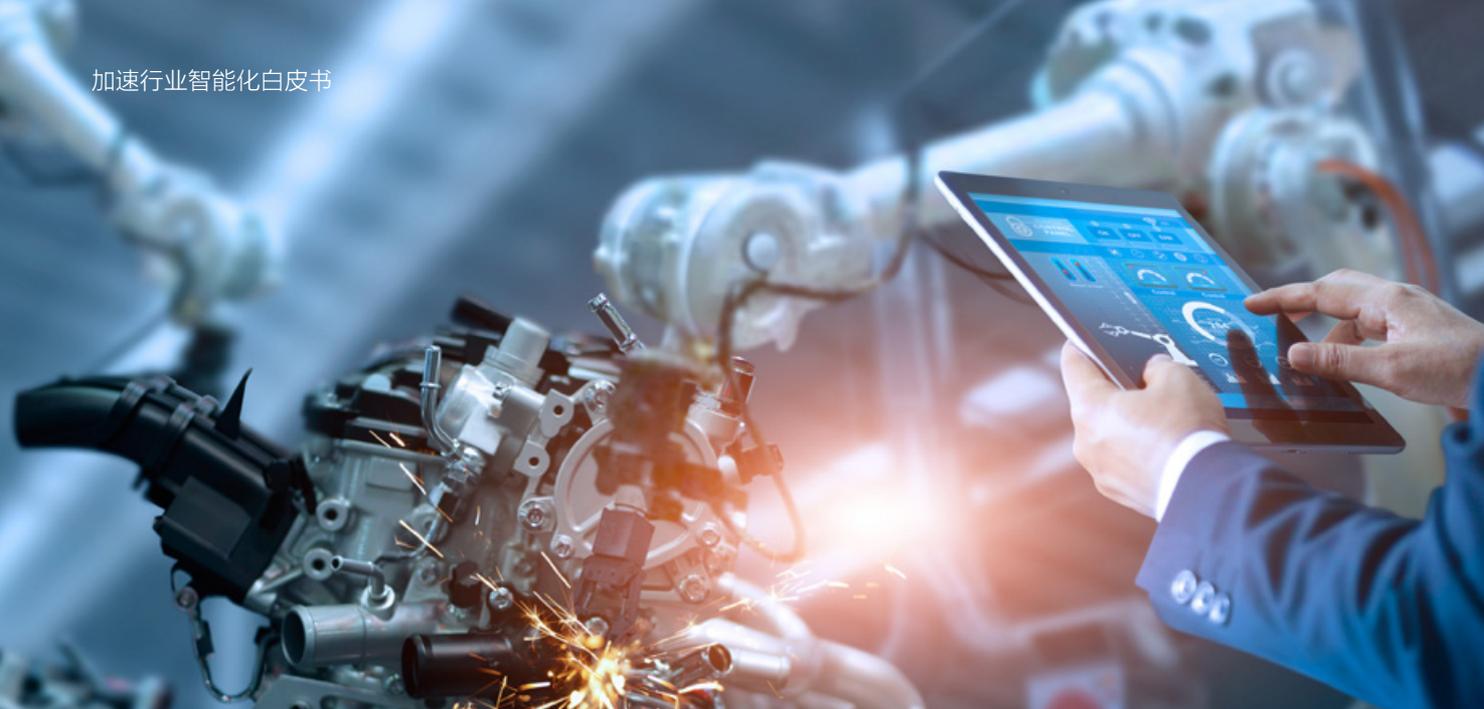
华为数字内容生产线 MetaStudio，提供 3D 数字内容开发，应用开发和实时互动框架。通过华为云 MetaStudio 数字内容生产线，各行业客户都可以便捷地在云上生产数字内容、开发 3D 应用，打造虚拟演唱会、虚拟展会、办公协作、工业数字孪生等 3D 虚拟空间，同时支撑海量用户的实时互动，让虚拟世界和现实世界无缝融合。





**加速智能化
带来新价值**

第二篇



第三章 智能化使能企业生产

▶ 3.1 智能制造

制造业是国家经济命脉所系，是立国之本、强国之基。国家经济要以智能制造为主攻方向推动产业技术变革和优化升级，推动制造业产业模式和企业形态根本性转变，以“鼎新”带动“革故”，以增量带动存量，促进中国产业迈向全球价值链中高端。

发改委等 8 部委在《“十四五”智能制造发展规划》中明确提出，当前中国站在新一轮科技革命、产业变革与加快高质量发展的历史性交汇点上，要坚定不移地以智能制造为主攻方向，推动产业技术变革和优化升级，加大对先进制造、精密制造、高端制造的投入。只有在数字

化技术的加持下，企业才能在产品质量、生产效率、生产成本间进行有效平衡。到2025年，制造企业70%以上要基本实现数字化，重点行业骨干企业初步形成智能化，建成500个智能制造示范工厂。

随着新一代信息技术加速拥抱千行万业，智能制造正在多领域多场景落地开花。最新的市场调研数据表明，人工智能在中国制造业呈现快速发展的趋势，智能制造市场规模自2019年

开始每年保持40%以上的增长率，2025年将超过400亿人民币。

毋庸置疑，AI正在更加深入地融入到工业生产的各个环节中，实现更程度的自动化和智能化。如果对制造主流程进行分析，将不难发现AI应用场景分布在“研产供销服”的各个环节。其中典型的AI场景在各个环节的汇总情况，如图3-1所示。



图 3-1 制造主流程与 AI 应用场景

在各个智能制造的应用场景中，AI 往往作为核心技术，与云、计算、物联网、大数据等技术相结合，共同形成更为强大的应用能力，最终推动工业领域的创新和转型，带来新的商业模式和机会。以下就几个典型应用场景做进一步介绍。

3.1.1 工业 AI 质检

质量检测的目的在于确保产品的质量、安全性和稳定性。制造行业传统质检以人工检测为主，存在检测效率低、工作量大、成本高的问题。而且人工检测结果受主观因素影响，检测标准不统一，导致质量控制困难，容易出现漏检和误判。

工业 AI 质检基于深度学习技术，通过对生产线上的图像、声音和视频等数据进行实时分析和处理，能够自动、高效地识别和检测制造过程中的产品缺陷和工作异常，实现对产品质量的自动检测和结果判定。相比传统计算机视觉的 AOI，AI 具有更高的准确性、灵活性和自我学习能力，可广泛应用到工业制造场景中，如图 3-2。

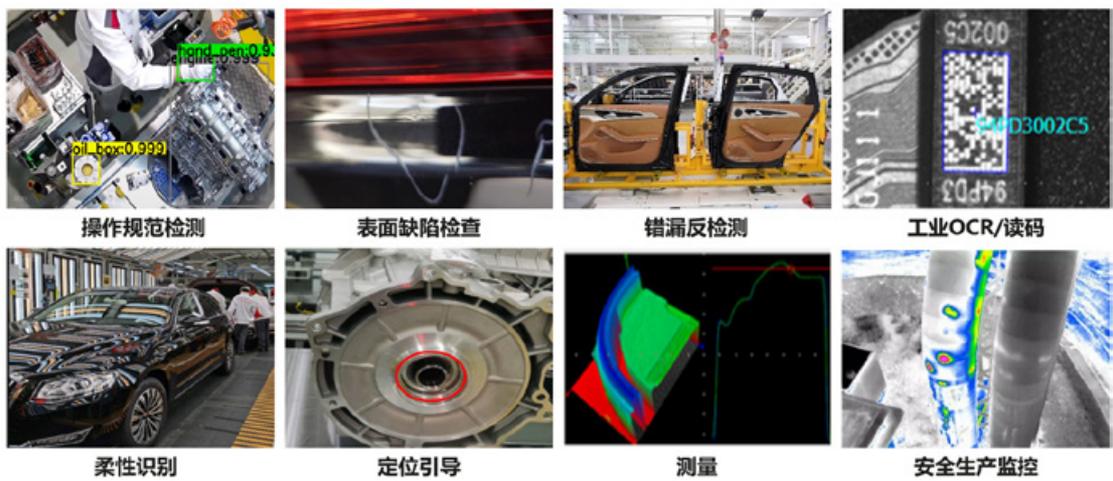


图 3-2 工业 AI 质检应用场景

3.1.2 智能资源优化

制造企业为了满足自身发展，必须不断提高效率、减少浪费和增加产值。随着生产流程、供应链和客户需求变得越来越复杂，传统的资源管理方法往往无法满足现代制造业的需求。在这种背景下，使用 AI 技术实现智能资源优化显得尤为重要和必要。常见的智能资源优化场景有智能排产、供应链优化和能源管理等。

在智能资源优化场景中，AI 使企业能够更加高效、灵活地管理其资源和工作流程，优化企业资源利用率，提高生产效率。以具有代表性的智能排产为例，AI 结合运筹学，使用求解器，完成先进算法迭代，在自动适应生产变换，减少冲突和瓶颈的基础上，可进一步提供决策支持。而在供应链优化和能源管理方面，AI 可以提供实时数据可视化，预测和响应潜在的需求和风险，并自动执行优化操作，如自动化库存管理、自动调节生产参数以降低能源消耗等，最终实现资源的优化配置，同时确保持续和稳定的生产。

3.1.3 工业智能机器人

工业智能机器人是一种在制造和生产线上执行各种任务的自动化机器，例如装配、焊接、搬运、质检等。与传统的工业机器人相比，智能机器人不仅可以执行预定的动作，还可以通过感知、学习和决策能力，使其能够执行复杂的任务并在多变的环境中自主工作。

以工业智能巡检机器人为例，如图 3-3 所示，在生产现场的广域环境范围内，基于融合感知技术，将多种类别的感知需求集成，借助自动化物流载具平台，以及物流系统和统一调度的智能机器人平台，实现对人体感知、人工检查的补强和提升，输出处理数据，提升现场综合管理的流程数字化和管理精细化。

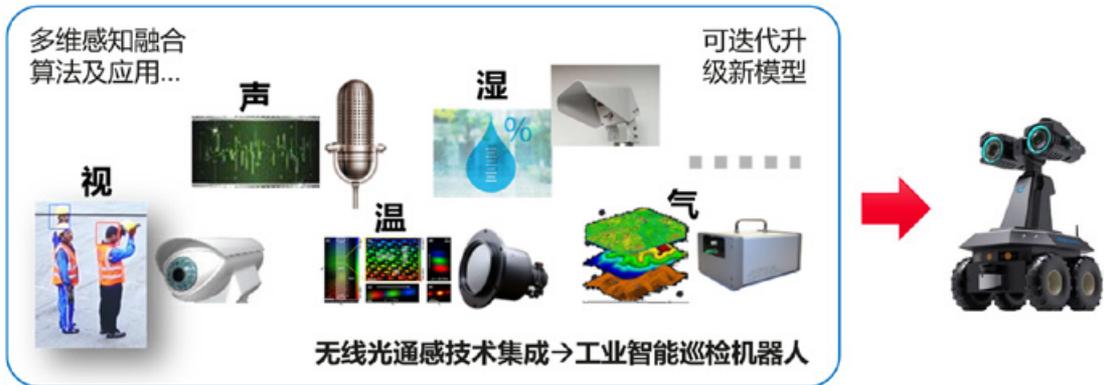


图 3-3 工业智能巡检机器人

3.1.4 预测性维护

在过去，维护通常是反应式的，即在设备出现故障后进行修复。为了避免设备故障停机，工业界转向了预防性维护，即定期检查和维护设备。随着传感器技术、数据存储和分析能力的进步，工业界开始具备实时收集和分析大量设备数据的能力。现代的数据分析和机器学习技术使我们能够从历史数据中识别出设备故障的先兆和模式，从而预测未发生的故障。

预测性维护的核心思想就是使用数据驱动的方法预测和预防设备故障，而不是等到设备真正出现问题后再进行维护。预测性维护可以帮助企业解决生产中设备相关的问题，包括减少停机时间、优化维护计划、延长设备寿命、降低维护成本和提高生产效率。掌握实时的设备健康数据，准确进行故障预测，为管理者提供有价值的洞察，帮助他们做出更明智的运营决策，提高安全性和客户满意度。

3.1.5 企业知识库与智能客服

企业知识库是企业内部的宝贵资产，它集中存储了企业的各种经验、案例、问题解决方案、操作手册等重要信息。在制造领域，企业知识

库可整合制造流程、生产设备操作手册、质量标准 and 故障处理等关键信息。这不仅提高员工获取信息的效率，缩短新员工培训时长，使其能快速融入生产环境，更重要的是保持了企业知识的连续性，有助于生产的顺畅运行，减少因信息缺失或错误而导致的停机时间。

大模型技术能够实现企业知识库的智能检索和内容优化。例如，当产线工人遇到设备故障时，他们可以用自然语言描述问题，大模型则能迅速返回相关的解决方案或故障处理指南。基于与人员的交互式问答，大模型可以基于现有知识库，结合提示词，自动生成新的内容，如 FAQs、教程等。利用大模型的自然语言处理能力，企业的智能客服可以更准确地理解客户的问题，并结合企业知识库，提供具有针对性的答案或解决方案，大大缩短了响应时间和解决问题的周期。大模型还可以在与用户的交互中持续学习，通过实时更新的知识库，不断完善其答案和建议，使服务水平持续提升。

▶ 3.2 智慧煤矿

煤炭作为中国重要的传统能源行业，是中国国民经济的重要组成部分，是长期以来推动中国现代化进程的基础和动力。加快推进智能化建设，是推进煤炭工业高质量发展的必然选择。2020年3月，国家发展改革委、能源局、应急部、煤监局、工信部、财政部、科技部、教育部8部委联合印发了《关于加快煤矿智能化发展的指导意见》，明确了煤矿智能化是煤炭工业高质量发展的核心技术支撑，并规划了3个阶段性目标：到2021年，建成多种类型、不同模式的智能化示范煤矿；到2025年，大型煤矿和灾害严重煤矿基本实现智能化，形成煤矿智能化建设技术规范与标准体系；到

2035年，各类煤矿基本实现智能化，构建多产业链、多系统集成的煤矿智能化系统，建成智能感知、智能决策、自动执行的煤矿智能化体系。该《意见》标志着煤矿行业正式步入建设新阶段。

煤炭行业作为高危行业，为了保护矿工生命安全，煤炭行业少人化、无人化是大势所趋。人工智能技术深度融入矿山，则是行业发展(少人、无人)的必然选择。从机械化到自动化、信息化，再到智能化，从应用、ICT设施到装备，都在向着智能化、协同化、在线化升级。煤炭生产从机械化向智能化演进过程，如图3-4所示。



图 3-4 煤炭生产演进路线

煤炭行业面临着开采地质条件复杂、灾害多发（如煤尘、水、火、瓦斯和顶板等自然灾害）、生产效率较低、作业环境恶劣、人员短缺等挑战，为应对这些挑战，实现增安、提效、少人无人，亟需AI技术的大规模应用。通过对采煤、掘进、通风、运输、排水、机电等多个业务系统进行调研分析，人工智能在煤矿场景存在海量的应用场景。



图 3-5 海量人工智能应用场景

如图 3-5 所示，一方面人工智能存在海量的应用场景，但另一方面这些应用场景都是很碎片化的，试图通过传统的“小作坊”开发模式无法覆盖这些海量的碎片化场景，亟需新的人工智能开发和应用模式，实现人工智能大规模应用。当前探索实践证明，大模型是实现人工智能“工厂化”生产及规模化应用有效手段，通过引入视觉大模型、预测大模型、多模态大模型和自然语言（NLP）大模型，可实现人工智能在安监、生产、决策指挥、经营管理等业务场景下的规模应用，如图 3-6 所示。



图 3-6 大模型四大能力支撑不同业务场景

当前人工智能技术在煤矿行业的应用还处在起步阶段，未来将在安监生产、综合管控、决策指挥、经营管理等方向会持续进行更多的探索，为煤矿的智能化建设提供有效的支撑。预计未来 3~5 年，煤矿行业人工智能市场空间规模达到近 150 亿。

3.2.1 采煤工作面：瓦斯隐患预测

近年来，随着煤矿生产装备技术水平的不断提升，同时在煤矿监管部门的严格监管下，国内煤矿瓦斯安全事故总量和事故死亡人数逐年下降，但相较于其他类型灾害而言，瓦斯仍然是煤矿安全“第一杀手”。按照国家安全监察局等监管部门的要求，国内煤矿须锚定瓦斯“零超限”等一系列目标，完善瓦斯综合治理，实现超前治理、超前预警和预防。

借助大数据平台和 AI 技术，可以对井下采煤设备、矿压、通风、安全监测、地质、瓦斯抽采等系统数据进行智能融合分析，通过人工智能算法结合综采工作面瓦斯涌出机理，实现井下关键位置瓦斯浓度的预测，进而实现瓦斯隐患的超前预警，帮助煤矿提前发现瓦斯安全隐患，同时给出处置预案，避免隐患进一步发展成事故，帮助煤矿提升生产安全性，减少设备停机停产的时间，提升生产效率。

3.2.2 掘进工作面：掘进作业序列智能监测

掘进作业作为煤矿井下采煤生产前进行巷道开采及安全支护的工作环节，是煤矿日常生产中最重要、最危险的环节之一。掘进的速度和质量直接影响到矿井合理的生产布局、稳定的采掘接续，并关系到矿井重大灾害的超前治理。

当前，掘进面装备智能化程度偏低，空间狭小，环境恶劣，人员密集，多工序协同作业，劳动强度大。现场采用“人盯人”的管理方式，监测管理难度大，作业规范性、作业质量难保证，安全风险巨大。近年来，井下重大安全事故 30% 以上发生在掘进面，而其中 50% 以上是由作业不规范、人员监管不到位导致，亟需引入新技术手段突破人工监管的屏障，提升掘进

面安全性，保证稳定生产。

掘进作业序列智能监测系统通过对作业人员及其相关动作进行智能识别，对作业规范性进行 AI 推理分析，结合传感器及设备运行状态数据，实现对危险区域人员入侵、人员摔倒、截割头落地、敲帮问顶、钻眼深度、锚固剂使用数量、搅拌时间等关键动作的实时监测，及时分析作业工序完整性，作业规范性和作业质量，发现问题第一时间对掘进队长进行告警提醒，并可联动停机，确保掘进作业安全规范，作业质量达标。

3.2.3 智能主煤流运输

矿井主煤流运输系统负责将井下采掘工作面生产的原煤运输到井上，是煤矿生产作业系统的大动脉。当前，矿井主煤流运输系统在实际运行中存在诸多安全隐患，例如皮带撕裂、异物、跑偏、堆煤等问题，需要人工巡检并介入处理，避免主运皮带撕裂、断带等事故，造成全矿井停产停工。因矿井主煤流运输皮带距离长，巡检人员每天需行走 5-10 公里，耗时耗力，且容易因问题发现不及时导致事故发生。

通过“计算机视觉 + 人工智能”技术，对主煤流进行可视化精准监测，能够在第一时间发现主煤流运输各种安全隐患，并且通过联动 GIS、融合通信、手机工单、皮带集控等周边业务系统，提升问题处理效率，在减少巡检人员工作量的同时，助力矿井主煤流运输系统更加高效安全运行。

3.2.4 智能洗选煤

选煤厂是煤炭加工的重要场所，选煤厂的主要任务是利用煤和矸石的物理性质不同，通过一系列生产工艺对煤炭进行分选，去除原煤

中的杂质，获得工业生产所需的精煤产品。选煤厂分选系统的运行情况对煤矿具有非常重要的意义。

重介旋流器分选是选煤生产的核心生产工艺，当前的分选工艺参数依赖人工经验设定，产品质量参差不齐，经常出现产品质量过剩和产品灰分超标的问题，给煤矿企业带来直接或间接的经济损失。

人工智能的引入为工艺参数优化提供了新的突破思路。以原煤煤质、入洗煤量、重介密度、入料压力和精煤产品灰分为基础，通过对生产数据建模和训练，生成工艺参数预测模型，然后利用参数预测模型对实时生产数据分析和推理，预测出最优的生产工艺参数，并将其应用到生产系统中，实现最优的控制效果。同时利用新的生产数据对参数预测模型持续优化迭代，从而逐步实现选煤生产数字化、精细化。

3.2.5 露天矿无人驾驶

露天矿开采一直以来面临着安全、效率、驾驶

员短缺等方面的挑战：

(1) 露天矿通常位置偏远、环境恶劣，招人难、人力成本高；

(2) 传统作业需要频繁换班，人工调度经常存在矿卡排队等待情况，影响生产效率；

(3) 日常生产作业存在落石、碾压、碰撞以及塌方等重大安全隐患。

露天矿无人驾驶系统的应用将有效解决以上问题。露天矿无人驾驶以智能驾驶计算平台、融合感知系统为基础，结合人工智能、5G、云控平台、高精地图等技术，可实现矿用卡车自主行驶、精准停靠、自动装卸、停车避让等功能，支撑露天矿采装、运输、排/卸等环节的全天候全流程无人化作业，有效提高矿车作业效率，减少危险区域作业人员数量，提升矿区生产作业的安全性。同时，矿卡可采集车辆自身状态、运输状态，实现矿用卡车保养提醒、车况检查、故障报警、车队管理等功能，有效提高矿用卡车营运效率，降低成本和事故风险。

▶ 3.3 智慧钢铁

钢铁行业是中国国民经济的重要基础产业，为国家建设提供了重要原材料保障，有力支撑了中国工业化、现代化进程。2022年2月，工信部、国家发展改革委、生态环境部联合发布《关于促进钢铁工业高质量发展的指导意见》（以下简称《意见》），明确钢铁工业实现高质量发展的目标和路径。《意见》中明确大力发展智能制造是主要任务之一，并要求推进人工智能、5G、工业互联网等技术在钢铁行业的应用，开展智能制造示范推广，打造一批智能制造示范工厂。

钢铁工业具有生产流程连续、工艺体系复杂、产品中间态多样化、大型高温高压设备集中、人员安全要求高等特征，面临着严峻的资源、市场、环保、竞争等挑战，亟需通过人工智能、云、5G等先进技术及场景化创新应用，提升行业的绿色环保、安全保障、生产效率。钢铁行业生产演进路线，如图3-7所示。



图 3-7 钢铁行业生产演进路线

当前，大部分钢铁企业已经完成了自动化建设，进入信息化和数字化的建设阶段，并通过信息化和自动化项目提升了生产运行效率。但是信息化的建设由于缺乏整体规划，基于装置和业务板块进行垂直的系统建设，无法更进一步支撑多业务、多系统、多数据的综合分析、联动和决策，停留在信息化阶段，无法进入数字化阶段；通过统一数字平台的建设，将原有系统的数据进行统一的收集、清洗、入湖，使数字化变革更加体系化、集成化和协同化；如何使得数字资产更有效应对行业挑战，亟需 AI 技术的大规模应用。



图 3-8 钢铁行业人工智能应用场景

如图 3-8 所示，钢铁行业生产流程长，工艺复杂，存在海量人工智能应用场景。但由于存在数据样本少、算法通用性差、开发门槛高、无持续迭代优化能力的缺点，导致在钢铁行业中 AI 无法大规模应用，亟需新的 AI 开发和应用模式。



图 3-9 大模型支撑钢铁场景工业化生产

基于华为盘古大模型的视觉大模型、预测大模型、多模态大模型和 NLP 大模型，以及工业智能中枢的工作流，可以帮助钢铁企业实现 AI 应用的高效开发和大规模部署，结合边学边用，实现越用越准。目前华为已经结合钢铁企业基于大模型实现智能配煤、废钢智能判级、行车智能调度、自动转钢、智能精炼等智能化应用，加速钢铁行业的智能化升级。

从各个应用场景的技术维度来看，AI 往往作为核心技术，与 5G、大数据等技术相结合，共同形成更为强大的应用能力，最终推动钢铁智能化转型，带来新的商业模式和机会。预计未来 3~5 年，铁钢行业人工智能市场空间规模达到近 200 亿。

3.3.1 智慧焦化配煤

吨钢成本中焦炭所占的比例达到 38%，与铁矿石相当，焦化厂急需在保证焦炭质量要求的前提下，综合考虑焦炭价格和化工产品综合收益，寻找最优配比使得利润最大化。通过人工智能技术，将配煤大数据与配煤机理结合，并融入

专家经验，深度挖掘原料煤之间的配伍性和特征相关性，结合业界先进配煤理论科学配煤，准确预测焦炭质量，通过分析对焦炭质量劣化情况及时给出预警，且能在达到焦炭质量要求的前提下，优化原煤配比，降低配煤成本，平均每吨焦炭降低 5 元~20 元。

3.3.2 废钢智能判级

传统定级作业环境恶劣，定级员常常需要多次攀高，与重型机械设备协作，人工近距离观察和检测，劳动强度大，作业风险高。同时，由于废钢类型繁杂，定级参数多等特点，往往只能依赖人工经验综合判断，受个人差异、抽检次数等因素影响，还存在定级结果不准确的情况。通过计算机视觉+AI，实现统一客观准确判级，降低定级的人为依赖，减少人为影响造成的定级偏差损失。部署废钢智能判级应用后，可实现废钢远程集控管理，减少人员现场工作安全风险；判级过程全程可视可回溯，减少判级结果争议。

3.3.3 行车智能调度

目前钢铁企业的炼钢车间冶金吊行车作业调度主要依赖人工调度。调度员基于预先制定的生产计划，通过视频监控现场运行情况，通过固定电话或对讲了解各工位生产实际情况，再进行调度决策。决策指令通过对讲机、固定电话等方式下达到行车操作人员和各岗位工作人员，实现行车吊运钢包任务分派，以及炼钢各个环节的串联。现有人工调度严重依赖调度员的经验和能力，由于炼钢过程高度复杂且生产过程持续动态变化，调度员难以全面感知各个环节中的情况，调度决策无法全面考虑后续各个环节，会造成工艺环节衔接不顺，从而导致生产过程中不必要的等待，生产效率下降。行车智能调度综合考虑炼钢各环节生产计划，实时采集分析炼钢厂转炉、精炼、连铸各工位的生产数据，结合行车和钢包实时位置信息、各工位生产工艺规则和行车的检修、设备异常等情况，通过人工智能技术，整体统筹安排行车运行规划，智能生成行车调度指令。生成的指令采用图形化界面和语音播报等方式下发到对应生产工位。对于生产过程中的动态变化，可以在 1 分钟内完成未来 30 分钟的决策并完成指令下发。行车智能调度系统的应用，可以有效提升钢包周转率，减少冶炼过程中的钢包传搁等待时间，降低过程温降，减少钢企在炼钢环节的吨钢成本，提升炼钢厂运转效率和产能。

3.3.4 自动转钢

宽厚板粗轧环节需要进行多个道次往复轧制，将固定规格的钢坯轧制到目标厚度和宽度。粗轧过程采用的横轧 - 纵轧方式需要对红热钢坯在多个道次进行 90 度旋转，当前由人工操作控制粗轧机前转钢辊道区域的两组锥形辊做不同方向旋转，从而带动红热钢坯旋转 90 度，旋转完成后再送入轧机进行轧制。转钢环节是

宽厚板粗轧生产效率提升的瓶颈。这个环节中红热钢坯亮度高，还有除鳞水气干扰，人工长时间目视操作容易疲劳。通过视觉采集转钢区域图像，采用高光抑制、水气去除等算法保证图像质量，实时智能识别钢坯位置、角度，AI 实时分析，并基于分析结果实时对接轧机控制系统，控制转钢辊道的转向转速，实现板坯转动精准控制；自动转钢系统还实现了结合电子围栏的实时控制，确保旋转过程中钢坯始终在转钢区域内，保障生产安全。通过大量生产实绩统计，该系统投用后实现平均每次转钢时间比人工操作减少 30%，转钢过程完全自动化，极大提升了生产效率。

3.3.5 智慧能耗管理

钢铁厂的能源系统复杂，既有一次能源消耗（焦煤、喷吹煤），又有二次能源消耗（煤气、蒸汽、余热、余压），还有自发电和外购电。如何管好能源，看清楚能源供需与生产之间的关联，从而找到优化和平衡的方法，是很多钢厂关心的话题。通过实时采集产线上的煤气、水、电、压缩空气、氮气等能耗，结合产线生产实绩，借助 AI 大模型建立起能耗需量的基线模型和预测模型。从单钢坯耗能预测出发，利用 AI 技术，分步骤实现物质流和能量流在车间级（局部）和工厂级（全局）的协同优化，在保证生产按质按量完成的前提下，实现能耗最优，成本最低。

3.3.6 智能精炼

通过大数据、云、AI 等技术，根据物理平衡、热平衡原理，以机理 + 数据模型相结合，实现合金辅料添加自动计算、LF 炉送电吹氩自动规划、钢水成分实时预测，避免人工操作带来的随意性和不确定性。通过智能精炼的实施，可以实现质量最优（钢水成分内控合格率

99.5%+)、效率最高(LF 精炼终点温度一次命中率 95%+)、成本最低(吨钢合金、渣料、用电成本减少 2 元,以年产 400 万吨钢炼钢厂为例,每年节约炼钢成本 800 万元)。

▶ 3.4 智慧银行

金融业是指经营金融商品的特殊行业,它包括银行业、证券业、保险业等。金融业在国民经济中处于牵一发而动全身的地位,关系到经济发展和社会稳定,具有优化资金配置和调节、反映、监督经济的作用。金融业的独特地位和固有特点,使得各国政府都非常重视本国金融业的发展。

与传统人工处理业务相比较,人工智能在金融领域中的应用有诸多独特优势,例如其强大的数据处理能力以及自我学习能力,更好地处理大量数据,进行精准的风险评估,提供个性化的服务,以及实现更高效的决策制定。另外金融领域数据复杂多变、易受外部环境影响,数据为各种政策的制定实施提供了强有力的理论支持,时时掌握最新的金融数据,把握经济动态走向显得尤为重要,利用深度学习算法模拟人类大脑的思维过程,就能够对数据进行及时响应并预警。

数据是数字经济时代的核心生产要素,围绕数据要素的生产转化,释放数据生产力,是各行各业当面要面对的问题。银行作为数字化水平领先的行业,已经识别到了客户需求和市场的变化,各银行机构纷纷启动了数字化转型战略,推进“一切业务数据化,一切数据业务化”的建设,加速实现从数据到数据资源的转化。

数据的价值在于应用,而应用的关键在于行动。从数据化到行动化,一般需要经历四个阶段的跃升:

数据化:这是基础阶段,区域银行要做好数据的采集、存储和管理,保证数据的完整性和准确性。

信息化:这是提升阶段,区域银行要通过数据分析和可视化,将数据转化为有价值的信息,为决策提供依据。

知识化:这是深化阶段,区域银行要构建知识图谱,深入挖掘客户画像、关系和行为,为业务创新和个性化服务提供支持。

行动化:这是高级阶段,区域银行要利用 AI 技术优化业务流程,智能响应市场变化,提高业务效率和客户满意度。



图 3-10 金融科技演进路线

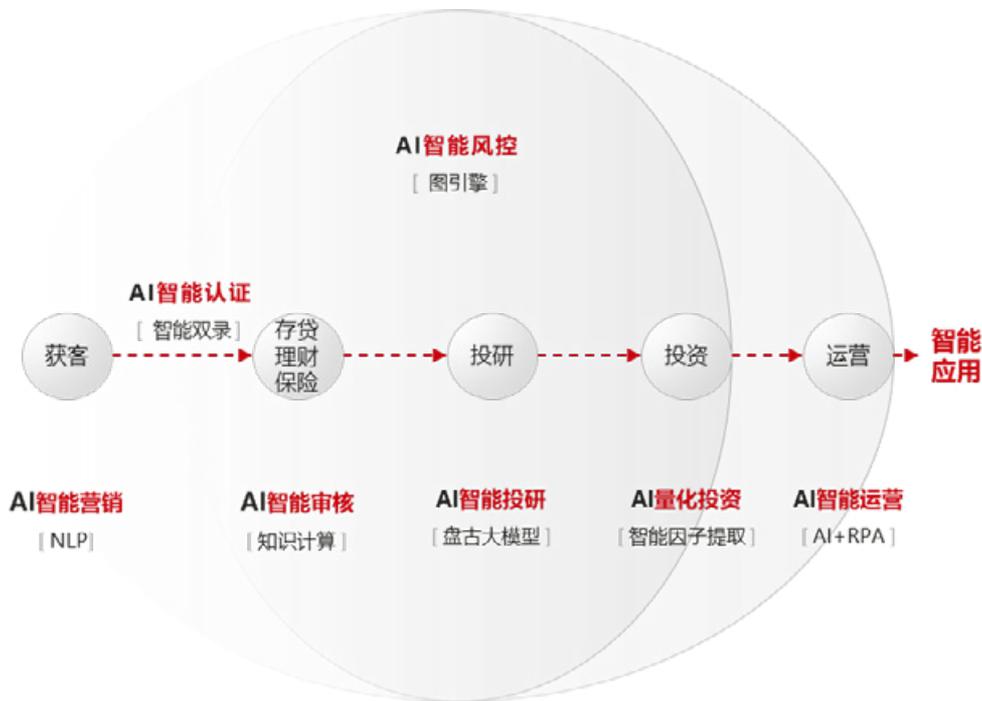


图 3-11 银行人工智能应用场景

AI 能力正在重新定义银行业生产力，大量银行已广泛应用 AI 生产力进化思路框架，建立持续迭代升级的闭环体系，适应 AI 技术的持续突飞猛进。AI 在银行业的应用场景非常广泛，在提升金融产品营销能力、提升风险处置能力、提升运营与开发效率、使能新业态四个方面已经显现出不同的价值。

在提升金融产品营销能力方面，AI 技术可以帮助银行实现客户画像的精准构建和个性化推荐和定制化服务。基于对客户画像的深入理解，AI 系统可以自动推荐符合客户需求的金融产品和服务，营销时客户圈选就能够更加有针对性，提高营销精确性。

在提升风险处置能力方面，通过“信用报告自动生成”、“舆情监测”等风控领域 AI 应用可以提升银行的风险防范与处置能力，使其能够加大面向中小微实体经济的资金要素供给，逐

步成为区域经济发展的新动力，切实提高金融服务实体经济效率和支持经济转型发展的能力。

在提升运营与开发效率方面，通过人机交互打通智能客服，智能网点等服务界面的业务阻隔与流程断点，实现“端到端”业务流程触发、路径选择、任务派发与质量监测，提高业务处理标准化、智能化水平，大幅提升银行客户体验。如工行在全国有四万多个网点，20 万多个网点员工，通过与华为盘古大模型合作，让每个银行柜员拥有自己的智慧助手。根据客户问题生成操作指引，将以前 5 次操作缩短为 1 次，单次办结时间缩短 5 分钟。

在使能新业态方面，AI 结合 IoT、区块链等新技术，将直接服务对象由“人”延伸到“智能物”，构建以场景为核心的业务模式，以及基于客观数据信用体系的风险管理模式。从而使得基于动产可信的质押模式成为银行的新业态。

| 应用类型 | 业务前中 | | | | 业务中 | | | 业务后 | |
|-----------|---|---------------------------------|---|-------------------------------------|--------------------------------------|---|--|--|--|
| | 客户服务 | 投顾 | 投研 | 核保理赔 | 营销 | 风控 | 数据分析 | 研发 | 内控合规 |
| 当前场景智能化现状 | 当前以“搜索”为核心应对用户疑问，重在快速解答 | 大数据对投资者画像与风险洞察，专家规则自动推荐投资产品与组合 | 实现数据资产以及非结构化数据洞见赋能，并利用知识图谱等技术，实现静态分析 | 从依赖人工审核，向利用核赔、理赔、控费等规则引擎 | 实时+数据驱动营销，营销内容以模板库建设与积累为主，营销策略依赖专家规则 | 专家规则+特征工程建模，AI应用于身份识别、贷前辅助审核等，全流程智能风控仍在探索 | 基于数据自助分析平台，支撑业务人员通过可视化、拖拽的方式进行数据分析 | 应用敏捷开发模式进行软件和产品研发，但满足研发敏捷性、业务合规性与客户个性化需求的探索仍处于初期 | 普遍运用RPA技术实现文件关键信息的抽取与分类归档，但与文件理解相关的任务需寻求更多技术支持 |
| 大模型+的探索 | 结合多轮对话与KYC探查能力，有效提升用户对话体验；但需完备的伦理、合规等内容审查策略 | 基于财富管理专业知识进行预训练模型，建立围绕个体的智能投顾服务 | 通过海量投资标的的信息分析，实现对各类投资数据大模型预训练的分析以及时间序列数据的预测 | 提升信息抽取、聚类与分析的实时效率，提升核保理赔用户体验，降低人力投入 | 营销物料快速生成，支撑个性化营销；基于大模型KYC的智能营销 | 风控全流程引入大模型进行新特征规则发现，结合因果学习探索智能风控策略制定与追因 | 大模型将业务问题转化为数据分析问题，进一步转化为SQL，从人人用数据走向AI助理分析用数 | 大模型协助制定符合行业标准的软件研发计划、代码编写与测试等，提升研发敏捷性；自动生成多样化产品定价与组合策略 | 大模型形成对文件的整体认知与理解，大幅降低对人工审核的依赖，同时给出风险提示以及解决方案 |

图 3-12 银行大模型应用场景

随着预训练基础语言大模型初具与真人相似的交互聊天、文稿撰写、逻辑推理、计算机编程以及综合性思考能力，其展示出强大的通用智能潜力，有可能引发新一轮的技术竞赛、产业革命，甚至改变未来人类的生活方式。大模型的核心优势在于其高认知水平、强泛化能力、深厚知识储备、有逻辑推理和有温度的交互。在多轮对话中，它们的交互能力实现了飞跃性的提升，意图识别的准确率也更高。大模型的相关技术潜在应用场景广泛，例如在财富管理领域，模型可以辅助生成研报、理解研报内容并脱水，还能根据外部各类信息解析出 alpha 因子；在客服服务方面，模型可以提高知识库的丰富度，逐步实现个性化服务；在营销领域，模型可以辅助生成话术，提升数字人的交互能力；在运营领域，模型可以提高智能化审核质检能力，并逐渐演化成多模态超级 RPA 的能力；在系统研发领域，模型可以辅助生成代码、监测代码质量，并提高自动化测试的能力；在办公领域，模型可以辅助记录会议、编写报告和 PPT，从而提高工作效率。大模型将激活银行的非结构化数据，更高效地释放数据价值，应用将渗透到业务前中后段，带来新的生产力

升级。经过市场洞察和测算，未来 3 年内大模型将加速金融科技智能化升级，带来银行科技的市场空间约为 260 亿，主要包括 AI 新算力集群建设、金融行业大模型建设和相应的实施服务。

3.4.1 智能营销 – 提升金融产品转化率与用户体验

传统的营销模式依赖传统媒体渠道，缺少与银行客户间的互动和即时反馈，目标客户营销精准度不高，很难评估营销活动对产品销售和品牌宣传的影响，效能较差；对于客户而言，没有参与感，处于被动获取信息的状态，无法满足个性化需求。随着金融科技的快速发展和竞争的加剧，银行正重构营销体系，以客户为中心，进行客户全周期管理，提供更加精准的产品和更优质的服务体验。通过利用大数据分析和机器学习算法，银行可以深入了解客户的消费习惯、偏好和行为模式，从而更好地把握市场趋势和客户需求。

AI 时代银行数字化营销将实现客户需求与商机

主动挖掘，人机互融，带给客户沉浸式极致体验。人工智能技术可在多个银行数字化营销场景中广泛应用，营销内容生成，AIGC 改变文案与图片产生方式，快速生成不同风格营销内容，已达到专业文创水平，改变了银行营销活动策划与推广模式，大幅提升生产效率与活动创意。数字人交互，AI 生成虚拟人物，通过大规模语料学习，实现客户问题自主回答，满足多样化场景，提升用户体验与沟通效率，提升复杂缠绕问题解决率，减少大量人工客服工作。产品推荐，传统数字营销方式依赖专家规则进行用户与产品匹配，无法进行精细评估。人工智能与大数据技术融合，通过银行全域数据、全量用户、全产品数据学习训练，获取最优参数模型，帮助金融机构实现客更精确地识别出潜在的高价值客户和目标群体，主动识别客户需求，根据客户的消费习惯和偏好，自动精准匹配产品。商机挖掘，通过对海量数据的学习分析，AI 算法模型可以发现潜在的市场机会和趋势，帮助金融机构及时调整营销策略。例如，通过对社交媒体平台上的用户评论和反馈进行分析，AI 算法可以预警客户对产品或服务的满意度和需求变化，从而帮助机构快速做出反应并进行相应的改进。

3.4.2 智能风控 – 风险传导识别

银行的风险传导识别场景是指在金融市场中，银行通过对各种风险因素的识别和分析，评估并预测可能对银行业务产生不利影响的风险传导路径。这些风险传导路径可能包括信贷风险、市场风险、操作风险等，它们通过不同的渠道和机制相互影响，可能导致银行的资产负债表出现不良变化，进而影响到银行的盈利能力和稳健经营。

随着金融市场的复杂性和全球化程度的增加，传统的人工识别方法已经难以满足需求，而通

过知识图谱建出复杂的风险传导模型，结合 AI 智能检测算法，识别信贷风险的传播模式和传导路径，预测未来可能出现的风险传导路径和影响范围，降低风险的扩散率。体系化地覆盖客户预筛、事前审查、事中决策、事后预警等全流程多方面，提升银行的风险防范能力。

3.4.3 智能风控 – 实时智能反欺诈

金融欺诈是指以非法占有为目的，采用隐瞒真相或虚构事实的欺诈手段，用以骗取公私财物或者金融机构信用、破坏金融管理秩序的违法犯罪行为。随着科技的发展和普及，产生了一系列新型的金融欺诈行为，如网贷平台欺诈、大数据精准欺诈等，逐渐形成了“黑色产业链”，所带来的社会危害也在不断加深，亟需借助更加智能化的手段来实现对欺诈交易的精准识别，更加高效地识别和预防欺诈行为，保障交易的安全性和可靠性，降低对正常用户的打扰。

首先，通过机器学习和深度等智能算法，构建更加复杂的分析模型从而更精准的识别欺诈行为。传统的欺诈检测方法往往依赖于专家经验设定的规则和模型，而人工智能技术可以通过机器学习和深度等智能算法，从海量的交易数据中自动发现异常模式和规律，迅速识别出异常的交易模式，从而更准确地识别出潜在的欺诈行为，预警机构及时采取措施进行干预和阻止。

3.4.4 智能运营 – 智能票表卡章识别

金融行业日常业务中涉及大量表单、凭证、票据和图片等材料，如对公对私开户资料、信贷业务资料、客户财务报表，运营票据、合同、档案等。由此带来大量重复的、低效率的录入、核对工作，需要大量的人力成本是金融行业面临的普遍痛点。

随着金融科技快速发展，金融机构对于高效、准确的文字识别需求日益增加。通过结合 OCR 文字模型和 NLP 自然语言处理模型，使用深度学习算法对大量的金融文档进行训练和学习，从而建立起一个强大的文字识别模型，能够自动地分析和理解文本内容，并准确地将其转化为结构化数据，即可将金融机构可以将大量的纸质文件、电子文档等转化为可编辑、可搜索的电子数据，可以实现对金融文档的自动分类、关键词提取、情感分析等功能，方便后续的数据分析和处理工作，能够有效的为其节约人力、物力成本，提升效率，增加营收。随着大模型技术发展，构建的金融 OCR 大模型，可以通过高精度文字检测与多模态表格还原等技术，可以大大提高手写体票据文字识别的准确性和效率，同时实现用一个模型覆盖多个通用的文字识别场景，更进一步降低应用场景拓展门槛。

3.4.5 智能运营 – 智慧客服

银行通过设置 7*24 小时的线上客服，通过电话、短信、邮件或在线聊天等方式，为客户提供全天候的金融咨询、业务办理和问题解决服务，能够快速准确地解答客户的疑问，协助客户完成各种银行业务，如查询账户余额、转账汇款、信用卡还款等。在处理客户投诉和收集建议的过程中，确保客户满意度。

随着科技的发展，许多银行还推出了智能客服系统，通过人工智能技术提供更高效、便捷的服务。通过应用 AI 自然语言处理和机器学习等算法，模拟人类客服的工作方式，为客户提供高效、准确、便捷的服务。智能客服可以理解并回应客户的查询，如账户余额、交易记录、贷款申请等。同时还可以根据客户的问题或需求，提供个性化的建议和解决方案。此外，智能客服可以 24 小时全天候在线，无需等待人

工客服的接听，大大提高了客户的满意度和银行的服务质量。

3.4.6 智能运营 – 智慧网点

在线上 + 线下渠道全面融合的趋势下，银行网点仍然有不可替代的作用，银行线下网点必须向更高效服务于全渠道销售转型。要考虑基于边缘计算技术、AI 技术、信息互联网技术、多媒体技术等多方面技术为银行网点智能化提供技术支撑，提供更好的服务。

网点多种场景对智能分析能力提出了很高的要求，借助智能化、自助化的设备实现银行的营运、合规、安全等场景的实时监测，从而提升用户的体验；借助智能化服务和安防系统，降低网点相关业务的人员成本。具体的智能分析相关服务包括：1) 客群分析为不同客群的客户提供差异化信息服务；2) 客流分析提供布局优化、人员调配的科学决策支撑；3) 服务分析，提升合规性、效率和质量；4) 区域监测，异常行为智能识别，实现及时告警等。

多媒体技术可以提升到网点办理业务的客户智能互动体验。网点保留的最重要职能就是和客户的物理接触体验，线下网点配置交互式多媒体体验设备，如全息投影、VR 技术，给客户提供内容丰富生动的服务和展示平台。借助智慧办公新物种的全新生产力工具，集白板、投影、麦克风、音箱、视频会议终端等设备于一身，加持多种智能技术，让智能协作水平满足网点智能化的要求。

3.4.7 新业态 – 动产融资

据世界银行集团估算，中国存货规模有 75 万亿，未来五年将达到 100 万亿，动产融资市场存在巨大的可开发空间，但一物多抵、以次充

好现象时有发生，金融仓爆雷事件常有报道，影响动产融资业务的发展。传统金融仓通过人工监管成本高且无法识别高端作案手段，例如虚开凭据、重复质押、质押品真实性等问题，导致银行不敢贷不愿贷。

随着 AI 技术发展，金融仓建设有了新的解决方案。例如运用物联网、以大模型为代表的人工智能，和视频监控等技术实现对仓库环境及资产状况的智能监测；应用人工智能技术结合大数据分析，基于历史数据与实时监测指标，智能识别异常事件或风险智能预警险，并针对运营人员进行主动预警通知；利于 CV 模型赋能摄像头实现远程运维，实现对系统定期自动巡检；例如机器学习算法对设备进行健康度分析、软硬件自动智能运维更新升级等，简化人工运维工作，提高运维效率与系统稳定性。

▶ 3.5 智慧证券

证券业是金融业的重要组成部分，主要从事证券的发行、承销、交易和资产管理等业务。证券业通过证券市场有效调剂资金，促进资金优化配置，服务实体经济发展。随着互联网技术、移动终端和大数据的广泛应用，证券业正在经历智能化转型，人工智能技术为这一转型提供了有力支撑。

人工智能可以深度参与到证券业的各个环节，实现智能化操作。在投资决策环节，可以建立智能投顾系统，利用深度学习算法对交易数据和市场趋势进行复杂的模式识别，辅助投资者进行资产配置和投资决策。在股票分析方面，可以使用自然语言处理技术解析海量的财经资讯、研报和社交数据，辅助判断个股质量。在量化交易环节，可以构建智能交易系统，利用强化学习等算法进行策略优化，实现自动化交易。在风险管理环节，可以使用机器学习算法进行市场风险和信用风险的评估。

在证券监管方面，监管部门可以运用人工智能技术进行智能监管。例如，监管部门可以建立智能分析系统，利用机器学习等技术，对海量的交易数据、文本数据、图形数据等进行智能

分析，实现异常交易的自动识别和市场滥用行为的智能监测，提高监管效率。

人工智能是证券业数字化转型的关键技术。它提高了投资决策、交易执行、风险控制和监管效率，创造了更优质的用户体验，将助力证券业由高速增长向高质量发展转变。如果运用得当，人工智能的应用将使证券业更智能、更高效、更安全、更公平和更普惠。预计未来 3~5 年，中国证券行业人工智能的市场规模预期会达到 180 亿元，其中大模型带来的增量空间预计约为 30 亿元。

3.5.1 智能投顾

证券行业的投资顾问业务是指证券公司等专业机构通过提供投资分析和建议，帮助投资者进行投资决策的金融服务。投顾业务旨在为投资者提供专业的资产配置和投资建议。传统投顾服务主要依靠人力来完成，包括市场分析、资产配置和客户服务等，面临多个痛点。首先，人工投顾需要大量时间和精力去分析复杂的市场数据和投资工具，这导致高昂的人力成本并限制了服务时间。其次，由于依赖个人经验，

传统投顾服务常受到分析师主观判断的影响，导致投资建议的质量参差不齐。

随着科技的发展，投资顾问服务也正在逐步智能化，并产生了“智能投顾”服务。与传统投顾相比，智能投顾通过应用人工智能和大数据分析，显著提高了投顾服务的效率和准确性。首先，它使用机器学习算法进行资产配置和选股，能根据投资者的个人需求和市场状况实时调整投资组合。其次，自然语言处理技术使得智能投顾能理解投资者的需求并提供个性化建议。此外，智能投顾还能实现7×24小时的全天候服务，大幅降低人力成本。智能投顾带来了更高效、更个性化的投资体验，正逐渐成为投资领域的新趋势。

3.5.2 智能投研

投资研究是证券公司的一项核心业务，通过对上市公司进行深入调研，预测其未来经营状况和发展前景，进而做出投资推荐或风险提示。传统投研面临多个挑战：首先，分析师需要花费大量时间进行数据收集和分析，导致工作效率低下。其次，海量的信息容易导致信息过载，增加决策难度。再者，过度依赖个人经验和主观判断可能导致认知偏见。最后，撰写高质量研报的工作量大，且难以保证内容的质量和一致性。

智能投研是指运用大数据和人工智能等先进技术来对投资研究过程进行改进的过程。相较于传统模式，智能投研通过应用人工智能和大数据技术，有效地解决了投研过程中的诸多问题。首先，通过自然语言处理技术，智能投研系统能迅速分析大量非结构化数据，如公司公告和行业报告，提高分析师的工作效率。其次，使用机器学习算法，智能投研系统能找出影响股价的关键因素，并辅助分析师制定更全面的投

资策略。此外，新兴的大模型技术进一步提升了投研的智能化水平，比如自动生成研报初稿，大幅提升研报撰写效率。在使用智能投研系统时，特别需要注意确保其输出符合道德和合规要求，以维护市场的公平性和透明性。

3.5.3 财务报表智能审核

对企业财务报表进行分析和审核是证券公司投行业务的重要一环。企业财务报表不仅反映企业的财务健康状况，还为投资者和监管机构提供了洞察企业运营的重要信息。然而，问题也随之而来：企业财务报表常常出现虚报或瞒报的情况。一些企业通过控制财务数据，企图欺骗投资者或规避监管。传统的解决方案，如人工数据核查，不仅需要高度的专业知识，而且工作效率低下，难以有效地识别和预防财务风险。

借助人工智能技术可以实现财务报表的智能审核和异常识别。具体来说，通过机器学习或深度学习的技术构建和训练智能审核模型，能够从历史财务数据中学习并识别异常模式。在实际操作中，系统能自动解析企业报表的数据，并依据模型来判断是否存在如虚报或瞒报等异常行为。更进一步，系统还能识别潜在的舞弊动机，比如偷税漏税或业绩舞弊。与传统的人工方法相比，智能审核模型能高效地处理大量数据，快速地识别异常并发出风险预警。这不仅大大提高了识别财务风险的效率，也为投资者和监管机构提供了更精准的决策依据。

▶ 3.6 智慧电力

电力系统主要由发电、输电、变电、配电、用电及调度组成，将各种一次能源转换成电能并输送和分配到用户，是国民经济的基础产业。电力系统发展主要经历了三个阶段：半自动化，自动化，智能化。生产系统从二遥到五遥再发展到现在的智能化交互，管理业务从人工到全面数字化。以数据为基础构建强大的软件平台，以软件定义电力系统，将传感技术、信息技术、计算技术、控制理论、人工智能与电力系统深度融合，推动实现电力系统的智能化。电力生产业务从半自动化向智能化演进过程，如图 3-13 所示。

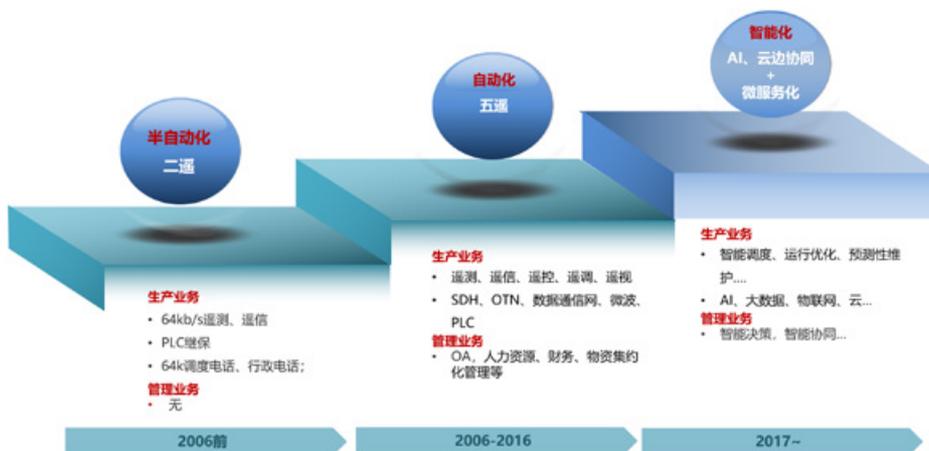


图 3-13 电网业务发展历程

2021 年 3 月 15 日，党的二十大报告强调加快规划建设新型能源体系，为新时代能源电力高质量发展指明了前进方向。能源数字化、智能化建设，是实现传统能源向新型能源体系转型的关键。

电力是重资产行业，系统非常复杂，电力设备是电力网络的重要组成部分，设备的状态直接关系到电网的稳定运行。随着新型电力系统建设的加速推进，电力系统越来越复杂，设备数量的不断增多，传统的管理与检修技术已经不能满足需求，如何有效感知设备状态、降低检修成本、提高设备使用寿命、提升技术人员的检修效率和质量是目前所面临的问题。如图 3-14 所示，电力巡检业务发展经历了人工为主、人工 + 机巡到现在的智能立体巡视阶段，智能巡检对电力设备进行预判性的检测及维修是保障设备安全、电网稳定运行的重要手段。

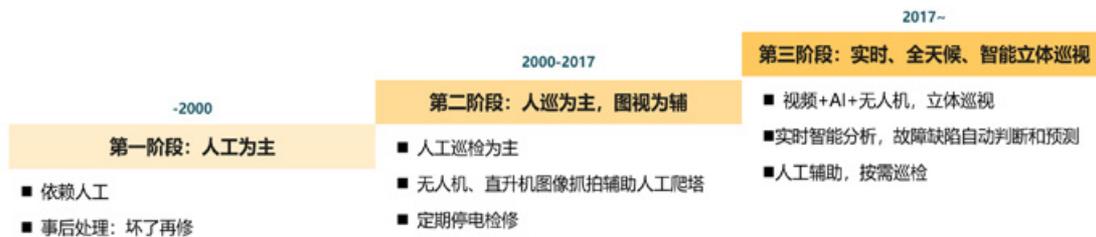


图 3-14 电力巡检业务发展历程

缺少智能巡视和智能分析的数字化手段，难免会造成电力设备安全风险未能及时监测和预警，从而影响运维人员对异常及时响应和处理的实时性和精准性。以物联、视频、AI 为代表的数字化技术可以为电网设备提供全天候、全方位、全自主的实时监视，实时掌控设备运行状态，及时识别设备异常风险，保障整个电力系统运行安全，减轻运维人员工作负荷。

AI 技术的创新和突破为电力设备的智能巡视带来了新的机遇。其强大的数据分析和推理能力，能快速检测异常信息，识别潜在风险，实现实时告警、实时监测及预警。基于 AI 视觉识别技术，结合视频监测等视频图像采集设备，运用 AI 视觉深度学习算法，实现电力设备状态实时监测、站端表计智能识别、环境智能实时监测、人员安全行为智能评估等智能巡检，提升了智能化巡视水平。特别是 AI 大模型的出现，进一步提升了 AI 技术的发展水平和速度。AI 大模型是预训练的模型，参数为数十亿甚至千亿元以上。大模型采用“工厂式”开发，泛化性好，行业用户二次开发门槛低，多个场景一个模型，上线快，开发周期为天级。对于开发者来说，通过大模型可以持续沉淀经验，沉淀能力，沉淀资产。

电网各场景积极拥抱 AI 技术，智能化进入规模应用阶段，迈入“场景 +AI”时代，电网场景包括：输电、变电、配电、用电、调度，如图 3-15 所示。



图 3-15 电网业务场景

大模型在电力行业的应用，其中基础大模型 L0 拥有百万级到千亿级别参数，结合场景可灵活调用基础大模型。在电力行业主要 CV 大模型、NLP 大模型、多模态大模型的应用，这些模型具备强大的泛化能力，通过与行业知识结合，能快速实现不同场景的适配，少量样本也能达到高精度，基于预训练 + 下游微调的工业化 AI 开发模式，加速 AI 行业应用。伙伴基于行业大模型之上，开发适配行业场景模型应用，提升电力发、输、变、配、用各环节 AI 应用实效，如图 3-16 所示。



图 3-16 电网业务大模型架构

随着新型电力系统发、输、变、配、用等领域的技术及管理变革，电网正演变为源、网、荷、储等多重因素双向互动，人工智能将发挥越来越重要的作用。预计未来 3~5 年，电力行业人工智能市场空间规模近 200 亿。

3.6.1 输电线路通道入侵识别

线路通道巡检重点针对线路通道、周边环境、“三跨”线路（跨铁路、跨公路、跨重要输电通道）、施工作业等情况进行检查，及时发现和掌握通道环境的动态变化情况。同时，需根据输电通道性质、地理气象环境条件等实际情况，对通道环境上的各类隐患或危险点安排定点检查，避免各种隐患因素对线路安全稳定运行造成威胁，产生事故。

传统线路通道巡检主要依赖人工，巡检劳动强度大，巡检质量会受到作业人员主观因素影响，在恶劣环境下巡检范围受到很大限制。目前架空输电线路巡线工作面广、线路长、沿线环境复杂，缺乏先进的辅助巡检装备支撑，对一些较为隐蔽的缺陷、隐患及故障点很难准确查找定位，增大了线路运行风险。

输电线路通道入侵识别主要部署在外力破坏易发区（跨高铁、跨高速公路、跨重要输电通道、违章建房、开山炸石、吊车施工等外力破坏易发区域）、火灾易发区、易覆冰区、通道树木（竹）易生长区、偏远不易到达区和其他线路危险点、缺陷易发区段的线路通道。在线路通道部署摄像机，拍摄通道画面，通过 AI 分析算法，重点针对通道环境中的安全风险，如施工作业（大型机械）、山火、线路悬浮物（塑料薄膜）等进行识别。实现架空输电线路巡检信息化、巡视少人化和检测智能化，从而大大提高巡检效率，降低人工巡检工作量，减少安全隐患，满足日益精细化的业务管理要求。

3.6.2 输电杆塔本体异常识别

输电杆塔本体监测，重点对绝缘子、线路金具、拉线、鸟巢异物、杆塔基础、接地装置等设施的完好情况进行全面检查。传统采用人工巡检，

以定期巡检、故障巡检等方式，对塔体设备和基础等设施，进行逐一检查。随着无人机的普及应用，无人机自动巡检采集杆塔本体设施图像的方式，正成为主流。

无人机巡视拍照，结合人工识别隐患的模式，相较传统人工巡检有节省巡检时长、降低人员巡检风险等优势，但由于无人机拍摄图片量大，人工分析效率低，存在人员成本高、巡检人员工作强度高，分析易出现漏检、误检的问题。电力行业正在探索借助 AI 分析算法，对杆塔本体（包括杆塔、绝缘子、大尺寸金具、小尺寸金具、导地线、基础、附属设施）和接地装置进行缺陷识别，缺陷类型主要包括塔身缺陷、部件缺陷、异物等。借助 AI 智能识别技术，可准确识别塔体缺陷类型和相应特性，并大大提升塔体缺陷识别准确度和速度，减少误报率，最终实现 AI 智能巡视代替人机巡视。



▶ 3.7 智慧油气

从最早的人工挖井采油发展至今，油气行业已细分为勘探开发、油气生产、管网储运、炼化化工、成品油零售等多领域全流程工业。数字化技术的应用和发展，为传统油气行业注入了新的生产力，成为油气工业生产新的增长引擎。从上世纪 90 年代开始，传统油气行业先后经历了自动化、数字化、智能化发展阶段，并正在向着智慧化的方向持续演进。油气行业的数字化建设历程，如图 3-17 所示。



图 3-17 油气行业数字化建设历程

中国上世纪曾戴着“贫油国”的帽子，在千千万万“王进喜”式的油气工作者的努力下，才有了现今各大油田产量年年增高的好势头。油气作为国家战略储备资源，重要性不言而喻。党中央也多次对保障国家能源安全做出部署安排。党的十九届五中全会强调要“保障能源和战略性矿产资源安全”。“十四五”规划和 2035 年远景目标纲要围绕“构建现代能源体系”、“提升重要功能性区域的保障能力”、“实施能源资源安全战略”等做出了一系列重要部署。此外，油气资源对于维持国家稳定运行至关重要，对经济发展有直接拉动作用，也是支撑国家基础设施建设和社会保障体系的顶梁柱。

随着行业数字化转型的深入和数字技术的不断发展，AI 逐渐成为了油气行业智能化发展进程中的重要抓手。AI 技术的应用，切实地改变了石油工人在恶劣环境下繁杂的生产工作。例如，在勘探开发场景中，利用 AI 数据处理、知识图谱、机器学习、模型训练调优等能力，可实现地震层位的快速解释、地震初至波的自动拾取、水平井定向控制、井筒钻探数据处理、钻井及井喷事件风险预测等。在生产管理场景中，面对井场“点多、线长、面广”的分布状态，AI 通过神经网络的自主学习建立工况诊断模型，对抽油机井的工况进行实时诊断预警，并通过周界安防、行为分析等智能算法助力油气田的安全生产防护，大幅减少日常人工巡检的

工作量，让工作人员“穿着西装、坐在办公室里”也能实时了解现场的生产动态。

AI 大模型技术正处在系统创新、深度融合以及智能引领的重大变革期，AI 大模型技术与油气工业的融合必将带动传统行业生产能力和管理效率的指数级增长，从而进一步加快实现油田增储上产、提质增效，实现油气行业智能化转型升级的目标。预计未来 3~5 年，油气行业人工智能市场空间预计可达 40 亿。如图 3-18 所示，华为公司推出的盘古大模型，可基于通用的基础大模型，叠加行业数据，快速训练出油气行业大模型，然后再通过细分场景数据的反复学习、训练及微调，进而构建出场景大模型，如地震大模型、测井大模型等。

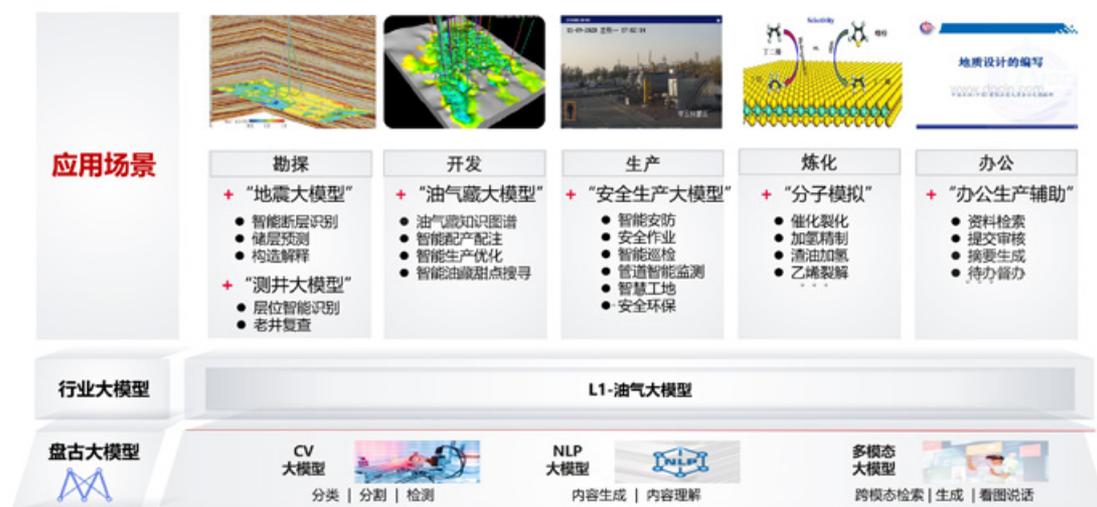


图 3-18 油气行业 AI 大模型能力和应用场景

3.7.1 油气井射孔压裂评价

射孔压裂是油气生产中常用的一种低渗透率油气储层增产技术，如图 3-19 所示，该技术是通过在油气井筒中进行射孔，将高压液体注入井筒，使岩石裂开，从而增加油气的产出量。射孔压裂空隙位置、大小和深度效果将直接影响采出量，因此射孔后需要对射孔压裂进行效果评价，类似给油气井出具一份 CT 报告，以查看射孔压裂是否达到最佳效果。评价内容包括射孔直径、周长、面积、圆度、裂缝分布和裂缝宽度等数据，确定是否需要后续的调整和优化，以期达到最佳油气增产采收量的目标。

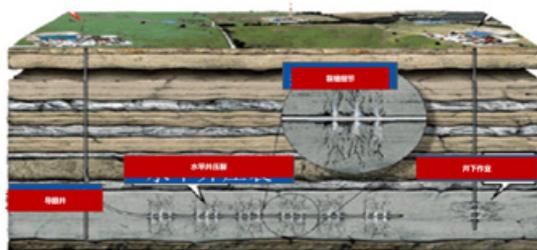


图 3-19 油气井射孔压裂剖面示意图

通常情况下，一个油气井筒中有多达几千个射孔，目前业内大多依靠人工看视频、看图像的方式进行手工标注、识别、量测、分析，工作量巨大、效率低，准确率也有待提升。

借助 AI 大模型技术的射孔压裂可视化 AI 评价方案，基于华为昇腾 NPU 算力、全场景 AI 计算框架 MindSpore 组建的 AI 评价平台，通过对媒体流进行云端模型训练、边端卷积推理，实现精准高效评价，射孔压裂精准可视，接箍识别、射孔簇识别的精确率可达 95%，射孔识别、射孔轮廓标注识别精确率可达 85%。相比传统方式，工作效率和准确率均得到大幅提升。

3.7.2 测井油气层识别

测井油气层识别是油气勘探开发中非常重要的一环，可以帮助勘探人员更准确地确定油气资源的分布和储量，从而指导油气勘探和开发工作。测井数据是通过在井孔中下放测井仪器，测量地下岩石的物理性质而得出的，如电阻率、自然伽马辐射、声波速度等，然后根据这些数据进行分析，确定地下岩石的类型和性质，从而判断是否存在油气层。

图 3-20 为油气井测井综合解释图，通过分析测井综合解释图，可以确定地层的岩性、油气储层性质、储层厚度、储量等信息，为油气勘探和开发提供重要的参考依据。

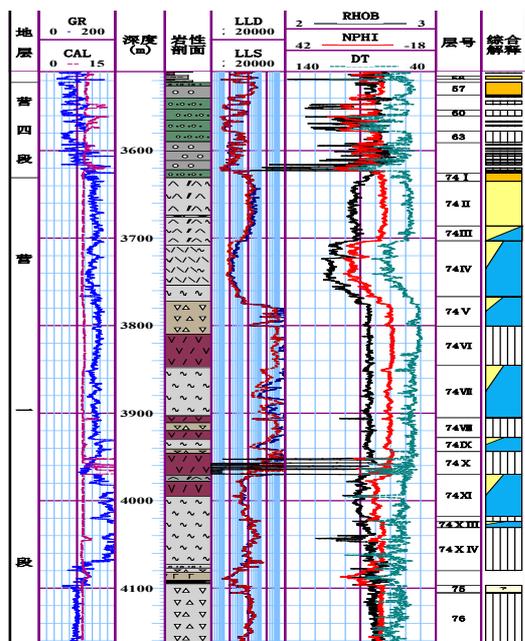


图 3-20 测井综合解释图

传统的油气层识别是采用测井资料进行人工识别和专业软件模拟解释，过程中依赖大量的专家经验，效率低、精确度不足。例如，探井解释一般需要一周左右，开发井需要 3 天左右，而区块老井复查一般需要多月完成。传统的油气层识别模型精度低、专家人力投入工作量大，严重影响油气开发开采效率。

融合知识图谱与深度学习等人工智能技术，油气田智能测井研究取得重大突破，在多个油田和区块通过人工智能测井解释发现遗漏油层。测井人工智能平台基于华为云 Stack ModelArts，全面适配通用算力和双栈 AI 算力，支持异构资源协同调度、云端协同。引入 NLP 生成式 AI 技术，在少量样本准备的基础上，可对测井资料实现高精度自动评价。对于评价过程和评价结论，以准确的测井术语和地质术语，以自然语言的方式进行专业质控，对异常信息可实现自动发现和评估分析，并自动生成报告，实现智能测井、增储上产。

3.7.3 抽油机井工况诊断

在油田生产中，抽油机是油井开采中使用最广泛的人工举升设备，主要由地面部分和井下部分两部分组成，即地面部分的电动系统、减速箱系统、连杆系统以及井下的抽油泵、抽油杆、

组合油管等。其中，抽油泵在井下上千米处工作，井下条件复杂，随着油田的长期开采，环境条件也会发生变化，抽油杆柱在井下往复运动的过程中会产生许多不明原因的磨损或故障，从而导致抽油效率降低，甚至造成油井事故。及时准确地了解抽油机的工作状况并进行综合诊断，对于保证抽油效率、降低机械采油成本、提高油井产量具有重要的意义。

目前虽然实现了对油井生产数据的采集，但还无法实现油井生产工况趋势的智能分析和预测。综合考虑油井生产动态、油藏物性、示功图数据、机抽设备等相关数据，可针对供液不足、结蜡等典型的渐变工况用 AI 手段进行预测分析。例如，示功图诊断方式主要是通过图像识别方式实现，先将示功图数据处理成二维灰度图像，再利用卷积神经网络对样本数据进行训练，并保存模型。使用实时生成的新示功图数据作为输入，通过模型预测所属分类，对示功图图像进行分类识别，实现示功图智能诊断的功能，从而实现抽油机井工况精细化诊断。工况诊断可有效发现、预测油井问题，为动态优化抽油机井并提供决策依据，确保油井长

期高效生产。

3.7.4 管道光纤监测预警

随着国家城镇化和油气管道建设进程的进一步加快，油气管道建设和运营的外部环境日趋复杂，油气管道隐患危及社会公共安全。部分油气管线营运多年，当年远离城镇的管道如今已被居民楼、学校、医院、工厂等建筑物包围，管道穿越人口密集区、安全防护距离不足、城市市政管网与地下油气管道交叉，尤其是第三方占压施工等情况时有发生，给国家能源安全和稳定供应造成诸多不安全因素。统计表明，中国 50% 以上的油气管道故障来自第三方入侵，包括意外机械施工、人工挖掘、打孔盗油等，单纯依靠人工巡检显然难以为继，运维压力不断增加，引入智慧巡线新技术，可提升管道的安全生产水平。

如图 3-21 所示，可利用油气管道同沟敷设的光纤，建立光纤预警系统，监测第三方活动信息，从而有助于加强管道的完整性管理，降低管道安全运行风险，是防患于未然的有力举措。

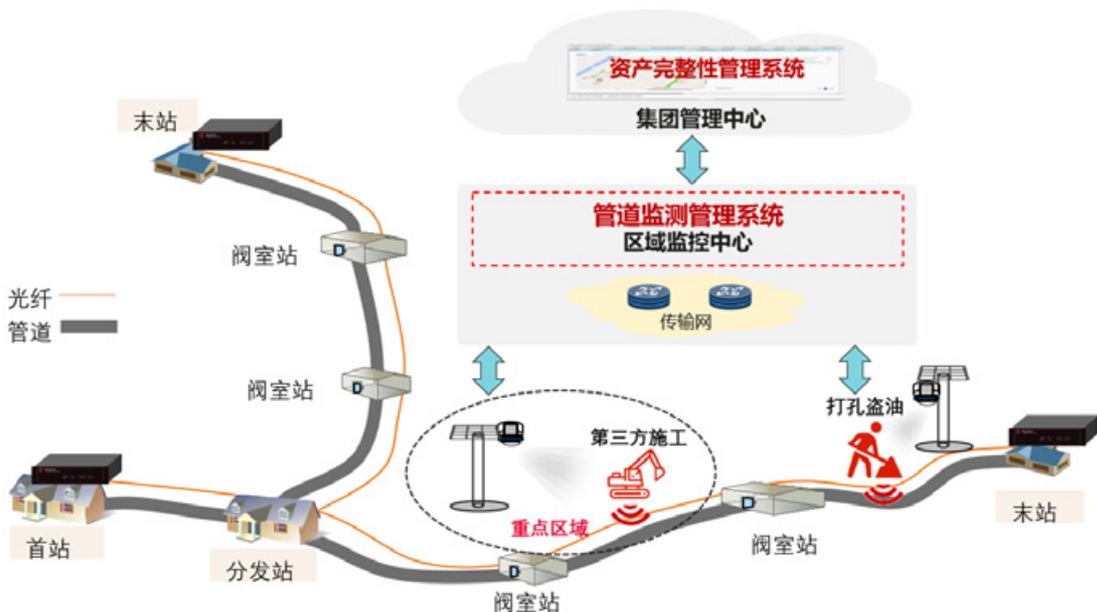


图 3-21 管道光纤监测预警

利用与管道同沟敷设的通信光缆中的冗余光纤，构成分布式光纤传感器，获取管道沿途的振动信号，利用 32 维振动波纹分析 AI 算法对油气管道场景下不同事件的波纹进行训练优化以及在线学习，在识别是否产生振动的基础上，还可以有效地过滤误报以及识别管道沿线的威胁事件（挖掘、钻孔、破坏）类型，实现快速准确定位，保障管道安全可靠高效输送。

3.7.5 化工工艺模拟与优化

化工反应是油气产业中最为重要的核心流程之一，其产物广泛应用于人类的生产生活中。然而，化工反应的复杂性和不可控性导致了其产物的品质和产量受到约束，从而更加注重反应过程的模拟和优化。化工工艺模拟是指利用计算机软件对化工生产过程进行模拟和预测。而化工优化则是指在化工生产过程中，通过对生产流程的数据分析和模拟，找到理想的工艺参数，以实现最佳的生产效益。化工工艺模拟与优化技术可以有效提高化工生产的效率和质量，降低成本和能耗碳耗，防止事故和污染，

推动化工生产向智能化、绿色化、低碳化等方向发展。

目前，化工工艺模拟和优化系统大部分基于数学机理模型，其核心是反应动力学及传递过程的计算，当前仍存在诸多问题，例如系统建模的准确性和可靠性仍然是最大难题之一；其次，化工工艺模拟和优化系统大部分需要手动构建模型，人工成本高、耗时长；同时，模型参数调整和精度评估等问题也亟待解决。

针对上述问题，化工工艺模拟与优化需要继续发展和完善，其中一个思路就是利用人工智能和大数据等先进信息技术进行优化建模，提高模型的准确度和可靠性。化工工艺模拟与优化应用“大数据+AI大模型+分子机理混合模型”，结合流程模拟软件，与实时 DCS 数据、LIMS 数据、在线分析仪数据通信，通过科学计算实现模型自动校正、实时操作优化方案，最终达到企业从原油到化工装置全流程在线优化，如图 3-22 所示。

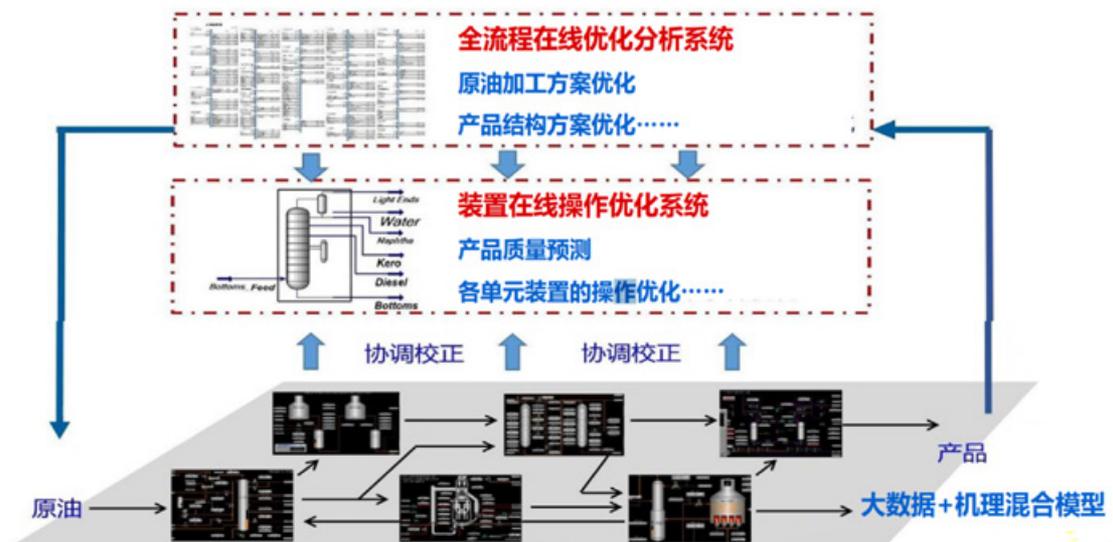


图 3-22 工艺流程模拟与优化示意图

首先，基于严格稳态流程模拟实现物料平衡、热量平衡、化学平衡等计算。其次，通过大数据分析和 AI 大模型在化工工艺模拟和优化系统场景中减少建模工作量，降低模型的复杂度且使之易收敛，缩短计算求解式时间，大幅度提高工艺优化模拟精度和效率。再次，利用双模技术有效提高模型准确度和鲁棒性，实现装置工艺运行数字孪生，智能推荐最佳关键工艺操作参数，助力技术人员基于实时数据进行模型在线校正、在线操作优化。

工艺模拟和优化涵盖常减压、催化裂化、加氢裂化、加氢精制、渣油加氢、重整、乙烯裂解等重点装置，稳步提升生产效率，提高收益。工艺模拟优化与大数据和 AI 的结合，最终可更优实现包括原油加工方案优化、产品结构方案优化、装置生产方案优化、产品质量预测、排放预测等功能，以实现炼化企业降本增效和装置的“安稳长满优”运行，将有效推动炼油化工生产过程迈入新的发展阶段。

3.7.6 智慧生产安防

“十四五”时期是加快构建现代能源体系的关键五年。在全球科技面向数字化和智能化发展

的背景下，为了保障国家的能源安全，降低油气对外依存度，完成降本增效的目标，油气行业数字化转型迫在眉睫。在此背景下，人工智能技术的应用已成为油气数字化转型的关键一步，不仅可以应用于提高生产运营效率、降低运营生产成本，还可以用来保障安全生产、降低环境风险。特别是计算机视觉技术在油气场站/炼厂、管道巡检、设备智能管理、数字油田/数字炼厂安全运营等领域得到广泛应用，工业应用效果突出，是目前行业内应用最广泛、成熟度最高的技术之一。

在油田勘探领域存在的环境复杂恶劣、钻井作业现场不可视和人员安全不可视、缺少实时告警等问题；在油气生产领域存在的缺少智能化、人工巡视效率低、系统联动性差的问题；在油气输送管道领域存在的环境复杂、设备老化、运维量大的问题；在炼化领域存在的炼化园区安全人防为主、智能化低、系统间联动差的问题；在销售板块加油站场景存在的安防智能化水平低与加油支付效率低的问题。针对上述问题，通过引入视频 AI 分析技术，实现现场环境的实时监测、告警及智能化联动分析功能。上述安防场景及 AI 视频技术应用如图 3-23 所示。



图 3-23 智慧生产安防应用场景

通过高空瞭望、电子围栏、目标识别、智能分析、枪球联动、视频拼接、多系统联动等新技术应用，对园区关键卡口、厂区周界、建筑物的关键区域（危化品存储、重要设施等）、输油管道进行全方位管控，实现全态感知和全域安防。借助视频 AI 技术可以端到端识别助力油气生产现场在第一时间、第一现场发现问题解决问题。使用的主要算法模型如下：

1. 危险行为防控：非法入侵、徘徊逗留、攀高、打电话、抽烟、遗留物品检测等；
2. 设备运行监管：漏水漏油检测、烟雾火焰检测、温度感知、可燃气体泄露识别等；
3. 作业安全规范：安全服、安全帽、装卸油安全、吊物下站人、管道上站人等。

目前，通过开展“视频 AI+ 大模型”技术在油气生产过程的应用研究，有效提高了石油化工领域的数字化、智能化程度，取得了初步成效。未来几年，石油化工领域的视频 AI 技术将在标准体系构建、融合 5G 技术的数据采集与传输、行业大模型算法等方面开展进一步的研究工作，加快推动油气行业数字化转型进程。

▶ 3.8 智慧航空

自 2005 年起，中国航空运输总周转量已连续 14 年位居世界第二。2019 年，民航全年完成运输总周转量 1292.7 亿吨公里、旅客运输量 6.6 亿人次、货邮运输量 752.6 万吨，同比分别增长 7.1%、7.9%、1.9%；千万级机场达 39 个，同比增加 2 个。从数据可以看出，中国已是名副其实的民航大国。

如今，航空客运正在由奢侈消费向大众消费转变，人均收入水平的增长和旅游业的发展让大众越来越倾向于选择更加便捷、舒适的交通出行方式。然而，旅客吞吐量的激增也给航空行业带来了挑战，单纯进行物理扩容或依靠人工进行调度指挥、监管运维，难以有效应对当前行业面临的供需矛盾、效率不足等诸多挑战，也难以匹配未来航空在客货运效率、安全、体验等方面高质量发展的需求。在数字化时代，利用人工智能、大数据等科技手段来建设智慧航空，全面推进行业数字化转型升级成为最优选择。

中国早在 2017 年正式提出推进智慧民航建设，开启民航强国的进阶之路。2018 年，中国民航局发布《新时代民航强国建设行动纲要》中提出：“从 2021 年到 2035 年，民航发展要实现从单一航空运输强国向多领域民航强国的跨越。同时指明，建设智慧民航是实现民航强国目标的重要一步。2020 年 12 月，民航局印发《推动新型基础设施建设促进高质量发展实施意见》提出“到 2035 年，全面建成国际一流的现代化民航基础设施体系，实现民航出行一张脸、物流一张单、通关一次检、运行一张网、监管一平台”。2022 年 1 月 21 日，为加强智慧民航建设顶层设计，落实多领域民航强国建设要求，促进民航高质量发展，民航局组织编制了《智慧民航建设路线图》，如图 3-24 所示，标志着中国智慧民航从顶层设计走向了全面实施阶段，智慧民航建设将有效提升民航安全发展水平以及行业运行效益、效率和发展质量。随着 AI 应用加速进入行业生产系统，预计未来 3~5 年，航空行业人工智能市场空间规模预计可达 43 亿。

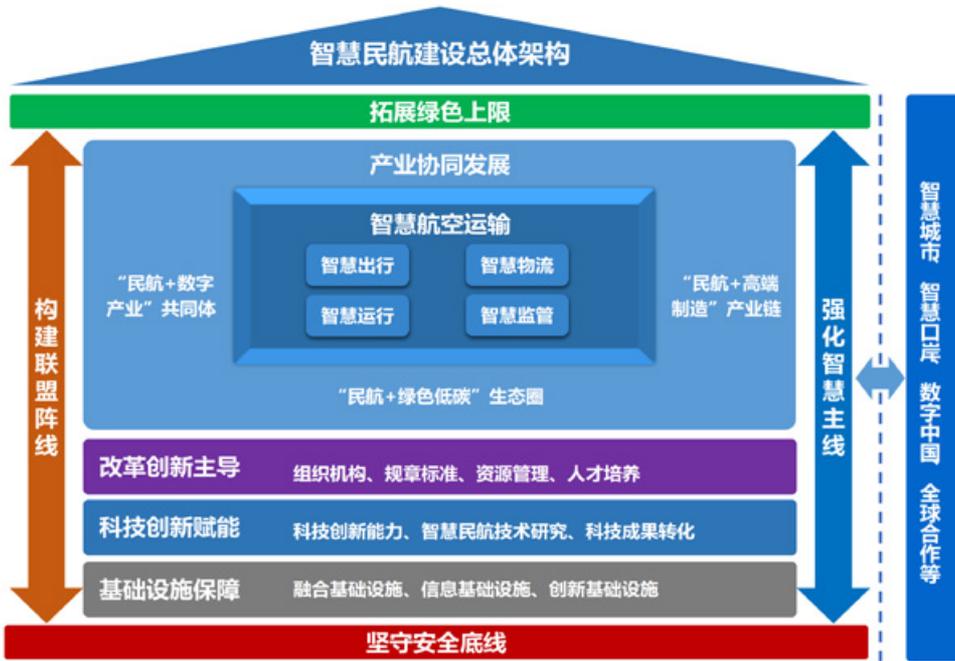


图 3-24 智慧民航建设路线图总体架构

对于机场而言，随着国际航空枢纽建设的快速推进，各项设施资源、空域资源等日趋饱和。现有机场一方面存在业务规模大、运行主体多、运行状况复杂、资源紧缺等问题；另一方面则面临着旅客量、航班量以及航线等业务快速发展带来的运行压力及挑战，迫切需要加速 AI 技术大规模的应用，进一步提高机场安全水平、运行效率、服务品质和建设管理能力。通过对运行、安全、服务、交通、商业、物流、管理等多个业务系统进行调研分析，人工智能在机场存在的典型应用场景如图 3-25 所示。



图 3-25 机场人工智能典型应用场景

在机场运行指挥方面，通过引入 AI 技术，构建机场运行一张图，建设智能机位分配、航班保障节点自动采集、主动运控等系统，实现运控业务的透明可视、全局态势感知、高效协同指挥，不断地促进机场信息化建设、数据质量提升、业务流程优化，打造主动、协同、高效、智能、一体化的大运控，有效促进机场行业的数字化转型。

3.8.1 机位资源分配

停机位是机场运行的资源中心，是航空器、旅客、行李以及地面运输的交集处，机位的高效分配是机场生产组织的核心环节。在机场航班信息管理作业中，机位分配（如图 3-26 所示）不仅直接影响飞机停放的安全，也影响登机口、摆渡车、行李转盘等周边相关资源的合理分配。在航班发生延误、取消、返航等特殊情况时，需要迅速对原有机位分配计划进行相应的调整。在传统的机位分配模式中，运行指挥人员需每天根据当日航班计划和机场设施状况，考虑繁多且复杂的机位分配规则，结合个人专家经验每天需要花费数个小时人工统筹安排飞机停机位分配计划，全天 24 小时监测调整，效率低下。

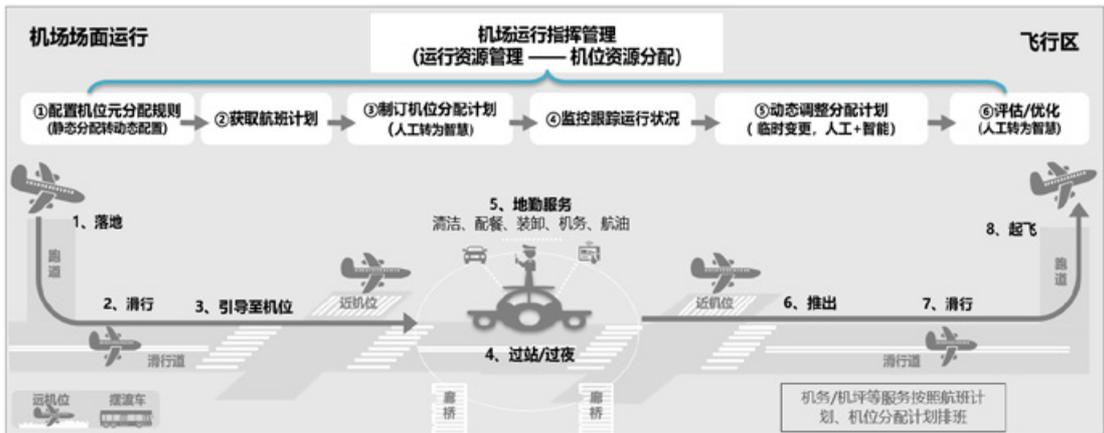


图 3-26 机场机位分配业务场景图

在当前客流量快速增长，机场运营面临安全、运行、服务体验等多重挑战的情况下，由于需考虑的因素众多，有时当场景变得复杂时，光靠人力已经无法得到最佳的机位分配解决方案，亟需对机场关键场面资源进行科学利用和合理配置，建设一套高效的机位资源智能分配系统。通过引入 AI 技术，实现“机器为主、人工为辅”的机位资源自动化、智能化分配，可在向指挥员提供机位智能预分配、实时分配等决策支持的同时，有效提升靠桥率、廊桥周转率等核心指标，并为值机柜台、登机口、行李转盘等运行资源的智能分配打下基础，进一步提升机场运行效率和旅客体验。

▶ 3.9 智能铁路

铁路运输是现代化运输业的主要运输方式之一，在社会物质生产过程中发挥着重要的作用。其特点是成本低、速度快、运量大、运输的准确性和连续性强，几乎不受气候影响。在旧中国，铁路运输能力十分薄弱，全国仅有 2.2 万公里的铁路。新中国成立几十年来，铁路运输有了比较大的发展。截至 2022 年底，中国铁路营业里程已高达 15.49 万公里，其中高速铁路达 4.2 万公里，稳居世界第一。中国已建成世界领先的现代化铁路网，并成为世界上唯一一个高铁成网运行的国家。但面对如此庞大的铁路网，如何提升运输组织效率，解决运输需求与能力供给之间的矛盾；如何实现风险隐患的自动排查，提升事故故障的超前防范能力，适应高速、重载的运输模式；作为重资产行业，如何掌握设备的劣化机理，提升设备的精细化管理水平，准确预判设备的维修时间和内容，大大降低设备的养护维修成本。面对这些问题，经过数十年的探索，智能化成为铁路行业实现高质量发展的必然选择。围绕铁路建造、装备、运营三大核心业务，铁路正在走向全面智能化，如图 3-27 所示。



图 3-27 铁路智能化

2017 年 6 月 29 日，铁路信息化总体规划首次提出了建设“智能铁路”的战略目标。2018 年两会期间，中国国家铁路集团有限公司（简称国铁集团）明确指出，基于 AI、5G 等先进技术，中国高铁将向智能化方向发展。2020 年，国铁集团出台了《新时代交通强国铁路先行规划纲要》，其中明确提出未来 30 年的两步走战略。第一步，到 2035 年率先建成服务安全优质、保障坚强有力、实力国际领先的现代化铁路强国；第二步，到 2050 年，全面建成更高水平的现代化铁路强国，全面服务和保障社会主义现代化强国建设。在时代趋势与政策引领下，智能铁路时代已经到来。随着 AI 技术的兴起，其在铁路智能化进程中发挥着重要的作用。在铁路客运、货运、机务、车辆、工务、供电等多业务场景中，都能利用 AI 实现智能化应用，提升效率、优化成本。典型 AI 场景如图 3-28 所示。



图 3-28 智能铁路 AI 典型场景

铁路 AI 场景多且碎片化，传统的小模型“作坊式”开发，周期长、经验无法继承、无法复用。大模型方式为解决这些问题提供了新的思路。大模型具有开发门槛低、泛化性好、经验可继承等特点，可有效实现人工智能“工厂式”开发，将 AI 规模应用于铁路车辆、工务、供电、客运等业务场景。如图 3-29 所示，基于盘古铁路大模型，在车辆运维场景中，使用 AI 可以“代替”人工看图，实现车辆故障智能识别；在工务线路巡防场景中，视频 AI+ 多传感技术可以实现 24 小时全面监测，提升人工巡防效率。随着 AI 应用加速进入行业生产系统，预计未来 3~5 年，铁路行业人工智能市场空间规模预计可达近 33 亿。

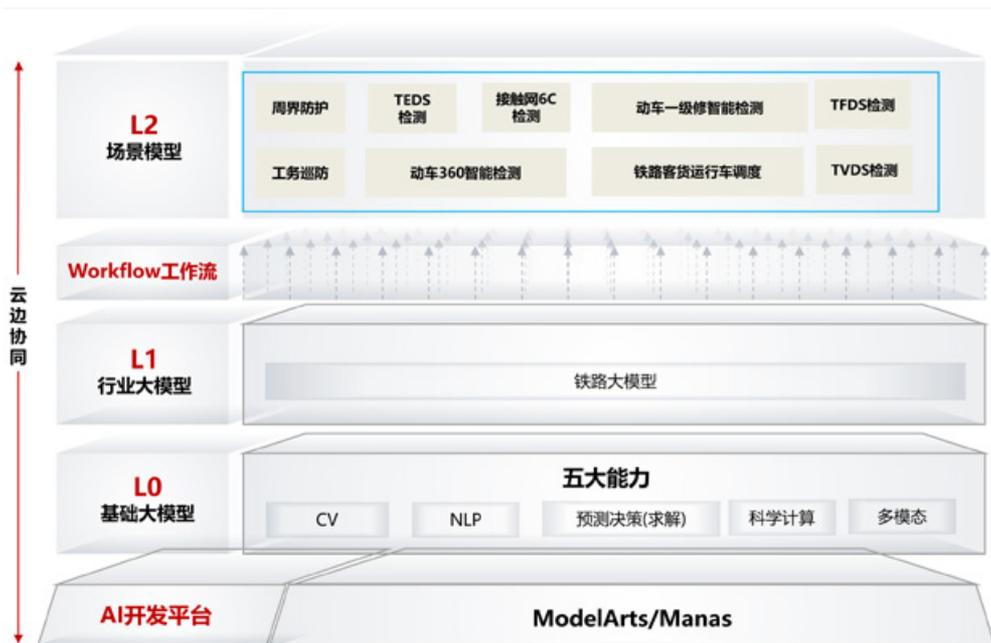


图 3-29 智能铁路 AI 技术架构

3.9.1 车辆运维场景

铁路货运是经济发展的“晴雨表”，展现着经济发展的活力与实力。中国拥有近 100 万辆货车，是平时大家所乘坐动车组数量的数十倍，因此铁路货车行车安全非常重要。多年来，铁路车辆部门立足科技保安全，研制、应用、推广了一系列先进的货车安全监测设备及系统，其中，TFDS（Trouble of moving Freight car Detection System，货车运行故障动态图像检测系统）是非常重要的系统之一。该系统利用高速相机拍摄列车底层、侧部照片，通过图像判断部件故障（如图 3-30 所示），有助于提高列检作业质量和效率，提高车辆安全防范的水平，在保障铁路货车运行安全方面发挥了积极巨大的作用。

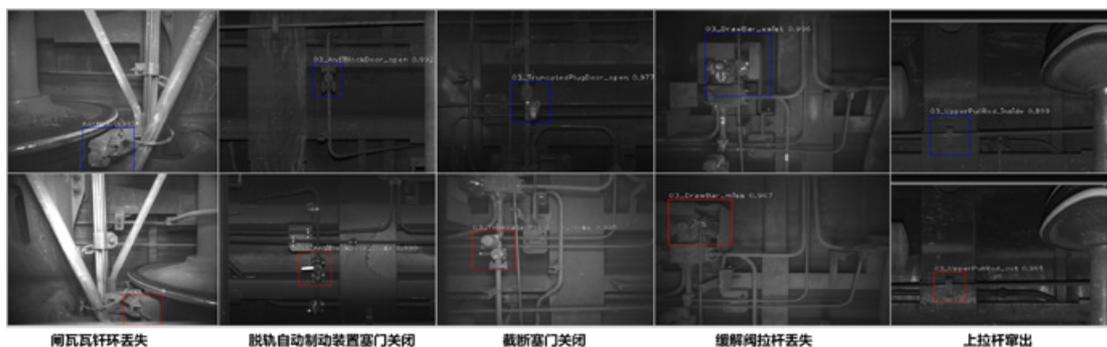


图 3-30 货车故障图像示例

目前，全国 TFDS 系统每天拍摄上亿张图像，主要依赖人工作业，工作强度大、效率低、人力投入大。AI 技术可以实现 TFDS 货车故障图像从人工识别转向智能识别，减少人工作业量。原本一列货车需要人工查看 4000 余张图片，经过 AI 识别后，可以过滤掉 95% 以上的无故障图片，人工只需对少量故障预报图片进行复核，作业效率大幅提升。

3.9.2 线路巡防场景

高铁已经成为人们跨城出行的首选交通方式之一，与我们的日常生活密不可分。高铁运行速度快、载客量大，全国“一张图”运行，一个点出现问题都有可能引起连锁反应，甚至会导致大面积晚点。国铁集团及其下属各铁路局投入了大量人力、物力、财力，全面守护高铁的运行环境不受外部因素干扰，线路巡防就是其中重点工作之一，全路投入的巡防人员数以万计。在每年的汛期，巡防人员更是 24 小时驻守在防洪点，时刻守护着铁路的基础设施和行车环境，承担着沉重的工作任务和巨大的精神压力。虽然可以利用沿线铁路综合视频监测系统，对风险点定期远程巡查，但每个铁路局都面临数万个摄像头，靠“肉眼”全面实时查看显然不现实。

随着 AI 智能分析技术的发展以及 AI 硬件算力的持续提升，为解决这一行业难题提供了可行的方向。基于强大的 AI 硬件算力，结合深度学习的视频 AI 算法，对铁路沿线海量视频数据进行全面实时的智能分析，一旦发现潜在问题，及时通知值班人员进行确认。同时，视频 AI 可叠加光纤传感、毫

米波雷达、激光雷达等多种感知技术，实现对入侵目标的更高精度识别。这样可大幅降低人工现场和远程巡查的工作量，大大提升线路巡防的工作效率。

▶ 3.10 智慧城轨

城镇化建设高速发展，带来了城镇居民日常通勤需求的激增，促进了城市轨道交通规模持续增长。截止到2022年底，中国已有55个城市，累计投运城轨交通线路10287公里。不过与世界发达城市相比，城轨交通在提升城市公共交通供给、促进都市圈发展等方面的作用可进一步提升。

2019年，国务院印发《交通强国建设纲要》，指出要以新兴信息技术与城轨交通深度融合为主线，推进城轨信息化，发展智能系统，建设智慧城轨，实现城轨交通由高速发展向高质量发展的跨越，助推交通强国的崛起。2020年3月，中国城市轨道交通协会发布《中国城市轨道交通智慧城轨发展纲要》，提出应用云计算、大数据、物联网、AI、5G、卫星通信、区块链等新兴信息技术，全面感知、深度互联和智能融合乘客、设施、设备、环境等实体信息，经自主进化，创新服务、运营、建设管理

模式，构建安全、便捷、高效、绿色、经济的新一代中国式智慧型城市轨道交通。2022年8月中国城市轨道交通协会发布《中国城市轨道交通绿色城轨发展行动方案》，提出推动绿色装备创新和新一代绿色智能技术装备的研发与应用，聚焦节能降耗潜力大、能效利用率高的新一代绿色智能技术装备。

随着智慧城轨建设的不断深化，城轨行业的数字化转型已形成共识并正在加速，行业正从业务上云向业务智能的阶段演进，如图3-31所示。在业务智能阶段，行业的客户开始不断尝试在城轨云上做业务的创新，设计和构建了支撑整个企业业务变革的数字底座，推进IT和OT的融合，并逐步在业务场景中融合AI，打通流程的断点、重构城轨业务，逐步向智能化、数字化大步迈进，如车辆运维的自动化诊断分析、企业经营的辅助决策等。



图 3-31 智慧城轨建设的演进过程

智慧城轨的建设要坚持智能化的实施战略，智能化方向重点推进云计算、大数据、AI 等新兴信息技术和城轨交通业务深度融合，推动城轨交通数字技术应用。随着城轨线路和车站数量增加，线网客流持续增长，城轨线网结构愈加复杂，建设、运营和运维压力随之加大，城轨面临可持续发展的挑战。结合中城协提出的智慧城轨体系，如图 3-32 所示，结合八大体系的内容，行业迫切需要围绕“安全”、“效率”和“体验”等目标，大力推广大数据、AI 等技术在城轨智慧乘客服务、智慧运输组织、智能能源系统、智能技术装备、智能基础设施和智慧运维安全等领域的深化应用，加速行业智能化。随着 AI 应用加速进入行业生产系统，预计未来 3~5 年，城轨行业人工智能市场空间预计可达 39 亿。



图 3-32 智慧城轨八大体系

3.10.1 智慧建造场景

城轨新建规模大、周期 3~5 年，多条线路同时施工，存在安全监管不到位、容易引发安全事故等问题。如何在建设中有效规避风险、保证施工安全，成为地铁建设必须解决的问题。目前地铁建设信息化管理较落后，存在如下短板：工地安全隐患识别不全、自动化监测排查能力有限、过程管理效率低、建设系统数据未统一、管理协同难等，企业层面无法及时获取线网建设的安全、质量和进度等信息。如某地铁集团需要依赖 15~20 人的巡检队伍完成 300 多个监测点的巡查，时间、效率难以保障。

基于此，在城轨建设领域，需要通过视频+AI+ 大数据等新 ICT 技术，融合城轨施工建

设数据，实现城轨建设一张图管理、安全隐患自动识别、轨行区安全调度、盾构安全大数据等应用，通过资源、技术、数据共享实现施工单位管理、施工安全管理、应急管理等功能，达成工地 AI 算法精确度超过 99%、巡检工作量减少 35%、安全事故降低 30% 等。

3.10.2 智慧客运场景

随着地铁线网规模的不断扩大，地铁对城市乘客出行产生虹吸效应，使得客流呈现快速增长，城轨运营管理在客流预测、行车调度、客流调度三大业务上面临着巨大挑战。当前的客流预测更多依靠历史客流数据统计分析来估算，客

流预测的准确率低于 70%，对突发事件引起的客流变化缺乏推演能力，无法有效支撑后续的运输策划和运营调度。在行车调度、客流调度业务上，目前只有大客流、行车延时 / 中断事件监测，没有针对这些突发事件产生后客流如何演变的预测，以及针对客流演变应该采取怎样的行车调度、客运组织的建议，仅依靠调度人员的个人判断，无法匹配线网级的行车调度和管控。

基于此，一方面需要基于城轨大数据建设，使用 AI 算法进行精准的客流分析和预测，实现短时客流预测精度超过 90%、实时客流预测精度超过 95%；另一方面，需要分析实际业务的各种约束条件和优化目标，建立多目标、多约束条件的网络运行图整数规划模型，采用求解器和启发式算法实现精确求解，实现分钟级的运行图编制和运能运量的精准匹配，有效支撑大客流时的线网、线路和车站一体联控。

3.10.3 智慧运维场景

城轨网络化运营下设施设备运维管理压力日渐严峻，传统的运维模式随着现网规模化发展已经不能满足运维业务的发展需要，面临“成本”和“效率”的双重挑战。当前，故障类型和故障程度的诊断，依赖个人经验判断，设备健康

状态监测难，老专家仅能识别约 70% 的故障、普通员工仅能识别约 30% 的故障。其次，运维检修作业计划依赖人工制定、人工编排，效率低下；同时，车辆大修 / 更换等按周期开展，易浪费维修资源和成本。

智慧运维可通过图像识别及目标检测等 AI 算法，动态采集车辆轮对、走行部、受电弓等关键部件图像，并提取特征对比，及时识别缺陷和异常报警，替代 80% 以上的人工日检内容，实现超过 95% 的识别精度并提高 35% 的日检效率。同时，基于设备设施的状态监测数据构建 PHM（故障预测和健康管理）模型，开展劣化预判和预测，探索从“计划修”向“状态修”，以及“预测修”的演进，以降低运维成本，提高运维效率，保障运维质量。

结合智慧建造、智慧客运、智慧运维等场景的业务场景需求和 AI 场景技术构思，城轨行业要构建云上训练、边端推理的整体架构，如图 3-33 所示。在城轨云上进行模型训练，训练好的模型通过平台下发到云边缘，云边缘进行模型推理，实现云端和边端的增量学习、联邦学习、协同推理等能力，以提升模型性能和精度，并降低模型建设成本。比如冷机节能算法，通过云上的实时训练，实现站段节能算法快速调整和小时级增量更新能力，保持模型精度大于 90%。



图 3-33 城轨行业 AI 能力应用思路

▶ 3.11 智慧港口

港口是一个国家或地区的门户，通常位于江、河、湖、海沿岸，有水、陆接运的作用。现代港口是水陆交通的集结点，具有水陆联运设备和条件，是供船舶安全进出和停泊的运输枢纽，同时也是工农业产品和外贸进出口物资的集散地，更是国际物流全程运输与国际贸易的枢纽和服务基地。

天津港面对港口行业数字化转型，提出更高的业务目标，“要志在万里，努力打造世界一流的智慧港口”。为响应党中央号召，港口行业普遍开始基于人工智能、大数据、5G、云、北斗等新技术向智慧港口进行转型，同时提出

以“海港为龙头、陆港为基础、空港为特色、信息港为纽带”的四港融合战略。其中智慧港口以信息物理系统 CPS (Cyber-Physical Systems) 为结构框架，通过高新技术的创新应用，使物流供给方和需求方共同融入集疏运一体化系统；极大提升港口及其相关物流园区对信息的综合处理能力和对相关资源的优化配置能力；智能监管、智能服务、自主装卸成为其主要呈现形式，并为现代物流业提供高安全、高效率和高品质服务。全球 TOP10 集装箱港口中，中国占有 7 个，自动化码头建设运营水平全球领先。自动化码头通过将人工智能、大数据、5G、云、北斗等新技术和码头自动化作业技术相结合，真正实现了码头生产作业的全自动化、少人、无人。集装箱码头核心生产作业环节，如图 3-34 所示。



图 3-34 集装箱码头核心生产作业环节示意图

港口的智能化演进可以分为三个阶段：早期的信息化阶段，中期的数字化 + 场景化 AI 阶段，以及未来的智能化阶段。港口智能化三个阶段的演进，如下图 3-35 所示。

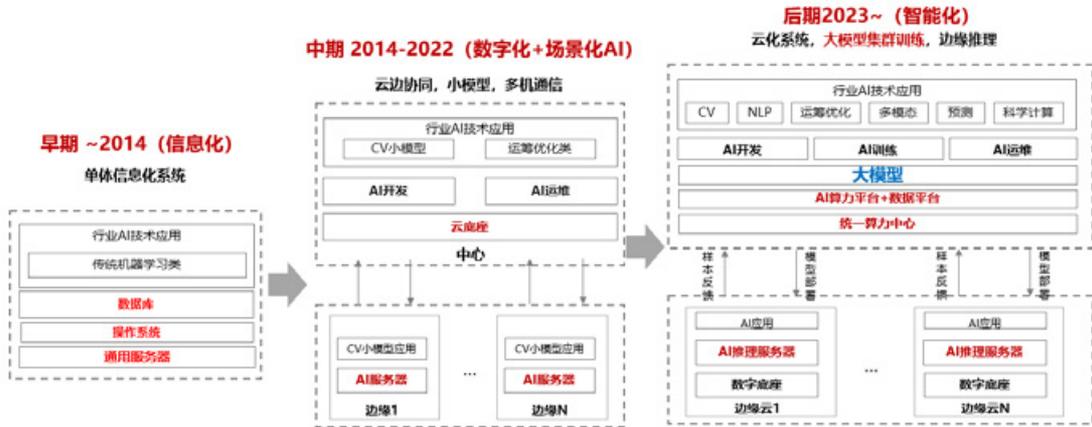


图 3-35 港口智能化转型的演进路径

2014 年前主要通过传统机器学习方法支撑了一些港机设备故障预防预测，2014 年开始随着深度学习的逐渐普及，基于卷积神经网络 CNN（Convolutional Neural Network）的视觉小模型逐步在码头安防、监测等领域得到了大量的应用，同一时期传统的运筹优化算法结合求解器也开始在计划调度及路径规划等场景开始持续发力，在这个阶段内，以磁钉 + 自动导引车 AGV（Automated Guided Vehicle）为代表特征的智慧港口 1.0 逐步演进为以自动驾驶 + 高精度地图 + 5G+ 智能导引车 IGV（Intelligent Guided Vehicle）为代表特征的智慧港口 2.0；2023 年随着大模型的出现，神经网络的智能水平得到显著提升，使得基于一个统一大模型底座全面支撑港口生产、管理、服务、安全各领域应用成为可能，智慧

港口也开始从 2.0 向 3.0 开始演进，装卸生产智能化、物流贸易协同化、运营管理数字化、生态服务一体化成为这一阶段的特征及趋势。据初步测算，未来 5 年内大模型等 AI 新技术为港口领域带来的市场空间约为 25 亿，主要包括省港集团 AI 训练中心、大模型训推一体机以及相应的专业服务等。

自动化集装箱码头作业可分为装船、卸船、集港、提箱四大流程。以卸船流程为例，码头会提前制定生产作业计划，在船舶到港后，使用岸桥进行集装箱卸载，经理货后由水平运输系统运送至堆场，再由场桥移进堆垛，最后由外集卡运输至目的地，作业流程中由安防系统进行实时监测。

3.11.1 智能计划

码头通过制定生产作业计划，为进出港船舶及集装箱装卸、运输和存放提前分配泊位、岸桥、场桥、集卡、堆场、舱位等资源。传统方式采用人工制定计划，存在人工耗时长、效率提升难等问题。通过使用运筹优化技术进行场景化建模并调用求解器，可在考虑多种约束的情况下，在给定优化目标上快速计算出最优解或近优解，将计划生成耗时从小时级降低到分钟级甚至秒级，同时在泊位利用率、船舶在港时间、船时作业效率、倒箱量等关键指标上带来一定优化。集装箱码头智能计划通常包含智能泊位计划、岸桥作业计划、智能配载计划以及智能堆场计划，如图 3-36 所示。

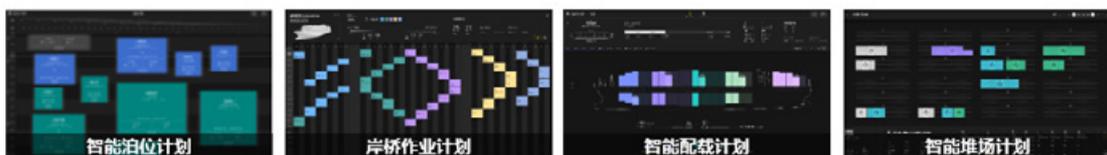


图 3-36 集装箱智能计划结果

3.11.2 智能操控

港机操控发展一般可分为现场操作、码头自动化操作、超远程集中控制、人工智能控制四个阶段，如图 3-37 所示。



图 3-37 港机操控发展趋势

传统码头通过人工现场操控岸桥和场桥等吊机设备，存在作业艰苦、招工难、人力需求高、资源不能共享的问题。每台吊车 24 小时作业需配备多名司机轮换，而吊车司机是特殊工种，要求高、培训时间长，上岗后需要持续在数十米高空低头弯腰操作，身体损伤大，一般工作到 40 岁就会转岗。自动化码头一般通过视频回传 +F5G/5G 远控的方式支持司机远程作业，1 名吊车司机可远程操控 1 台岸桥或多台场桥，改善工作环境的同时也提高了作业效率。一些自动化码头已经基于视觉 AI 技术初步实现了岸桥作业的部分自动化，以及场桥作业的全程自动化，但在控制方面仍有不足，岸

桥作业时集装箱运行的线路不够优化，场桥作业时会出现频繁二次着箱、人工介入的现象。目前国内已有个别企业通过使用深度强化学习技术，通过作业视频学习桥吊司机的操作经验，而在实际作业中根据预设目标、集装箱历史运行路径、当前运行速度、方向、晃动等情况进行实时控制调整，选取最优路径，避免二次着箱，减少人工介入，有效提升了吊车作业效率。

3.11.3 智能理货

集装箱在进行船舶装卸作业时，需要进行箱号核对、箱损和铅封检查等操作，如图 3-38 所示。



图 3-38 传统理货与智能理货

传统理货需要理货员在装卸现场通过理货单或者理货终端进行，因码头机械、集卡、货物环境复杂，存在一定安全隐患。部分码头已将理货作业后移到码头理货中心，前端通过光学字符识别 OCR（Optical Character Recognition）识别箱号，后端通过查看照片，检查铅封、残损状况，此时理货环境得到改善，但人工工作量仍较大。部分码头已采用智能理货系统，通过 CNN 等视觉 AI 技术，在岸桥装卸过程中自动识别集装箱号、ISO 号、危险品标识、铅封等信息，支撑理货人员从露天站位盯箱到室内轻点鼠标，这样就在有效改善理货环境的同时，也提升了理货效率，并实现了理货安全。

3.11.4 智能水平运输

水平运输系统是指针对集装箱码头的岸边与堆场间的集装箱集运输环节，构建集内集卡、上层 TOS 调度系统与通信系统于一体的管理与控制系统。传统方式依赖人工司机 7*24 小时作业，存在司机招募难、效率提升难、安全风险高的问题。智能水平运输系统使用监督学习、深度学习、强化学习等 AI 算法，基于 5G、北斗、高精地图，通过云端智能化调度，实现了动态排泊、IGV 实时路径规划、基于运动学模型的车辆控制、超 300+ 车辆的大车队统一管理以及车辆在码头内安全自

动驾驶。提高岸线资源利用率的同时，有效保障了码头的作业效率和生产安全。水平运输系统在实际作业中无人集卡运行的情况，如图 3-39 所示。



图 3-39 无人集卡作业

3.11.5 智能闸口

如图 3-40 所示，集装箱在进出闸口时需要检测集卡货箱信息，传统方式下采用人工现场检查，存在效率低、易拥堵的问题。通过使用卷积神经网络等视觉 AI 技术，可自动识别集装箱号、ISO 号、装卸状态、船上贝位、装车位置、车顶号码、车牌号码、单双箱类型、危险品标识、铅封等信息，进而结合业务规则进行实时智能分析决策，实现闸口自动检测及放行，有效减少了集卡等待时间，降低了拥堵发生概率。



图 3-40 闸口检测

3.11.6 智能安防

港口安防包含了泊位、堆场、缓冲区、辅建区、闸口、码头楼宇内部以及港区道路等区域安全管理。传统方式依赖人工监测，工作量大，效率和可靠性均难以保障。通过使用卷积神经网络等 AI 技术，可实时发现车辆逆行、周界入侵、人员未戴安全帽等不安全行为和不安全状态，大幅提升监测效率，有效保障码头生产安全。智能安防典型场景如图 3-41 所示。

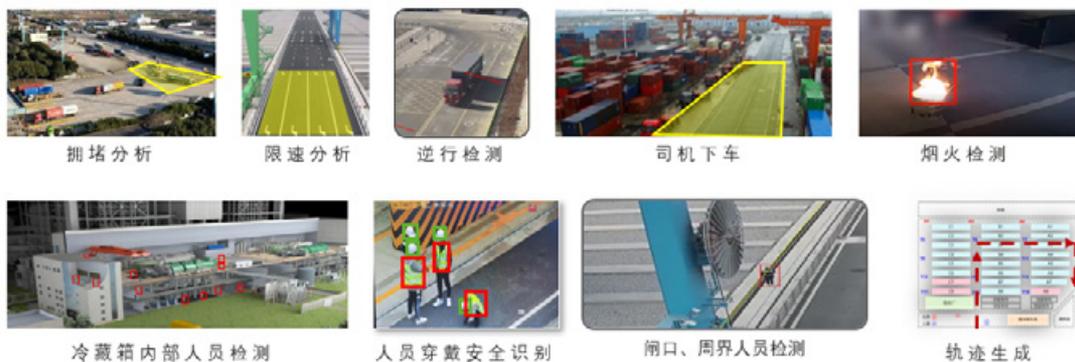


图 3-41 智能安防典型场景

▶ 3.12 城市智慧交通

2019年9月，中共中央、国务院印发《交通强国建设纲要》，作为新时代交通工作的顶层战略统筹文件，《交通强国建设纲要》提出从交通大国到交通强国，形成“全国123出行交通圈”与“全球123快货物流圈”的总体目标，要求推动大数据、互联网、人工智能、区块链、超级计算等新技术与交通行业深度融合，构建泛在先进的交通信息基础设施，建设富有活力、智慧引领的科技创新体系，到2035年，基本建成交通强国。

2021年10月，为贯彻党的十九届五中全会精神，落实《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》要求，按照《交通强国建设纲要》《国家

综合立体交通网规划纲要》相关战略部署，交通运输部印发《数字交通“十四五”发展规划》，进一步明确发展目标：到2025年，“交通设施数字感知，信息网络广泛覆盖，运输服务便捷智能，行业治理在线协同，技术创新活跃，网络安全保障有力”的数字交通体系深入推进，“一脑、五网、两体系”的发展格局基本建成，交通新基建取得重要进展，行业数字化、网络化、智能化水平显著提升，有力支撑交通运输行业高质量发展和交通强国建设。

城市智慧交通经历了多年的发展，可以总结为交通控制自动化、大数据辅助管控、大模型智慧运营三个阶段，如图3-42所示：



图 3-42 城市智慧交通发展历程

第一个阶段是交通控制自动化阶段。在该阶段重点建设信号控制、过车抓拍、违法抓拍、流量检测、视频监测等外场设备，以及配套的设备管理及信息采集系统，实现了交通信号自动化控制、交通违法自动抓拍、交通过车自动采集、交通流量自动采集、视频监测自动回传，初步构建交通自动化控制与监测体系。但系统之间烟囱式建设，数据利用相对单一，缺少统一的管理平台。

第二个阶段进入了大数据辅助管控阶段。随着云计算、大数据、人工智能等新技术的出现，智慧交通信息化迎来了新的发展方向，在进一步完善监测预警手段的基础上，围绕着大数据资源中心、大数据辅助管控、视频智能分析开展建设，深入挖掘数据价值，进一步提升辅助决策、管控能力，同时提供基本的公众信息服务，基本实现交通态势可视、可测、可服务。但道路交通管理、综合交通管理、公交管理行业系统之间仍然相对独立，数据共享利用率低。

第三个阶段进入了大模型智慧运营阶段。随着物联网、5G、高精定位、AI大模型技术的进

一步成熟，带动了城市智慧交通的进一步发展，通过打破交管、交通、应急等行业数据壁垒，汇聚全域全量精准数据，基于大模型保障城市交通安全、高效运行。将高质量 prompt 标注、AI 训练芯片、人工反馈强化学习训练等技术与能力，与交通行业管理的需求深度融合，构建交通千亿级参数规模大模型，以多模态多能力向未来的新一代智慧交通体系服务。

城市交通系统的管理职责主要由城市交通运输部门与道路交通管理部门承担，由于参与主体包含社会大众、从业人员、相关交通企业，管理对象包含各类交通运载工具、固定设施系统、道路管理系统，按照不同的类目，城市智能交通系统以相关政策及标准牵引，以职能部门的行政许可、行政审批、行政监管、服务考核等核心职责为导向，面向综合规划、工程建设、设施养护、交通管理、公共交通、道路客运、道路货运、交通执法、安全应急、静态交通等不同的业务场景提供基于业务流程的信息化服务。

城市智慧交通系统包括道路交通、城市公交、

相关配套设施等，是一个开放的复杂巨型系统，需要从交通需求、交通供给、交通组织、交通管理、交通行为等各方面综合制定顶层方案。



图 3-43 城市智慧交通业务场景全景图

在智慧交通快速发展的情况下，交通行业管理服务不但面临着交通事故预防、交通拥堵治理等复杂难题，同时也面临着公共交通、交通枢纽供需匹配不平衡、不精准等问题。建设的大量设备设施健康状态主要采取人工巡查的方式和定期检测的方式开展工作，导致基础设施类灾害预警评估所依据的数据来源不全面、不准确且滞后，基础设施类灾害的预警预报缺乏及时性、有效性和准确性。这些问题可以基于视频监控数据、全域全量标准化数据依靠人工智能的 CV 能力、预测能力、多模态能力、NLP 能力实现交通事故预防、交通拥堵治理、智慧公共交通、智慧交通枢纽、设施状态监测等 AI 场景，支撑多种运输方式的调度协调，提供交通行政管理和应急处置的信息保障，提升交通事件发现、预警、预防、处置的能力，降低交通运输隐患和风险，提升交通运输效率。



图 3-44 城市智慧交通大模型应用场景

3.12.1 交通事故预防

“十四五”时期处于“两个一百年”奋斗目标的历史交汇点，迈入高质量发展新阶段，预计到2025年，中国机动车保有量、驾驶人数量、公路通车里程将超过4.6亿辆、5.5亿人和550万公里。人、车、路等道路交通要素将持续快速增长，导致中国道路交通安全整体形势依然不容乐观，道路交通安全工作基础仍然比较薄弱，存在不少短板弱项，地区和领域发展不平衡不充分问题仍然突出，农村交通安全问题凸显，道路交通事故时有发生。根据国务院安委会办公室关于印发《“十四五”全国道路交通安全规划》的要求，道路交通事故万车死亡率“十四五”期间相较于“十三五”末年年均下降3%左右，较大道路交通事故起数“十四五”期间相较于“十三五”末年年均下降4%左右，“十四五”期间重特大道路交通事故起数年均控制在4起左右，国省干线交通安全设施技术状况优良率2025年达到85%，电动自行车和摩托车头盔佩戴率2025年摩托车骑乘人员头盔佩戴率达90%，电动自行车骑行人员佩戴率达到80%，汽车安全带佩戴率2025年前排达到95%，后排达到70%。可以依托AI技术保障交通安全、排查交通隐患，降低交通事故的发生概率。

交通事故预防的AI场景包括机动车特征识别（含危化品车辆检测）、交通事件检测（包括交通事故、交通拥堵、逆行/倒车、异常停车、路面抛洒物、路面积水、隧道烟火检测）、交通违法检测（包括机动车闯红灯、不按导向行驶、疲劳驾驶、超高、超载、闯禁行、非机动车闯红灯、非机动车不带头盔）等场景。

3.12.2 交通拥堵治理

交通拥堵是指由于机动车、非机动车、行人交

织在一起，造成通行效率低下的情况，通常出现在节假日或者上下班的高峰时期，在经济发达、人口密集、车辆众多、的大中城市，造成交通拥堵的主要原因是交通系统的承载量小于区域交通负荷量、交通供需不匹配、交通组织不良等。可以借助AI技术手段对交通拥堵路段、区域进行缓堵治理、引导分流，提升交通通行效率。

常见的交通拥堵治理AI场景包括交通流量检测（包括过车流量、客货运流量、出行规律）、城市交通规划、交通组织优化（包括信号控制优化、信控绿波带、HOV车道、潮汐车道）、交通流量分析与预测、OD分析、预约出行等场景。

3.12.3 智慧公共交通

根据交通运输部和国家发改委印发的《绿色出行创建行动考核评价标准》，“十四五”期间，超/特大城市的公交机动化出行分担率要求达到50%以上，大城市、中小城市分别为40%、30%；清洁能源公交车比例不低于50%，有必要持续提升公交出行品质，强化公交行业的交通服务竞争力。由于公共交通属于传统型交通行业，在行车计划编制、运营调度、充电管理流程中往往依托于人工经验，存在方案编制周期久，人工填写时间长、充电效能待提升等短板。

智慧公共交通以城市地面公交的数字化、智能化为依托，全面提升行业效能、强化公交服务。其中，公交智能排班调度基于公交车辆运力与客运需求的精准匹配，通过AI算法快速输出排班计划，为车辆提供调度灵活的发车时刻；公交智慧充电根据波峰、波谷不同的电价成本，推导智能充电计算模型，从而合理安排公交车

辆班次和充电的计划，降低新能源公交企业整体用电成本；公交线网优化融合城市出行大数据，通过客流 OD 与线网模型快速生成轨道接驳、高快干线，为行业决策提供数字化依据。

3.12.4 智慧交通枢纽

交通枢纽是城市内外交通的骨干节点，承担交通出行换乘、接驳能力。由于枢纽内部的到发客流量大，市内交通接驳服务需求高，在高峰客流抵达、发送的时间节点出现时，极容易出现枢纽周边交通拥堵、枢纽内部旅客滞留、乘客等待出租、公交等时间过长等各类问题。

智慧枢纽以“多位一体”综合枢纽建设理念，通过海量出行数据的分析，建立一套枢纽覆盖服务区域、乃至更大的时空范围内具备感知、互联、分析、预测、控制等能力。包括基于枢纽各类交通方式的 OD 时空矩阵，通过枢纽出

行算法在高峰出行时期对各类交通方式的接驳运力进行预测，并根据旅客的出行需求，通过 AIGC 大模型推荐合理的出行方案等。从而实现高铁、长短途公交巴士、地铁、社会车辆等各类交通无缝换乘，以充分发挥交通基础设施效能、提升交通系统运行效率和管理水平，为通畅的公众出行和可持续的经济发展服务。

3.12.5 交通基础设施健康监测

交通基础设施监测是预测设施健康状态的重要依据，是保障交通安全、畅通的重要保障，是减灾防灾的重要内容，其核心是通过直接观察和仪器测量记录基础设施灾害发生前各种前兆现象的变化过程和基础设施灾害发生后的活动过程。交通基础设施监测包括对桥梁健康状态、隧道健康结构、山体滑坡、护栏位移、标识标线不清晰、设施进行自动监测等场景，这些场景同样可以使用 AI 技术来实现。





第四章 智能化使能民生

▶ 4.1 城市治理

随着城市人口数量和人口密度的持续增长，城市的资源承载能力受到极大考验。城市空间无序开发、城市管理效率不高、城市服务供给不足的矛盾日益凸显。城市发展在交通、住房、环境、资源等领域面临挑战，城市治理效率亟待提升。运用 AI 技术赋能城市治理，可以有效提升城市管理精细化水平，是实现高质量发展、创造高品质生活、打造高效能治理的重要方式。

城市治理业务涉及城市运行、城市管理、社会治理、应急管理等重点领域，按照业务横向协同、纵向联动的建设思路，通过建立省、市、县（区）三级基础平台，形成“统一领导、统筹管理、专业运营”的管理架构，面向省、市、县（区）、镇（街）、村（社区）五级用户提供创新管理模式，实现城市治理的纵横向联勤联动。通过实现事件的智能发现、智能分拨、智能处置，全面提升城市运行精细化管理水平。

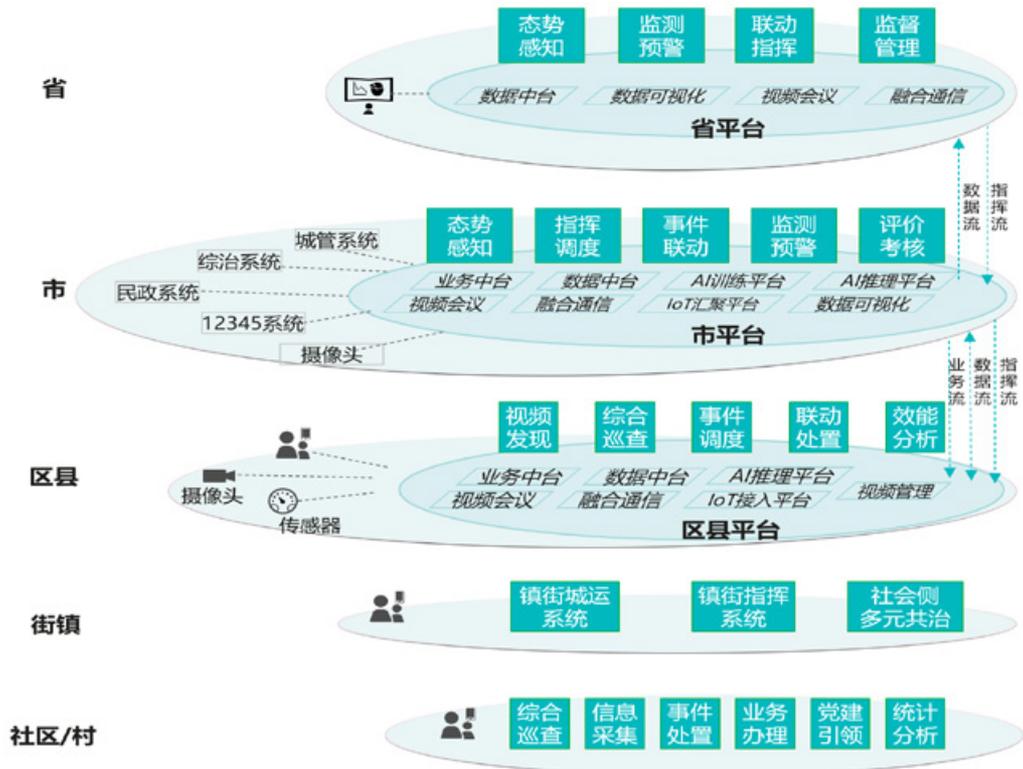


图 4-1 城市治理业务体系

城市治理面向区域内的问题、事件，对于采集、分拨、处置、结案各环节通常采用人工经验判断的方式进行处理，往往被动响应、处置力量到达现场时间较长，事件处置效率低、群众满意度差。

通过使用人工智能技术，促进业务与 AI 的不断融合，实现城市事件智能发现、城运工单自动分拨、城市问题智能分类，充分提升城市事件的处置效率、提升城市治理智能化水平。

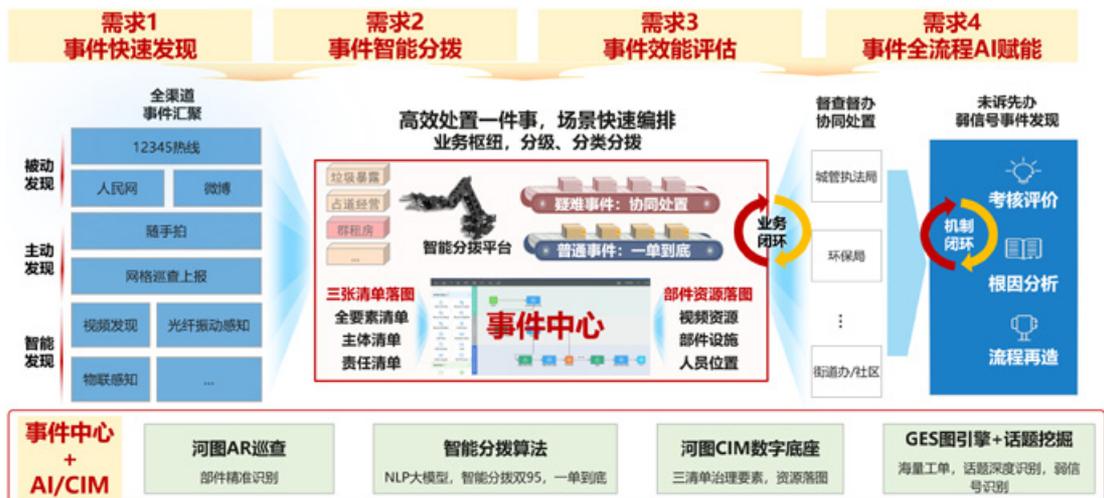


图 4-2 城市治理事件处置的需求

利用 AI 技术从事前、事中、事后三个阶段分别赋能和改善传统城市治理模式，强化事件及时发现、问题及时处置、事后科学评价的闭环管理。

一是在事件发现阶段，以往主要依赖于人工巡查、市民举报、12345 热线等途径，事件发现的时效性较低，影响了群众的生活幸福感和办事体验感。利用视频 AI 技术自动发现事件，不仅能够帮助基层工作人员实现市政、环卫、违停等多个场景事件的识别，还能够自动上报至事件中心完成派遣，从而有效提高事件的处置效率，减轻基层工作人员的工作压力。

二是在事件处置阶段，以往需要通过人工查找

各部门的“三定”方案及相关法律、法规明确事件归属责任单位，再派单至对应部门。此过程耗时较长且错误率较高，影响了事件处置效率。通过运用 AI 技术赋能工单自动分拨，能够快速将事件与权责部门进行匹配，同时为复杂事件提供相似事件的解决建议，大幅缩短人工派单时间，提升事件处置精确度和处置效率。

三是在事件处置完成后，传统管理模式需要依靠人工给出事件处置的效能评估结果，用时较长、工作量大。通过运用 AI 技术，对处置效果进行跟踪、挖掘热点问题、挖掘责任主体和问题根因，持续提升政府服务效率，增强人民群众的获得感和满意度。



图 4-3 AI 赋能城市治理

高效处置城市事件，是推进城市治理新模式改革的重要突破口。未来，城市治理将持续运用 AI 技术赋能城市治理事件全周期管理。作为提升城市风险防控能力和精细化管理水平的重要途径，构建城市发展新格局的重要抓手，AI、大模型等新技术赋能的城市治理场景的市场空间将超过百亿，将对促进城市高质量发展、推进城市治理现代化发挥巨大作用。

4.1.1 事件智能发现

城市运行管理中心汇聚了城市各委办局和社会上百万路视频数据，但事件智能发现率较低。事件的主动发现基本依靠人工方式或市民投诉，中等规模的城市每个区约有 500-1000 名网格员，网格员工作强度高但往往不能快速发现城市治理问题；各委办局、企事业单位的视觉智

能化系统分散建设，基础设施重复投资且建设水平参差不齐，算法标准不统一。为了解决此类问题，视频智能发现利用可演进的 AI 能力，提供多场景的城市治理类算法，使能视频识别和自动发现，并通过事件发现模型与事件中心联动，实现自动识别违规行为、自动报警推送，同时通过网格员中心巡查的方式，减少人工巡查时间，提高城市事件多环节处理效率。

目前视频的智能发现 AI 算法可以为多场景提供服务，如出店经营、暴露垃圾、绿地脏乱、机动车违停、车辆或大型物体占用消防通道、道路积水、人群密度监测、特殊车辆识别、渣土车识别等场景，有效提升事件发现时效性。

4.1.2 工单智能分拨

在事件发现之后，如何高效、准确地将事件分拨给正确的处置部门是当前城市治理的一大难题。负责事件分拨的派单员需掌握约 40 多个部门的“三定”方案，1000 多项权责清单、事项清单，2000 多条法律、法规。单纯依靠人工分拨导致出错率高，疑难工单识别难度大、处置周期长，热线中心长期处于被动管理的状态。城市治理工单接入的渠道丰富，包括 12345 热线、小程序、APP 等，事件的类型、影响程度、严重程度以及涉及的部门需要大量的人工进行预判。采用 AI 技术进行语义分析，智能派单模型自动识别工单内容，并通过学习历史工单，掌握各类工单对应处置部门的专业知识，由此自动派发工单到相应处置部门，提升工单流转和处置效率。同时，智能派单模型针对新领域工单、疑难工单，能够识别工单内容、群众诉求、地址等信息，标识出疑难工单类型，并推送至派单员提醒其关注。另外，针对疑难工单，AI 平台能够基于历史典型工单的处理结果，向派单员进行相似案例的推荐，帮助其提高疑难工单的分析能力和处置效率。

4.1.3 效能智能分析

在城市治理领域，往往由于缺乏科学、全面的评价体系，导致城市管理人员疲于完成事件巡查、事件分拨和事件处置等海量工作任务，治标不治本。基于 AI 技术，构建城市治理效能分析能力，围绕“一件事”的评价目标，以工单量和满意度等维度对城市治理效能进行评价，通过对每日事件相关的数据进行挖掘，分析各类民生诉求渠道汇聚的事件数量、事件类型、处置情况等关键要素，辅以数据可视化图表的形式生成智能报告，通过“每日一报”的方式揭示城市运行的堵点、难点问题，协调推进问题整改。以 AI 智能分析作为城市运行的晴雨表，实时反映城市运行状态、老百姓的热点诉求、政府工作的短板，以评促改、持续优化。

▶ 4.2 气象预报

气象预报从最初的玄学，纯粹依靠生活经验判断，到后来传统天气图的诞生，天气预报开始变为应用科学，再到现在的数值天气预报和 AI 预报，预报准确率大幅提升，经历了人工经验、信息化、数字化和智能化四个阶段，如图 4-4 所示。



图 4-4 气象预报发展史

现代气象预报是气象台(站)运用现代科学技术(如卫星、雷达等)收集全国甚至全世界的气象资料，根据天气演变规律，进行综合分析、科学判断后，提前发出的关于未来一定时期内的天气变化和趋势的报告。准确及时的气象预报在经济建设、国防建设趋利避害，保障人民生命财产安全等方面有极大的社会和经济效益。

按气象预报的时效长短，可分为：短时预报、短/中期预报、气候预测。短时预报是根据雷达、卫星探测资料，对局地强风暴系统进行实况监测，预报未来 1 ~ 6 小时的动向。传统的短时预报主要是用雷达监测资料外推的方式开展，即根据估算前几天的雨区今天会移动到哪里。短、中期预报指预报未来 1 ~ 15 天的预报，传统的短、中期天气预报根据大气运动的规律数值化成偏微分方程组，根据初始的各个站点的气象数据，带入方程组中，利用超级计算机(HPC)不断循环迭代计算未来时刻天气。气候预测指未来 1 个月到数年的预报，传统气候预测主要应用统计方法，根据各月气象要素平均值与多年平均值的偏差进行预报，用数值预报方法制作长期预报的方法也在试验之中，已有了一定的进展。

AI 气象模型在过去几年取得了显著的进展，在短时，短、中期和气候预测领域都有探索。AI 气象模型通过处理大量的气象数据，如卫星图像、气象观测数据和气候模拟模型数据，挖掘天气系统的规律。这些模型能够识别和分析各种天气模式、气候变化和极端天气事件，并提供准确的天气预报。AI 气象预报的演进如图 4-5 所示。



图 4-5 AI 气象预报演进

视觉和科学计算两大模型的能力可以为三大类 AI 气象预报场景提供有力的支撑，高效率的开发出场景适用的业务模型，如图 4-6 所示。



图 4-6 大模型两大能力支撑不同业务场景

AI 气象预报在短临预报、中短期预报和气候预测三个方向未来会持续进行更多的探索，为行业的数智化转型提供有效的支撑，预计未来 3 年市场空间近 50 亿。

4.2.1 短时（短临）预报

传统临近预报算法主要以雷达监测数据通过光流法等方法进行外推平移的形式呈现降水云团的移动，其无法利用长期积累的历史气象大数据，同时在实际的场景中，影响雷达云图运动的因素肯定不止反射率因子信息，风向风速、温度、湿度、气压等对云的移动与变化都会有影响，这是传统外推法的局限性。

而基于人工智能的短临预报算法则通过构建神经网络对多年的雷达观测资料完成模型训练，把风向、风速、温度、湿度、气压等信息也放入模型中进行训练，则可清晰地预测降水云团的位置移动和强度变化趋势，与实况更为吻合。人工智能和机理融合的方式也有积极的成果，如通过端到端建模降水物理过程的神经演变算子，来实现深度学习与物理规律的无缝融合。

4.2.2 短、中期预报

在传统短、中期天气预报的制作过程中，有两个关键的组成部分都需要利用大规模高性能计算（HPC）集群进行模拟。一个是数据同化，用数值模型从卫星、雷达和探空气球等观测数据中推导出采样数据，进而估测真实数据，以提供预报模型的初始场；另外一个是通过数值天气预报（NWP）系统建立预测天气相关变量将如何随时间变化的模型，即数值预报模型。然而，随着数据量的显著增加，数值预报模型却无法得到有效的扩展。也就是说，虽然现在有大量的观测数据，但我们却很难直接利用这些数据来提高预报模型的质量。而改进的方法，一般是由训练有素的专家手动创造更好的模型、算法和近似值，这个过程耗时耗力，成本高昂。

数据驱动的 AI 短、中期预报模型可以通过学习长时间序列高质量的再分析数据来更全面的掌握物理规律，做出高质量的气象预报，而推理计算资源却低得多。业界 AI 短、中期预报模型当前主要基于欧洲中期天气预报中心 60 余年高质量的再分析数据来开展训练。

4.2.3 气候预测

气候预测是根据过去气候的演变规律，推断未来某一时期内气候发展的可能趋势，传统气候预测可以分为两类，一类采用统计方法，另一类为动力学数值预报，即采用多初值和多模式的集合预报方法，因而从本质上看，气候预测是一种概率预报。虽然准确率不高，但因气候变化对粮食生产、能源供应以至人类活动等都有明显的影响，因此对未来气候变化的预测已引起人们的普遍重视。

AI 在气候预测方面利用深度学习、卷积神经网络等技术，对海量的气象观测数据、卫星图像、雷达信号等进行快速处理和分析，从中提取有用的特征和信息，为气候模型的输入提供更高质量的数据。另外利用机器学习、强化学习等技术，对传统的物理方程式或统计方法构建的气候模型进行优化和改进，从而降低模型的误差和偏差。最后和短、中期气象预报类似，通过机器学习的方法来构建数据驱动的气候预测模型，世界气候研究计划的气候数据集可作为 AI 气候预测模型的训练基础。

▶ 4.3 智慧医疗

医疗行业是促进国家长远健康发展、提高人民健康水平、保障和改善民生的战略性行业，行业长期存在医疗专家资源稀缺与医疗诊断能力不足的公众医疗问题。伴随人工智能、5G、大数据等新兴技术的快速发展，医疗行业受到深刻影响，此问题也得到了初步改善，AI+ 医疗场景应用可以进一步有效缓解医疗资源分布不均、数据价值利用程度低、数据标准不统一等问题，推动医疗行业从最初的电子化、单系统应用，逐步向数字化、智能化不断演进，如图 4-7 所示。

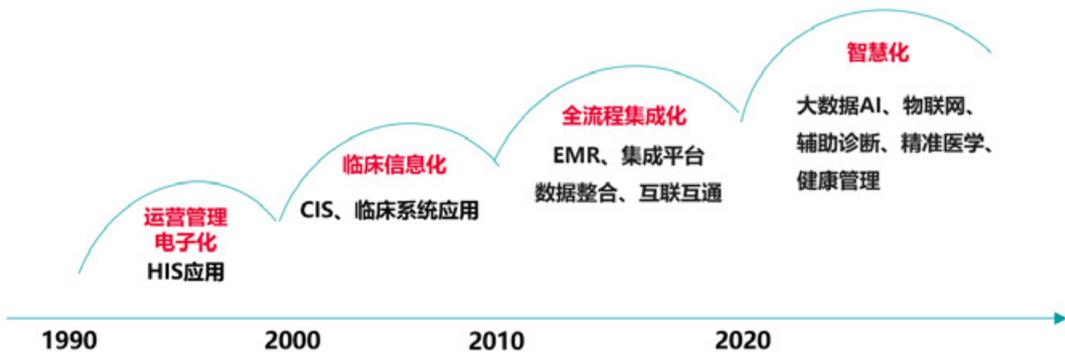


图 4-7 医疗信息化发展历程

AI+ 医疗的融合发展共经历了四个关键阶段。从二十世纪七十年代起，人类就开始尝试通过 AI 辅助医疗，提升诊疗效率和质量，改善医患关系。二十世纪八十年代，AI 研究聚焦临床决策专家系统，建立临床知识库模仿医生决策过程。二十一世纪后，医学成像设备逐渐成熟，AI 应用开始聚焦于辅助医生提取复杂的多维医学影像数据特性，完成诊断分析。伴随具备人工智能的医疗器械获批上市，AI 也逐步用于辅助医疗诊断。至今，AI+ 医疗逐步走向成熟，大模型技术横空出世，扮演起重要角色。AI 技术开始渗透到医疗各阶段及场景中去，扩展医疗健康服务边界、改变诊疗路径并改善医患关系，如图 4-8 所示。



图 4-8 医疗行业人工智能应用场景

AI+ 医疗是指医疗健康的生命周期内，在协助人或解放人的状态下，以提升院内外医疗服务效率为目的、以人工智能技术为手段干预到传统的院内外医疗环节的一种新型辅助技术。院前，人工智能技术可以应用于药物研发、基因检测等医疗健康管理环节，也可以应用于医疗环境监测、患者预约挂号、智能分诊等场景改善患者院前就医体验。院中，AI 可以实现影像辅助诊疗、辅助病理诊断、精准医疗等，减少医生工作量，提升诊断效率和诊断质量。院后，AI 通过健康管理、知识问答等功能，协助患者进行健康管理，从被动治疗转向主动预防。如今 AI 已经介入医疗领域的诸多场景，为医学领域的技术发展与医疗服务提供了重要支撑，但是仍存在场景单一、碎片化严重、模型维护成本高、模型参数量小、应用范围狭窄等问题，华为通过研发具备超强理解能力，可完成更多通用性任务的盘古大模型来解决上述难题，如图 4-9 所示。

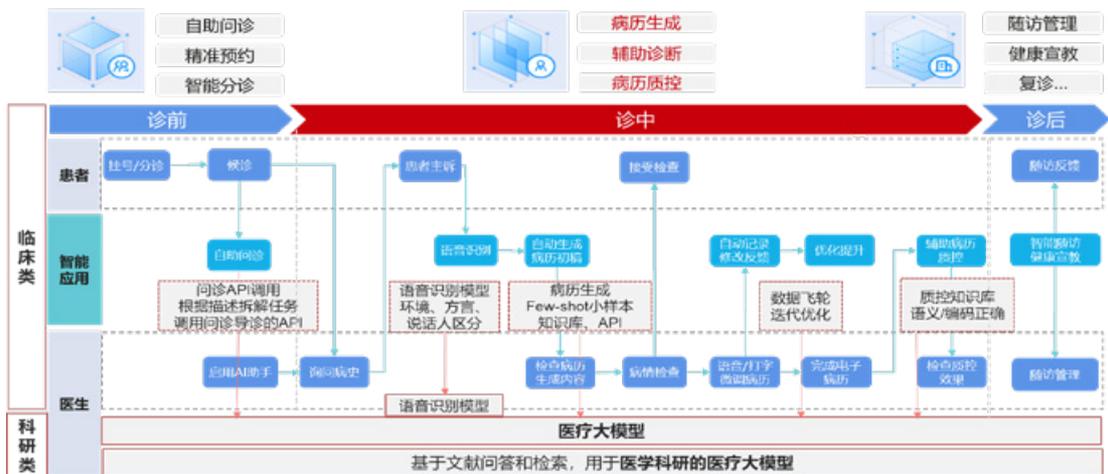


图 4-9 大模型支撑的医疗业务场景

未来智慧医疗伴随 AI、大模型等技术不断发展，以及信息技术与医疗服务深度融合，必将在智慧影像、精准医疗、药物研发、知识问答等场景提供更加有力的数字化、智能化技术支撑，预计未来 3 年市场空间近 70 亿，可有效促进生态有序发展和商业繁荣。

4.3.1 智慧影像

医学影像是指为了医疗或医学研究，对人体或人体某部分，以非侵入方式取得内部组织影像的技术与处理过程。医学影像使许多疾病都能看得到、看得准，从而使患者可以得到及时有效的治疗，这是现代临床诊疗体系中不可或缺的一环。但是对于医院来说，医学影像也存在着专业医生缺口巨大，存在误诊漏诊的现象，诊断速度有限等问题。

人工智能医学影像，将 AI 图像识别领域不断取得的前沿性突破技术应用在医学影像领域，解决上述医学影像诊断痛点。通过人工智能技术将医学影像进行 3D 重建，治疗过程中，影像显示病灶的对比分析，如图 4-10 所示。

自动3D重建

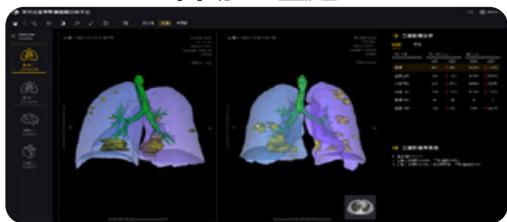


图 4-10 医学影像进行 3D 重建

对比分析

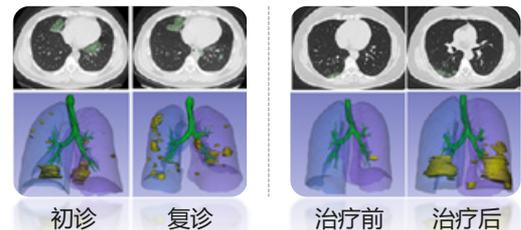


图 4-11 病灶显示对比分析

在医学影像 3D 重建方面，AI 通过图片特征点提取、分层建模、精准测量等技术，展示直观、精确的 3D 医学影像；在病灶识别与标注方面，针对 X 线、CT、核磁共振等医学影像的病灶进行图像分割、特征提取、定量分析、对比分析等，为影像科医生阅片提供参考，帮助医生发现难以用肉眼发现和判断的早期病灶，大幅降低漏诊及误诊；同时大幅提升影像医生诊断效率，十万张以上的影像处理仅需数秒。

通过使用人工智能技术，可有效提高医学影像诊断的精准性、标准化和自动化。一图胜过千言万语，对医学图像的理解与诊断需要有医学知识库、图像诊断模型和特征比对策略等关键条件的支撑。通过对大量图像数据和医学知识的训练学习，医学影像智能辅助诊断系统能够快速识别不同的病症图像，定位病灶组织，为患者提供高质量的检查报告。并基于深度学习不断优化，通过大量已有的影像数据和临床诊断信息训练，在目前诊疗体系的基础上进一步

降低复杂疾病的误诊率及早期病灶的漏诊率，从而带来医学影像总体诊断水平的提升。

4.3.2 精准医疗

随着信息化手段的不断提升，以及人们对医疗健康的需求越来越高，未来专业化的医疗服务将会不断的升级迭代，精准医疗的概念也应运而生。精准医疗是依据患者内在生物学信息以及临床症状和体征，对患者实施关于健康医疗和临床决策而量身定制的医疗服务。旨在利用人类基因组及相关系列技术对疾病分子生物学基础的研究数据，整合个体或全部患者临床电子病历，为患者提供定制化治疗解决方案的新型医学模式。其本质是利用基因组特征、人工智能与大数据挖掘、基因检测等前沿技术，对大样本人群和特定疾病类型进行生物标记分析与鉴定，找到精确发病原因和作用靶点，并结合病患个人的实际身体状态，开展个性化精准治疗，提高疾病预防与治疗效果。

精准医疗主要包括基因测序、细胞免疫治疗和基因编辑三个层次。其中，基于大量细胞和分子级别的基因测序是精准医疗的基础；对免疫细胞进行功能强化与缺损修复是精准医疗在疾病治疗领域的常见应用方法，目前 CAR-T 和 TCR-T 疗法备受关注；对变异细胞进行批量改造治疗的基因编辑技术则为精准医疗的高阶应用层次，技术壁垒较高。

4.3.3 药物研发

一款创新药从研发到上市，平均成本超过 10 亿美元、研发周期大于 10 年——这是医药界公认的“双 10 定律”。同时，药物结构设计强烈依赖专家经验、新药筛选失败率高，如何通过大数据、人工智能等科学技术加速新药研发进程、平衡研发投入与成果产出之间的关系，成

为了医药公司在数字化改革道路上的重点之一。

将深度学习应用于药物研究，通过大数据分析技术快速准确地筛选出合适药物成分，从而缩短新药研发周期，降低研发成本，提高研发成功率。其中，在靶点筛选、药物挖掘、患者招募以及药晶预测等关键研发环节中，人工智能技术至关重要。人工智能通过自然语言处理、图像识别、及其学习等技术，可以大幅度缩减研发流程，如图 4-12 所示。

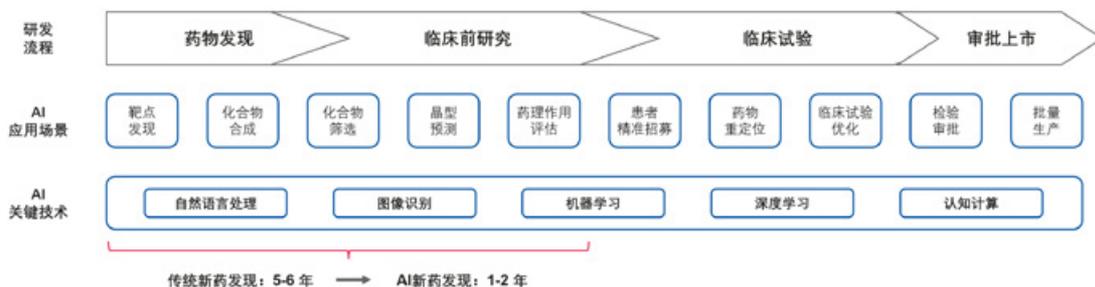


图 4-12 药物研发流程

药物研发主要是通过靶点发现、化合物合成、筛选、晶型预测等流程进行。通过人工智能的应用，可以在靶点发现阶段利用 NLP、DL 技术，快速识别患者数据中的身体变异数据，通过靶标数据库高效定位潜在的靶标分子。化合物合成 + 化合物筛选阶段，利用 ML、DL、CV 技术，参与到靶标三维结构预测、化合物从头设计、成药性预测及优化、药物虚拟筛选等过程中，大幅降低新药研发的时间和成本。晶型预测阶段，AI 结合实验的晶型预测和筛选技术可以在 2~3 周内预测出潜在的最佳药物晶型，有效加速研究和决策过程，大幅降低固体形态在后期应用的风险。

▶ 4.4 智慧水务

智慧水务是智慧城市理念在水务行业的延伸，是智慧城市的重要组成部分，更是未来水务现代化的基础支撑和重要标志。在智慧城市发展的时代浪潮下，对水务信息化的建设提出了更高的要求，同时也为水务信息化发展提供了前所未有的机遇，智慧水务是水务信息化发展的高级阶段，是数字经济环境下，实现数字化转型的必经之路。

智慧水务业务主要包含水安全、水资源、水环境、水生态、水工程和水事务等领域。水是人类生活的源泉，而随着城市的发展和全球生态环境的变化，水灾害问题频发，水污染问题也越来越严重，水资源短缺加剧，水业务的监管和治理也成为城市发展的一大困扰，水灾害、水污染、水资源等更是直接关系到民生问题，影响到人民的生命安全和财产安全。智慧水务业务如图 4-13 所示。



图 4-13 智慧水务业务图

智慧水务是运用人工智能、物联网、云计算、大数据等新一代信息通信技术，充分发掘数据价值和逻辑关系，促进水务规划、工程建设、运行管理和社会服务的智慧化，提升水旱灾害的防御能力和水资源的利用效率，改善水环境和水生态，保障国家水安全和经济社会的可持续发展。智慧水务技术架构如图 4-14 所示。



图 4-14 智慧水务技术架构图

近年来，随着云计算和大数据技术在水务行业的发展，算力的提升和算据的积累，使围绕机器学习和深度学习的人工智能技术在智慧水务的应用取得了突破性的进展。即利用人工智能、数据分析等技术，实现对城市、河流、湖泊、湿地、水源地和水利工程等区域的水文监测、水资源调度、水环境保护等智能化管理。AI 技术的引入，使水务工作人员能够更好地应对复杂多变的水文环境，优化水资源配置，保护水生态环境。人工智能技术在水务行业的应用场景如图 4-15 所示。

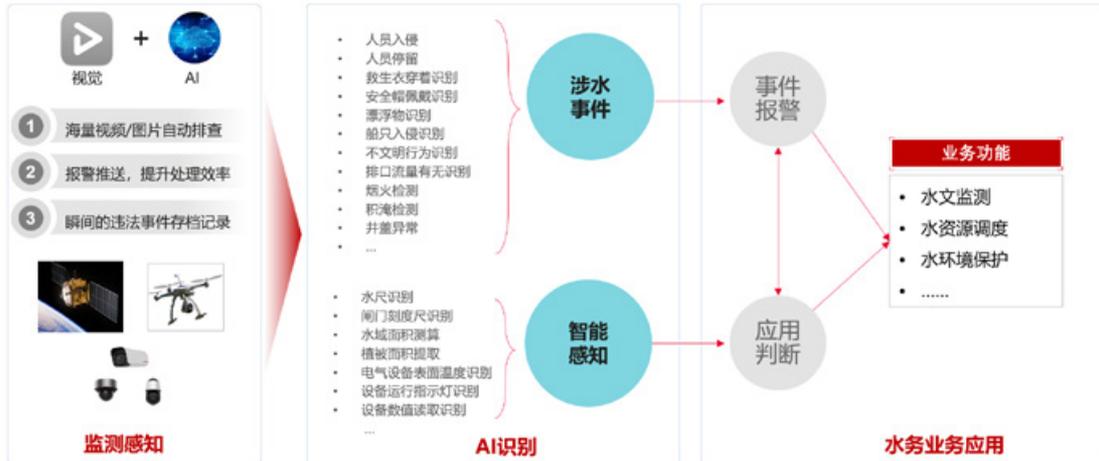


图 4-15 人工智能技术在智慧水务的应用

华为智慧水务 AI 解决方案采用 AI 技术和计算机视觉技术相结合，通过对视频图像进行特征分类学习、识别和分析，实现对河道区域内乱扔垃圾、倒排污水等不文明涉水行为的自动抓拍、城市内涝积水自动识别预警和水情信息智能监测等智能化识别。通过自动发现问题，为水务业务应用提供更全面、更快捷的感知能力，减少需要大量人力排查的工作。华为智慧水务 AI 解决方案架构图如图 4-16 所示。



图 4-16 智慧水务 AI 解决方案架构图

同时基于深度学习等智能算法在水文遥感数据的处理和信息提取、山洪灾害影响因素识别和风险分析、洪水计算模型调参、洪水预报预警和洪灾风险评价等方面有着比较大的优势和前景。

智慧水务 AI 解决方案为水务行业的智能化、数字化提供了强有力的支撑，未来在水资源管理、水安全预警、水环境监测和水工程规划等领域发挥越来越大的作用，进一步提升水务业务智能化水平和精细化管理能力，预计未来 3 年市场空间近 30 亿。

4.4.1 城市内涝预警

近年来，由于短历时强降雨或持续强降雨等极端天气导致城市内涝的现象渐趋严重，例如今年北京特大暴雨、2021年郑州特大暴雨等均导致了严重城市内涝灾害。暴雨内涝已经成为中国城市频繁发生、损失严重且影响较大的灾害，影响人们正常的生产生活，甚至对人们的生命财产安全产生了严重威胁。如何做到及时有效预警并通过一定的智能化手段来实现城市内涝从被动防御转向主动防御，是当前面临的重要课题。

城市内涝是因为高强度的降水或者连续性降水，超过了城市的排水能力，使城市内的一定区域产生积水而引起的灾害现象。城市内涝易发区有城区低洼地区、下凹式立交桥、地下轨道交通、地下商场与地下车库等。

目前城市建设大量的视频监控，主要用于安防、交通等，可充分共享这些视频，同时补齐易涝区域盲区的监测，构建城市内涝积水等智能视频分析模型。对城市内涝易发区积淹情况进行及时预警，将通知发送至相关监管部门和市民，将可大大提高城市内涝实时监测覆盖范围以及监测预警能力，及时采取相应防护措施，避免产生更大的损失和危害，助力城市内涝从被动防御转向主动防御转变。

4.4.2 涉水事件监管

目前全国流域涉水事件监管仍以人工巡河观察为主，如检查河道有无新增排污口、河道水体有无异味、河道设施是否完整、岸线是否存在四乱问题、是否存在非法电鱼等，涉及事件种类繁多、发生范围广，现有监管手段已经无法满足新形势下的要求。

随着视频 AI、遥感 AI 等技术发展，充分利用已建视频监控资源、遥感卫星资源，建立一套能够利用计算机智能分析和理解音频、遥感和视频的模型库。通过从海量存储视频、遥感影像中挖掘出人工难以获取的信息，为后续的事件搜索预测提供了科学依据，避免视频建设了但是无人看的问题、流域监管范围广监管难等问题，把工作人员从单调、重复、机械的被动巡查工作中解脱出来。有效实现水资源变化、水安全防御违规行为主动监测，形成水、物、人全方位的监测，实现大规模智能分析任务的调度、在线和离线分析、内容管理等，从而弥补传统的传感器和人工在处理涉水业务处理上的不足。遥感识别模型可实现河湖“四乱”、生产建设项目扰动、地表水体、土壤墒情、岸线变化、生态补水、下垫面信息、农业灌溉取用水等涉水事件发现；视频识别模型可实现采砂船、安全帽、漂浮物、垃圾堆、人员入侵等事件识别。

4.4.3 水情智能监测

水情是水务的重要感知监测要素之一，对于日常的水资源调度、汛期的防洪排涝至关重要。目前该数据监测手段主要有自动监测、人工监测两种方式。

自动监测依靠传统压力式水位计、雷达水位计、气泡水位计、浮子水位计进行水位自动监测并回传数据中心。其技术优点是比较成熟稳定，缺点是易受温度、湿度及风浪等影响，安装及维护成本较高。

人工监测常应用在防汛、灌区计算等业务场景中，需要人定时到现场通过仪器或读取水尺的方式。如在一些灌区监管部门，每天在 8 点、12 点、16 点派人在渠首读取水尺水位数据后，

通过水位流量关系计算出渠首的来水量，在特大暴雨情况下，受天气影响，会造成人员读数不准确，误差增大，对人员安全也会造成威胁。

为了解决以上问题，国内已有许多厂商实现基于水尺图像的智能识别，相对传统人工或自动监测方式，提高数据采集的效率，降低了安装成本及实施难度。虽然此技术也得到一定程度的应用，但受识别算法、识别模式、安装环境等因素影响，水尺识别的准确性表现仍不佳，且基于图像识别方法对每个摄像头的安装角度和位置要求较高，适应性不够。

随着大数据、大模型等技术兴起，基于人工智能水尺识别将成为趋势，即使用人工智能开发平台，对大量水尺图像数据采集及筛选、数据标注及模型训练等深度学习后，使水尺刻度的识别算法可适应不同新安装的现场环境，做到安装摄像头即可使用，无需专业技术调试，后期升级模型算法即可。同时通过人工智能方法也大大提高水尺识别准确性，可达到平均误差不超过 1CM，最大误差不超过 2CM 的精度。因此基于人工智能水尺识别，相对传统人工读数和自动采集、常规图像识别，具备采集快、识别准、易适应等特点，可广泛应用到水情监测中。







加速智能化 转型案例

第三篇



第五章 制造和大企业

▶ 5.1 冠捷科技

5.1.1 案例概述

冠捷科技是全球的大型高科技跨国企业，旗下拥有多个显示行业自有品牌 AOC、AGON 和 Envision 等，并长期获得飞利浦独家授权运营其显示器、电视及影音等业务。为了应对工业 4.0 的浪潮来袭，冠捷科技加快“制造”向“智造”转型，在各种场景中引入工业 AI 解决方案，为未来业绩长期增长注入动力。

冠捷科技多个工厂的 SMT 产线均采用 AOI 设备，针对线路板的贴片和焊接质量进行炉前和炉后的检查，原 AOI 设备直通率为 50% 左右，极大消耗人力。公司希望通过 AI 技术，大幅度降低 AOI 设备误判率，大幅度降低人工复判的比例，实现降本增效。

冠捷科技联合华为、博瀚智能，将基于华为昇

腾底座和盘古大模型的解决方案落地至冠捷多家工厂。该解决方案可检测移位、漏件、侧立、立碑、假焊、连锡等 10 余种 SMT 缺陷，覆盖客户所有 SMT 缺陷检测。该解决方案在不更换 AOI 设备、极少改造产线的前提下，降低了 80% 的人力，同时实现 AI 系统由工厂内部员工运维，取得节省人工、提升准确率、AI 系统自主运维的多重收益。

5.1.2 解决方案和价值

华为联合博瀚智能于业内首创基于“数据驱动+MLOps”的自适应 AI 质检系统，深度适配盘古大模型，通过 AI 模型在云端（集团侧）开发，在边侧（工厂）和端侧（车间）应用。“昇腾制造”PCBA AI 质检解决方案整体架构与检

测优化流程如图 5-1 所示。

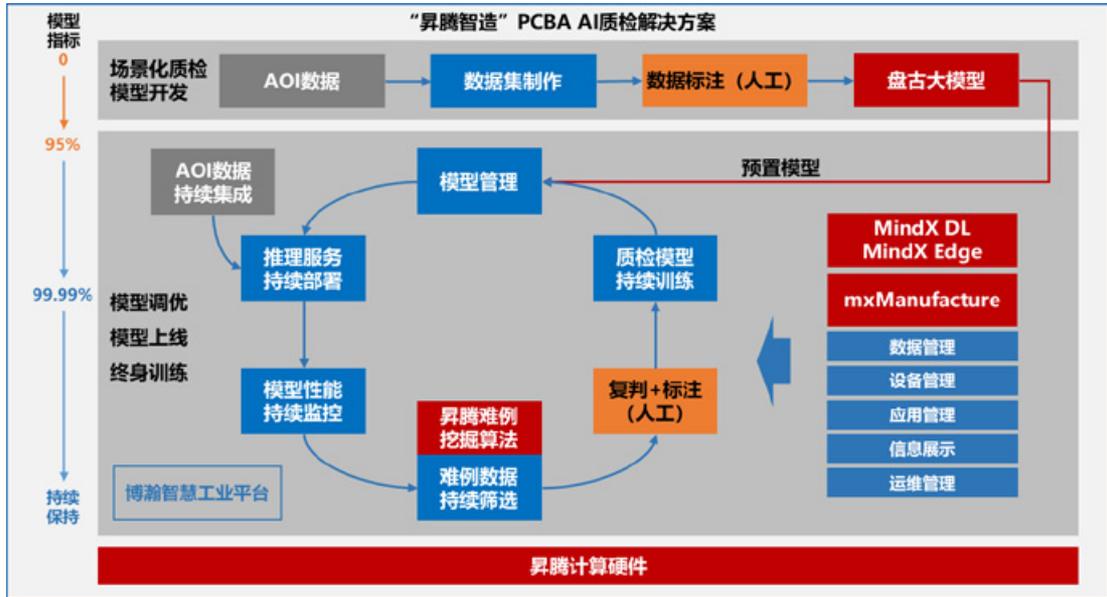


图 5-1 PCBA AI 质检解决方案

该解决方案在 AOI 检测后，加入 AI 质检，对 AOI 判断为 NG 的点进行二次检测，通过深度学习技术训练出具有极低误检率和超高检出率的 PCB 缺陷检测模型，从而有效降低后续人工复判工作量，节省 QC 成本。基于昇腾算力底座，实现业内首个基于“数据驱动+MLOPS”的自适应 AI 质检系统的落地。解决方案充分发挥昇腾 AI 硬件的澎湃算力和盘古大模型的优势，结合博瀚智能的自适应 AI 质检系统，有效实现 AI 系统终身持续学习，克服业内现有方案上线周期长、模型持续优化成本高的普遍痛点，使缺陷检测模型实现自学习和自迭代。

在缺陷类型方面，解决方案可检测移位、漏件、侧立、立碑、假焊、连锡等 10 余种 SMT 缺陷，覆盖冠捷科技所有 SMT 缺陷检测，在不更换 AOI 设备、极少改造产线的前提下，将模型开发和上线效率提升 5 至 10 倍，降低工厂和车间侧模型适配和运维成本超 80%，降低人力成

本超 90%，如图 5-2 所示。解决方案同时实现 AI 系统由工厂内部员工运维。取得节省人工、提升准确率、AI 系统自主运维的多重收益。

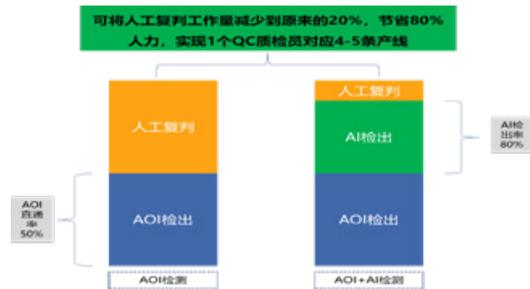


图 5-2 AOI+AI 检测效果对比

经过 50 天的试运行和模型调优，模型的缺陷检出率可稳定保持在 99.9% 以上，过滤率在 85% 左右，远远高于项目最初定下的预期目标。不仅实现了人力成本的节省，同时产线的缺陷检出率也大幅优于完全依赖人工的水平，从源头上避免了缺陷流至下一个环节。解决方案带来的具体价值包含以下几点：

一、产线改造成本低：解决方案使用现有的 AOI 设备获取图像，无需增加或调整成像设备。AI 检测的时间小于 300ms，不影响产线原有节拍。AI 检出的不需要人工复判的部分，通过接口直接将结果上报到 MES 系统，无其它软硬件对接成本。对于约 10% 需要人工复判的部分，工人复判方式与原来一致，无需进行复杂的产品培训。

二、大幅缩减 QC 投入：AOI+AI 双重检验过滤后，只有 10% 的产品需要人工复判，仅为原来的五分之一，节省 80% 人力，实现 1 个质检员对应 4-5 条产线。AI 检测的漏检率低于 0.01%，相比质检员更稳定可靠。同时，AI 的检测结果输出包含缺陷分类标签，对检出缺陷进行更精细化的分类、统计和分析，为工艺改进提供数据支撑。

三、在 AI 质检项目运维方面，传统 AI 质检项目受到手工模型升级、部署的限制，AI 系统精度常常波动，难以持续保持，如图 5-3 所示。

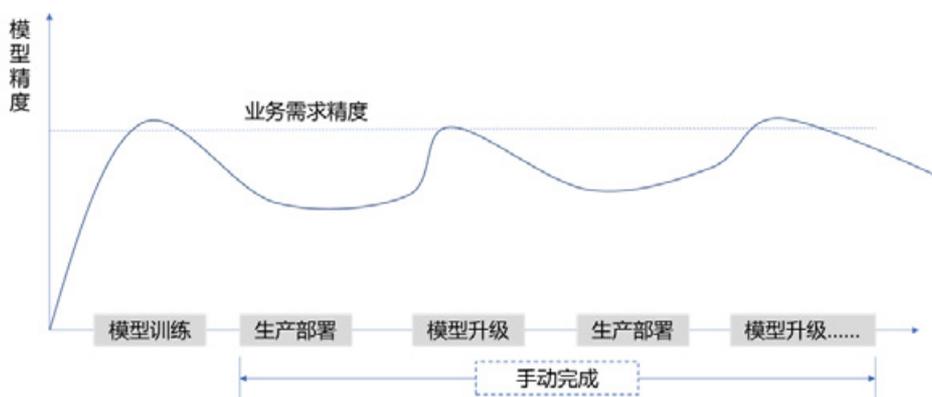


图 5-3 传统 AI 检测模型精度

华为与博瀚智能联合在业内首创基于“数据驱动 +MLOps”的 AI 质检系统，实现自动化的“数据持续集成”、“模型持续训练”、“模型持续部署”和“模型持续监测”。AI 检测无需手动完成模型升级和生产部署，可持续保持系统精度，如图 5-4 所示。

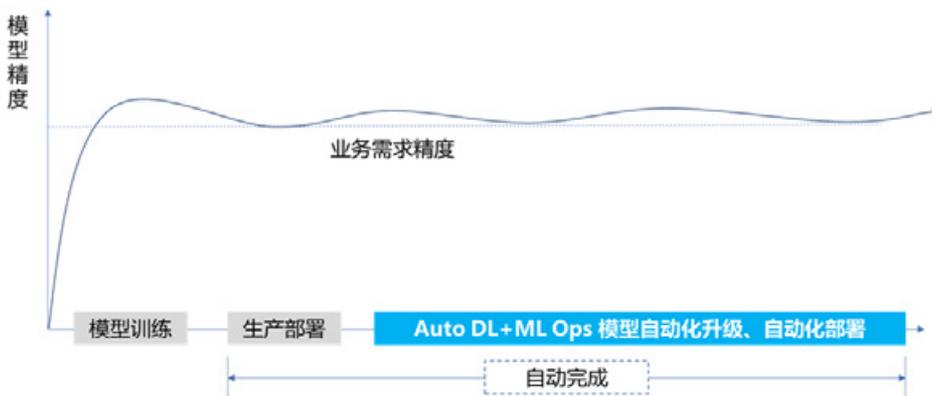


图 5-4 自适应 AI 质检系统自学习的模型精度

经过以上对比不难发现,实现“数据驱动+MLOps”,不仅避免了大量的手工模型升级和生产部署工作,大大降低运维成本,更使 AI 质检系统具备持续自学习能力,赋能 AI 质检系统在多场景下快速上线并持续保持性能高位。

5.1.3 总结和展望

在冠捷科技 PCBA 板 AI 检测项目中,华为与博瀚智能合作,充分支撑全场景 PCBA 质检应用,实现小样本训练和模型快速上线等功能,为大模型在工业 AI 质检领域广泛落地提供了新的范式和动能。冠捷科技基于该解决方案,后续将在全国范围的 PCBA 产线进行大规模推广复制。

与此同时,大模型技术在 PCBA 板 AI 检测中的成功应用,标志着大模型在工业 AI 质检领域打开了新的篇章。大模型的应用,使模型训练所需要的样本数量减少 50% 以上,训练周期更短,精度提升超过 1%,这不仅展示了小样本训练和快速上线模型的强大能力,而且为未来其他工业质检场景提供了先行经验。随着技术的进一步研究和完善,我们预期该模式将成为工业质检的主流做法。它不仅能够提高质检的准确性和效率,还可以为企业节省大量的时间和资源。在全球制造业日益追求智能生产和自动化的背景下,大模型技术的应用无疑会为企业带来更强的竞争优势,推动整个行业迈向新的高度。

► 5.2 华为南方工厂

5.2.1 案例概述

华为的制造体系每年都要处理亿级的排产,以往主要靠业务规则和人的经验。由于排产人员的经验和出发点不同,会出现各种问题,如资源冲突或浪费,导致订单延误。现在,华为已经在车间、线体、设备和工序等制造要素上实现了全面的数字化,并结合科学调度理论,构建了各种场景的排产模型,实现了自动化排产。

华为南方工厂是华为生产交换机、服务器、手机等终端产品的主要生产基地之一。华为南方工厂在制造领域已经全面应用了智能排产调度技术。

以单板温循为例,温循是指对单板进行高低温循环测试的过程。由于不同单板工艺要求的差异,不同批次的单板到达温循的时间也不同,且单板温循所用的测试装备存在共用情况等原因,单纯依赖人工经验进行排产往往难以将温箱装满,导致温箱装载率低。业务高峰期,温箱经常出现短缺现象,有时甚至需要租用外部资源。

针对该问题，华为自动化部研发了温箱智能排产算法，成功替代了人工排产，使任务令分批、温箱及温循时间推荐能够自动决策。华为南方工厂在某车间上线温箱智能排产应用后，温箱利用率提升了 17%，生产高峰期外租温箱的需求从 50 台降低到了 0 台，节省了上千万元的成本。

为了进一步实现多工厂生产计划排产综合最优，华为南方工厂也在积极探索“盘古大模型 + 天筹求解器”的大模型智能排产解决方案，实现智能排产从有限约束条件到亿级约束条件、工厂打包排产到多工厂协同、单目标牵引到多目标综合考虑这三方面的优化升级。

5.2.2 解决方案和价值

单板温循前需要确定温循的开始时间、温箱编号和温循时长。如图 5-5 所示，测试车容量、批次数量和温箱尺寸需要相互匹配。

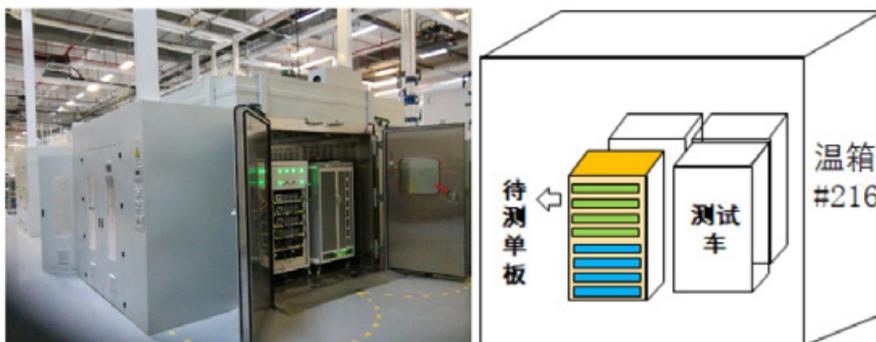


图 5-5 单板温循用温箱

解决这个排产科学问题难点在于，对于 10 个任务令要在 10 个温箱中按照任意顺序分配，排列组合数为 $(10!)^{10} = 3.96 \times 10^{65}$ 。这意味着即使计算机能够在 1 秒内列举出 2 亿个解，穷举所有可能的分配方案也需要约 6.27×10^{49} 年，相当于在有限时间内无法完成计算，考虑到任务令可能需要分批执行，计算复杂度将会更高。因此，必须设计高效的算法来解决此类问题。

华为南方工厂参照批次组合调度 + 并行机调度模型构建了实际的温循排产数学模型，决策变量、约束条件和目标函数如图 5-6 所示。

(1) 决策变量：温循任务令批次与温箱、测试装备的对应关系、温循批次大小、温循批次开始 / 结束时间。

(2) 约束条件：上下工序批量交接约束、批量连续测试约束、冻结排产约束、停机时间窗约束、专用温箱约束、温箱占用约束、测试装备可用资源约束、任务令插单约束、任务令批次混测约束、温箱容量约束等。

(3) 目标函数：温箱平均装载率最大化、任务令完工时间超期最小化。

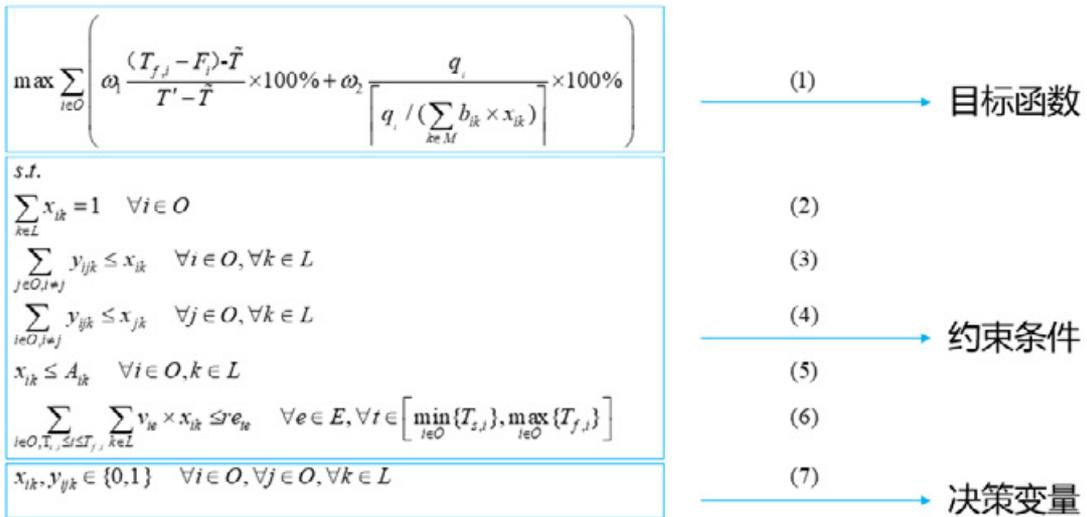


图 5-6 单板温循数学模型

针对以上数学模型，华为南方工厂采用了 ERT（Expected Residual Time，预期剩余时间）启发式规则 + 迭代贪婪算法解决了温循排产问题。具体来说，在预处理阶段，采用 ERT 启发式规则，对任务令的释放时间进行排序，优先处理最早释放的任务令。在迭代计算阶段，使用迭代贪婪算法，不断缩小搜索空间，逐步寻找最优解。为了提高算法的效率，华为南方工厂还采用了一些定向、剪枝等策略，对搜索空间进行压缩，从而有效降低了计算复杂度，实现了分钟级快速求解的目标。

在智能排产部分场景的算法实现方面，华为南方工厂也使用了天筹求解器为数学模型求最优解。求解器是一种计算机程序，用于解决数学

问题、逻辑问题或其他类型的问题。求解器可针对定义的数学模型快速分析和计算，找到问题在数学上的最优解或者近似解。

求解器的数学建模和算法开发有比较高的技术门槛，普通业务人员无法直接使用。为了进一步推广智能排产的应用，降低使用门槛，华为南方工厂积极探索大模型与求解器结合的智能协同决策方案。大模型善于捕捉和理解模糊的、复杂的和非结构化的信息；求解器善于处理确定性问题，执行逻辑运算和分析任务，对数学模型进行计算求解。华为南方工厂使用盘古自然语言大模型，辅助数学建模，将天筹的求解结果转换成业务人员能够理解的方案。具体的协同决策流程如图 5-7 所示：

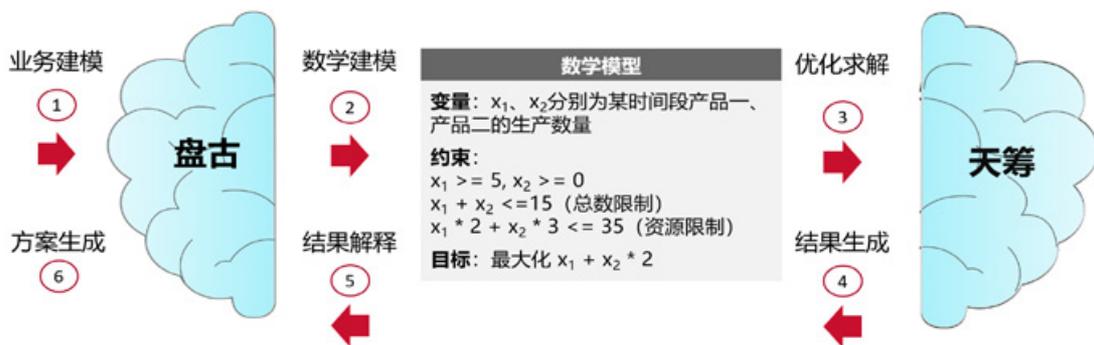


图 5-7 大模型与求解器协同决策流程

在该决策流程中，用户使用自然语言向盘古描述业务问题，盘古根据描述，将问题转换为数学模型，包括需解决问题的变量、约束和目标；用户确认数学模型后，盘古自动生成基于天筹求解器运行的算法代码；天筹求解器对问题优化求解后生成计算结果，如上图中的简单示例，优化结果为“ $x_1=10, x_2=5$ ”；盘古使用自然语言对于优化结果进行解释，便于用户理解；最后用户使用优化结果和描述信息生成方案。整个流程大大提高了自动化程度并降低了使用难度。

在大模型技术架构方面，华为南方工厂基于“昇腾 + 鲲鹏”打造统一算力中心，并在此基础上建立统一的华为 HIS 云平台。ModelArts 是基于云平台的统一 AI 平台，为大模型提供模型训练、模型推理和性能分析的完整端到端工具链，支持 L0 基础大模型、L1 行业大模型和 L2 场景模型的三层架构。南方工厂大模型技术架构如图 5-8 所示：



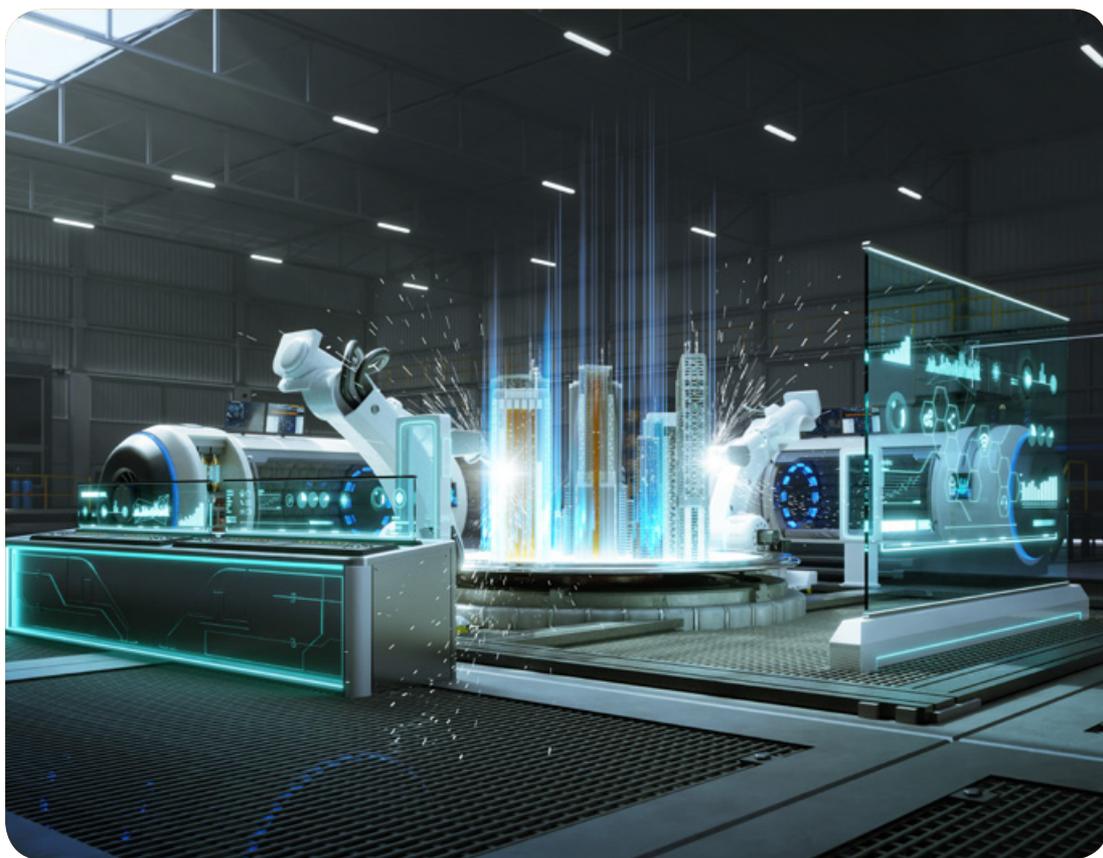
图 5-8 南方工厂大模型技术架构

这种 L0、L1 和 L2 的分层结构可以使得模型更易于理解和优化，同时也可以提高模型的泛化能力和效率。其中，L1 行业大模型是基于 L0 基础大模型加入行业数据进行微调得到的。华为南方工厂主要使用的是制造行业 L1 大模型，该 L1 模型可以直接应用于代码辅助、预测性维护和智能客服等场景。而 L2 场景模型是对 L1 行业大模型进行的压缩，保留数据关联性的同时降低了模型的复杂度和计算量。L2 场景模型常用于计算机视觉和预测性维护。

5.2.3 总结与展望

华为南方工厂通过各种资源的优化配置，应用智能排产进一步降低浪费，成功地缩短了产品上市周期，这不仅提高了经济效益，使生产更加适应了产品的多样性和生产流程的复杂化，满足生产环境不可预测的动态变化，增强了市场竞争力。

在华为南方工厂的智能排产应用过程中，求解器无疑是核心的使能技术，而大模型的应用，降低了求解器使用门槛，缩短了复杂技术与普通用户之间的鸿沟。展望未来，随着数据驱动和技术进步，以华为南方工厂为代表的华为制造，将进一步完善和扩展智能排产的应用范围，并完成智能排产系统与华为 MES+、MetaERP 的全面整合，实现更加精细化的生产调度、更智能的资源分配和更深入的供应链整合。这将进一步提升生产的敏捷性和弹性，使华为制造能更好地应对市场变化和不确定性，从而在行业竞争中保持持续领先，并推动其走向更高效、绿色和可持续的制造未来。





第六章 矿山及重工业制造

▶ 6.1 山东能源集团

6.1.1 案例概述

山东能源集团（以下简称山能集团或山能）是以矿业、高端化工、电力、新能源新材料、高端装备制造、现代物流贸易为主导产业的大型能源企业集团，位居世界 500 强第 69 位，煤炭产量位居全国煤矿行业第三位。

山能以科技兴安、少人则安、无人则安为指引，按照顶层设计、标准先行、分类施策、示范引领、有序推进的原则，如图 6-1 所示，于 2020 年启动智能化煤矿建设的统筹规划；2021 年建成 9 个国家级智能化示范煤矿；2022 年全面推进，建成 32 个一类矿井（高级智能化煤矿标准），43 个二类矿井（中级智能化煤矿标准）。当前山能集团已步入智能化建设深水区，亟需引入业界先进技术，保持自身智能化建设的领先优势。



图 6-1 山东能源集团智能化建设推进目标

2022年山能集团与华为技术有限公司（以下简称华为）立足各自产业和技术优势，建立全面战略合作伙伴关系，深度融合工业场景与数字技术，重点围绕煤矿智能化“少人、无人、安全、绿色、高效”的目标，开展20个课题的联合创新，积极探索人工智能、大数据、物联网、5G、F5G等新一代信息技术在煤炭产业的应用。

联合创新中心成立后，双方专家深入煤矿井下作业现场，了解井下生产时人、机、环、管的作业场景和安全规章要求，结合人工智能在井下应用的诸多问题，进行了充分的技术交流后发现，井下人工智能技术应用最大的问题包括如下几方面：

- 1) 负样本无法穷举，很多异常场景无法模拟，导致AI模型只能识别已经训练过的样本，新的异常场景无法识别，精度低。
- 2) 算法越用越差，AI模型无法自动迭代更新，长时间运行后，因为异常场景的增多导致精度会越来越差，算法误报、错报率高，导致生产

人员对算法普遍不信任。

3) 算法通用性差，AI模型在不同煤矿相同场景运行的时候，准确性差，必须在AI模型提供商的开发环境经过二次开发（甚至可能需要重新建模）和训练，开发和部署周期长。

4) 数据安全风险高，AI模型提供商必须将数据拿到自己的训练环境进行模型开发和训练，从而存在数据泄露的安全风险。

5) 煤矿客户普遍不具备人工智能开发和维护能力，AI模型出现问题后，只能等待厂家远程或者出差到现场进行问题定位和解决，问题解决时间长。

为解决以上问题，使人工智能应用大规模“下井”，山能集团与华为探索出了一套完整的可规模复制的工业化人工智能大模型解决方案，并通过实践形成了一批应用成果。

6.1.2 解决方案和价值

6.1.2.1 方案架构

基于双方在人工智能方向的探索与实践，创新性地采用了云边协同架构，并使用非正常即异常和边用边学的机制，成功解决了当前行业人工智能应用的难题。

(1) 云边协同架构

采用集团中心云和矿山边缘云两级架构。集团中心云建设基于盘古矿山大模型的人工智能训练中心、开发中心和运维中心，通过集团统建统管，避免了各单位重复投资和烟囱式建设，实现了集约化的公共训练算力服务。集团中心云使用矿山边缘采集的数据进行算法开发、模型训练，AI模型统一由集团中心云下发到矿山边缘云进行远程部署，并与应用系统通过标准化的AI框架平台进行集成，模型的推理信息与应用系统、生产系统实现联动，完成业务闭环，如下图6-2所示。

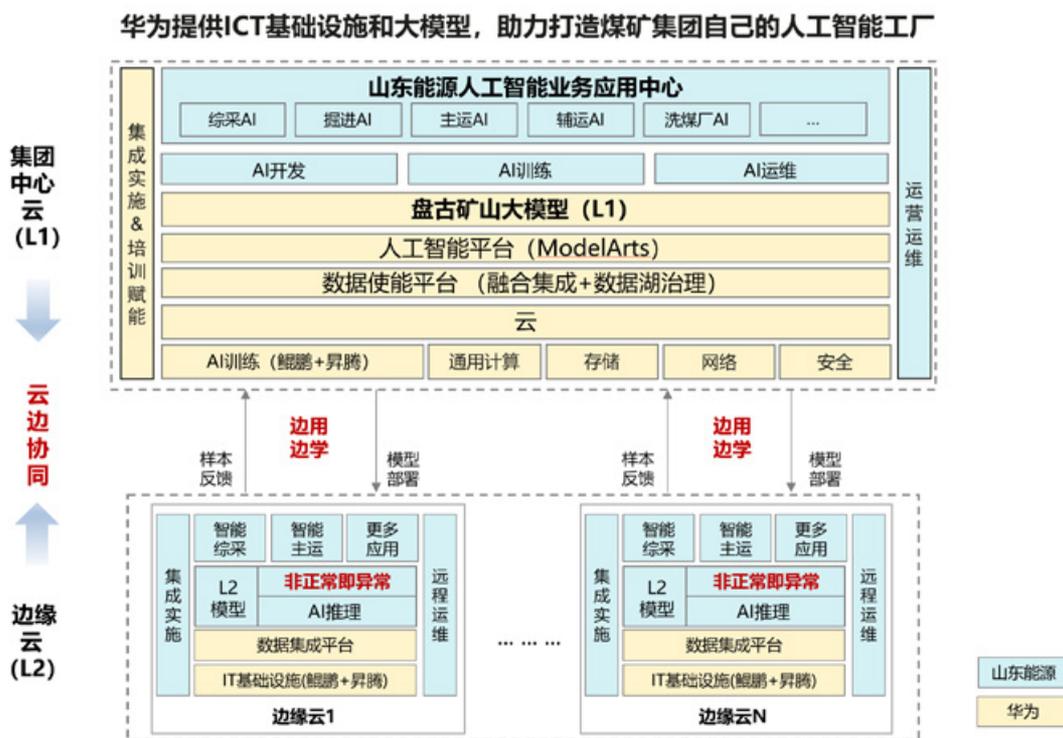


图 6-2 云边协同架构

(2) 非正常即异常、边用边学机制

非正常即异常：针对异常场景无法穷举的问题，通过学习大量的正常样本特征，在日常 AI 推理过程中将无法识别的异常样本提取出来（过滤掉 99% 以上的正常样本），彻底解决负样本无法穷举的问题。

边用边学：异常样本识别出来以后，上传到集团云大模型训练平台的增量样本库中，由行业专家进行人工判断标注，然后进行二次训练，并将训练完成的新版本算法下载到边缘云进行迭代升级，从而实现算法模型越用越好的能力。如图 6-3 所示。

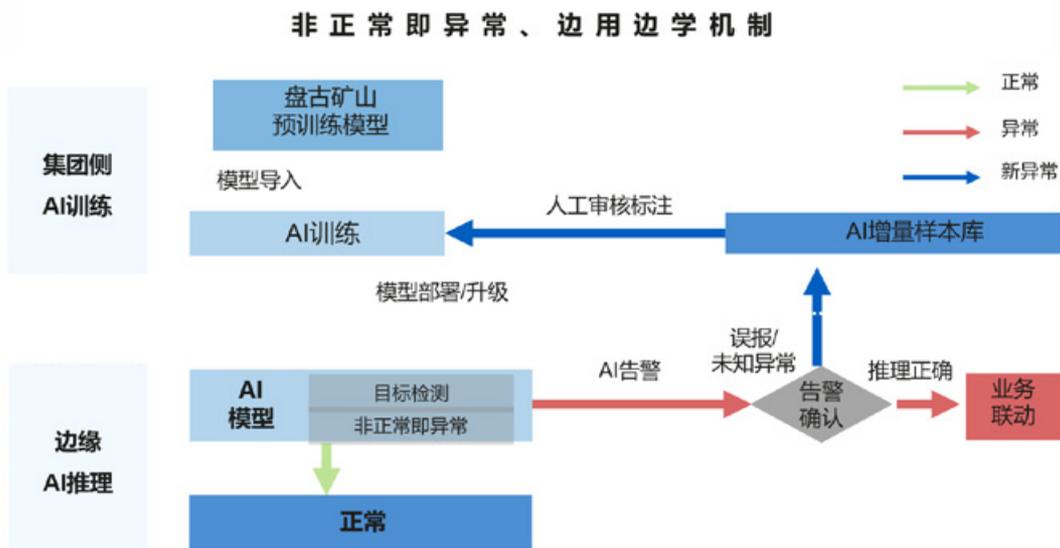


图 6-3 非正常即异常和边用边学机制示意

(3) 盘古矿山大模型：矿山行业预训练大模型，实现人工智能应用“工业化生产”

大模型“起点高”，小样本能力强，快速提升识别精度：盘古大模型通过海量数据预训练，学习大量矿工行业的特征，从而形成了行业大模型。在进行场景化 AI 算法开发时，通过少量样本的训练进行精调后即可开发出精度高的模型，相比传统模型识别精度提升 30% 以上。

大模型“见多识广”，通用性好，支持规模复制：盘古大模型通过海量数据的预训练以后，因参数量大，可以承载更多的类似场景，带来更好

的通用性。例如大模型对一百条皮带视频场景进行了预训练，那么基于大模型开发的皮带监测算法，如跑偏、大块煤、异物识别等场景的通用性要大大优于小模型，从而大大提升了算法的开发和应用效率，降低了开发和交付成本，实现模型开发产线化，推动人工智能应用从“作坊式”到“工厂式”升级。

大模型 + 工作流，降低 AI 开发门槛：提供向导式的算法开发、训练、操作平台，通过调用大模型自动调参（学习速率、批次大小等）和生成模型的能力，简化场景化 AI 模型开发难度，减少对专业 AI 开发经验的依赖。如图 6-4 所示。

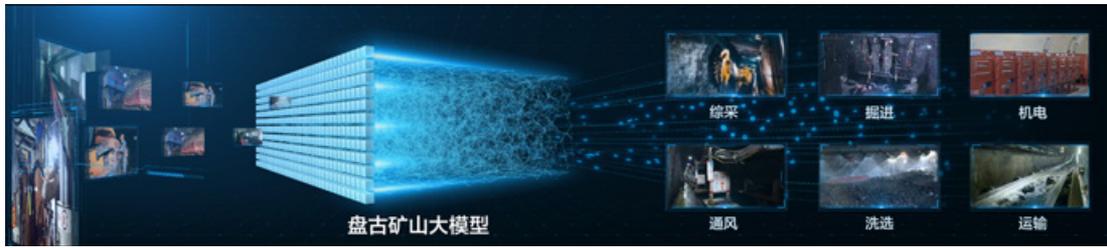


图 6-4 华为盘古矿山大模型

6.1.2.2 应用场景

当前，山能集团已在其下属的李楼煤业、兴隆庄煤矿、济宁二号煤矿等单位开发和实施首批场景化应用，涵盖了采煤、掘进、主运、辅运、提升、安监、防冲、洗选、焦化等 9 个专业 21 个应用场景，通过接入现场生产数据和监测视频，实时优化工艺参数、识别故障与异常、审核作业规范，以广播提醒、设备联动等方式实现了自动处置，形成了一批应用成果，部分场景介绍如下：

(1) 李楼煤业：防冲卸压施工孔深智能监管

山东是中国受冲击地压影响最深的区域，防冲卸压工程是冲击地压防治的重要手段，通过在采掘工作面等应力集中区进行卸压钻孔作业，可有效改善煤体内应力状态、降低冲击地压风险。

钻孔深度是防冲卸压工程关键的参数之一，是人工核验的重点。传统防冲卸压施工监管采用井下录制视频、井上对视频逐个人工核验的方式，存在审核工作量大、核验滞后时间长、标准不统一、监管有盲区等问题。

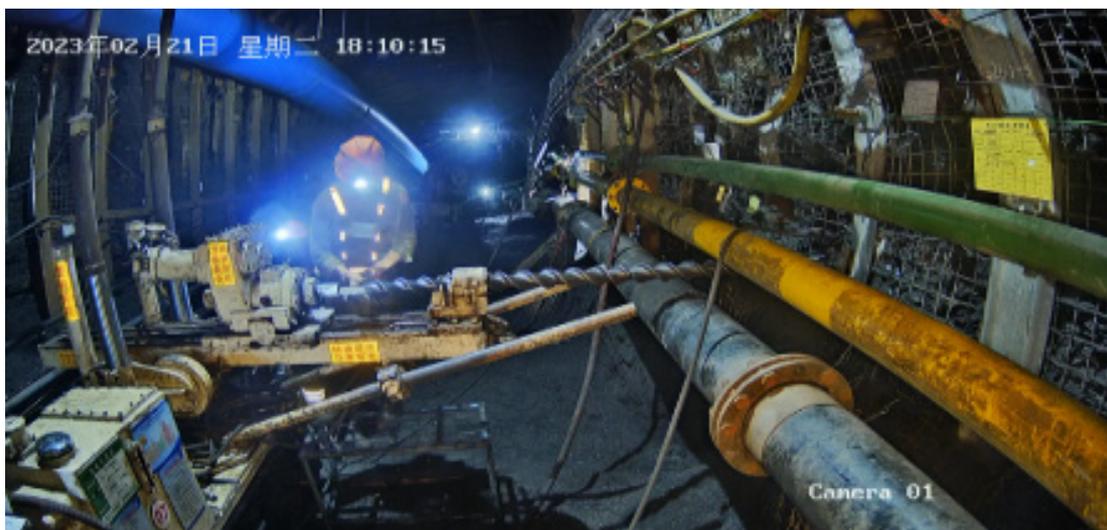


图 6-5 防冲卸压施工现场

通过在钻机上架设无线摄像仪，俯视拍摄钻机、钻杆和工人取杆动作，采用计算机视觉 AI 算法对打钻视频进行实时核验，识别有效钻杆，得到钻孔深度，对不满足施工要求及时告警，如图 6-5 所示。



图 6-6 防冲卸压施工孔深监管部署示意

如图 6-6 所示，该系统增强井下卸压工程的实时监管能力，实现井下卸压工程远程可视可管可控，做到工程质量当天审核。通过打钻深度智能监管，降低人工核验工作量 80%，实现卸压钻孔核验全覆盖，实现闭环管理，提升核验效率。

(2) 兴隆庄煤矿：掘进安全质量智能监管

掘进作业在安全管控上存在安全保障难、顶板管控难、行为规范难、质量保持难、责任追溯难等五大难点，也是制约掘进工作面实现智能化和安全高效生产的痛点。



图 6-7 防冲卸压施工现场

如图 6-7 所示，通过在掘进机机身前部、右部和侧面安装摄像仪，对掘进迎头作业区域进行实时监测和分析，以 AI 智能目标识别和动作识别为主，结合传感器及设备运行状态实现掘进工作面人员安全和作业质量的智能监管。监管范围包括截割期间入侵，作业区域人员摔倒，截割结束后截割部落地到位、敲帮问顶、临时支护、顶板支护等作业工序作业是否规范。并实时生产告警工单，通知掘进队长、安全员、调度室值班人员和安全矿长，及时进行问题处置和事后问题回溯、学习、改进。



图 6-8 掘进安全质量智能监管系统

如图 6-8 所示，掘进安全质量智能监管实现了人工智能与掘进工艺、技术、装备的深度融合。通过人工监管到自动监测的转变，保障人员安全，提升作业规范，保证支护质量，降低巷道维护成本，提高煤矿生产的安全系数。

(3) 济宁二号煤矿：重介选煤分选密度智能控制

重介选煤分选工艺复杂，选煤厂自动化和智能化水平普遍较低，当前主要依靠人工经验控制生产。由于人工经验参差不齐，经常出现精煤产品质量过剩的问题，造成精煤回收损失，影响选煤经济效益。

利用盘古矿山大模型在参数预测方面的优势，对重介分选生产数据的进行分析建模，构建出重介选煤分选密度预测模型。通过对实时数据分析处理，实现基模型自动调参优化和高效协同，预测出最优的分选密度，通过联动重介旋流控制系统完成精确控制，如图 6-9 所示。



图 6-9 重介选煤分选密度智能控制系统

重介选煤分选密度的智能预测实现了重介选煤由粗放式生产到精细化生产的转变，稳定了精煤灰分波动率，提高了精煤产率，提升了企业的经济收益。经过生产验证，精煤产率提升 0.2% 以上，按照年入洗量 400 万吨计算，精煤增产约 0.8 万吨，参考该矿精煤产品和中煤产品当前的市场差价 600 元 / 吨，粗略测算年增收约 480 万元人民币。

6.1.2.3 方案价值

人工智能新架构和盘古矿山大模型协同的矿山人工智能方案在山能集团商用落地，它给山能集团带来“四个转变”：

实现由“高危生产”向“本质安全”的转变。通过对人的不安全因素、物的不安全状态进行智能管控、提醒、消缺，进一步提升安全自管自治能力，降低安全事故发生概率。

实现由“劳动密集”向“精简高效”转变。基于大模型的视觉 AI 分析技术，优化日常巡检和安全监管工作，进一步提升生产效率，降低

劳动强度，减少作业人员。

实现由“粗放管理”向“质量效益”转变。通过对各厂矿生产数据建模，与人工智能相结合，进行生产系统参数预测和优化，提升产品质量，降低生产成本，有效提高生产效率和经济效益。

实现由“分散重复”向“集约高效”转变。通过对人工智能应用的统建统管，实现了集约化的人工智能算法开发、模型训练和迭代升级，业务场景创新规模负责、系统应用和人工智能人才培养，避免了各单位重复投资和烟囱式建设。

6.1.3 持续运营

人工智能的成功应用是一个完整的系统工程，不仅要搭建技术平台，还要建立全新的运营体系，通过产品运营、运营支撑和客户运营环节的不断优化，确保人工智能技术和产品能够更快地实现规模化复制。在此过程中，山能集团、云鼎科技股份有限公司（以下简称云鼎）和华为携手合作，在各自领域发挥优势，共同推动

人工智能的持续落地。

云鼎作为山能集团的产业化公司，主要负责场景化 AI 应用开发、交付和运营工作，其中山能集团主要提供应用场景、业务专家和政策保障方面的支持，华为作为山能集团的合作伙伴赋能云鼎，主要在场景化 AI 应用开发、维护和优化能力方面给予全面支持。

1) 赋能培养人工智能人才：华为提供标准化培训课程，赋能云鼎研发团队，具备基于大模型平台的场景化模型开发能力、平台运维运营能力，帮助云鼎、山能集团、矿企培养人工智能人才。

2) 结对帮扶支撑云鼎组织建设：云鼎整合组织，华为提供结对帮扶服务，围绕人工智能应用 / 项目的全生命周期为云鼎提供 E2E 的专家技术支持。

3) 流程支撑纵向做深、横向扩展：三方共同优化完善人工智能场景需求规划流程，覆盖矿企需求收集、需求评审、立项、解决方案设计、版本开发、交付运维等环节，支撑云鼎完成煤矿产业板块价值场景全面覆盖，后续向电力、新能源等板块横向扩展。

4) 运营驱动产品迭代演进：云鼎开发的场景化应用上线后，基于大模型平台的边用边学能力和持续运营流程，矿企持续收集反馈负样本，云鼎运营团队例行分析误报、漏报的反馈，统计模型指标，及时识别问题并优化模型，确保模型越用越好

5) 能力外溢：基于云鼎当前已具备的场景能力和持续规划，在已完成与华为产品的集成验证基础上，打造一套即适用于集团内部复制推广，

也能用于向集团外溢出能力的解决方案组合，通过提供应用，AI 算法模型和服务，助力煤矿行业提升智能化建设水平。

6.1.4 总结与展望

山东能源集团与华为此次基于盘古矿山大模型的联合创新，是人工智能技术与行业业务场景深度融合一次成功的探索实践，场景应用系统为盘古矿山大模型训练提供了多样化的模型数据，持续提升大模型的能力，而大模型训练出来的场景化 AI 模型又应用在煤矿安全生产场景中，不断提升煤矿智能化和安全水平。

通过本次实践，我们也深刻的认识到盘古矿山大模型不仅是一个产品、一个平台，更是一套系统工程，它以云边协同架构、边用边学和非正常即异常机制为基础，以持续运营机制作为保障，通过技术赋能，组织整合，使人工智能从理论走向实践，从联合创新走向实际应用，从研发中心走向了生产现场。

面向未来，山东能源集团将持续深化人工智能技术的创新应用，最大限度减少危险岗位作业人数，建成一批新型智能化示范矿井，坚持合作共享共赢，为提高煤矿行业智能化建设水平，助力国家经济社会发展做出行业贡献，展现担当。

一、深耕矿山智能应用，提升生产智能化水平。
基于盘古矿山大模型在企业的落地，进一步深化人工智能技术，加强人工智能技术业务场景的深度融合，训练开发更多的应用场景，提高煤矿安全生产保障水平和企业效益。

二、辐射核心产业板块，推动智能化建设横向拓展。立足当前矿业板块成熟实践，将盘古矿

山大模型能力辐射到企业高端化工、电力、新能源新材料、高端装备制造、现代物流贸易等业务板块，构建全产业链智能生态，持续推动“数智山能”建设。

三、坚持合作共享共赢，赋能行业高质量发展。加强行业内外智能化应用交流合作，分享智能化应用经验，联合更多业内同行和产业合作伙伴，携手创新，构建开放共享的创新生态圈，促进人工智能技术与矿山智能化建设的深度融合，推动行业实现安全、绿色、智能、高效发展。

▶ 6.2 华菱湘钢

6.2.1 案例概述

华菱湘钢始建于 1958 年，产品涵盖宽厚板、线材和棒材三大类 400 多个品种，具备年产钢 1600 万吨的综合生产能力。目前已经发展成为了湖南单体规模最大的国有企业，也是中国南方重要的精品钢材制造基地。2021 年以湘钢为主体的湖南钢铁集团取得了中国钢铁上市公司营收和利润排名双第 2 名的成绩，成为湖南首家世界 500 强企业。

2019 年 8 月，湘钢与华为公司、湖南移动携手合作，共同开启了湘钢“5G+智慧工厂”建设。历经 3 年，成功打造了“黑灯车间”、无人产线 3 条、建设 5G+远控天车、5G+AR 远程装配、5G+巡检无人机群等应用场景 21 个，成为中国钢铁行业 5G 全连接工厂的标杆案例。

2022 年 11 月，湘钢与华为公司、湖南移动进一步签订《钢铁人工智能项目合作框架协议》，在全厂 5G 高速互联的基础上，基于盘古大模型 + 云边协同架构，建设覆盖钢铁制造全流程的智能平台和标准体系，将湘钢建设成为钢铁行业“信息化、数字化、智能化”的示范标杆，引领钢铁行业从“制造”向“智造”转变。三方联合团队深入现场，结合生产关键问题和云

边协同架构，快速在焦化、炼铁、炼钢和轧制环节联合打造了行车智能调度、智慧配煤、转炉自动识别控制、皮带智能监测等场景应用，并在试点孵化智能精炼、能耗预测等场景应用以及在集团内的推广复制。

6.2.2 解决方案和价值

6.2.2.1 方案架构

依托华菱湘钢与华为技术有限公司在智能制造上的长期探索与实践，双方创新性采用了人工智能新架构和盘古行业大模型协同的工业化人工智能解决方案，实现了 AI 大模型在钢铁行业的快速构建和部署，缩短了 AI 建设周期，加速了湘钢 AI 应用的落地。

(1) 华为云和湘钢数智云两级协同架构，实现数据采集与算法训练的高效协同。

在华为云上建设基于盘古大模型的人工智能训练中心、开发中心和运维中心，实现湘钢智能化应用的快速训练开发和部署应用。盘古大模型使用湘钢数智边缘云采集的数据进行算法开发、模型训练，训练好的 AI 模型通过线上下发到湘钢边缘数智边缘云部署，湘钢边缘云运

行 AI 推理模型。结合非正常即异常、边用边学机制实现云边高效协同，越用越准。

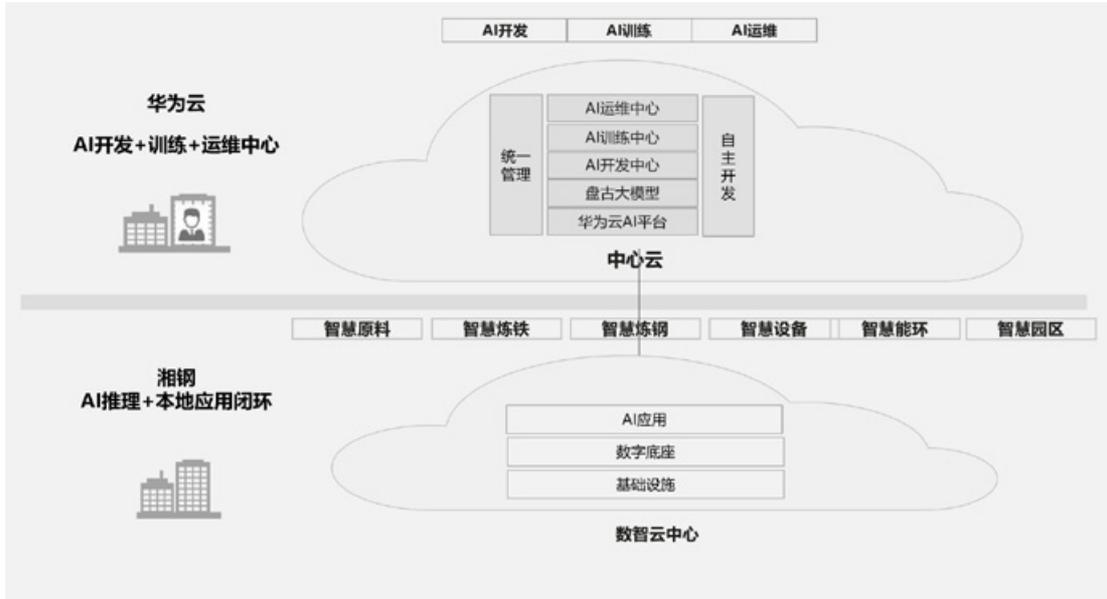


图 6-10 湘钢人工智能部署架构

盘古大模型采用大量预训练模型，实现湘钢 AI 应用的快速落地。盘古行业大模型通过海量数据训练，学习到大量工业领域的特征和经验，将其沉淀在大模型上亿的参数中。当针对钢铁行业某一具体场景开发时，由于大模型已经学习过类似场景的相关特征，而非从 0 开始，因此只需少量样本即可达到较高的识别精度，相比传统模型识别精度提升 10% 以上。智慧焦化配煤、智能皮带监测等应用场景在进行少量样本训练进行模型微调后，模型精度即可满足业务使用。

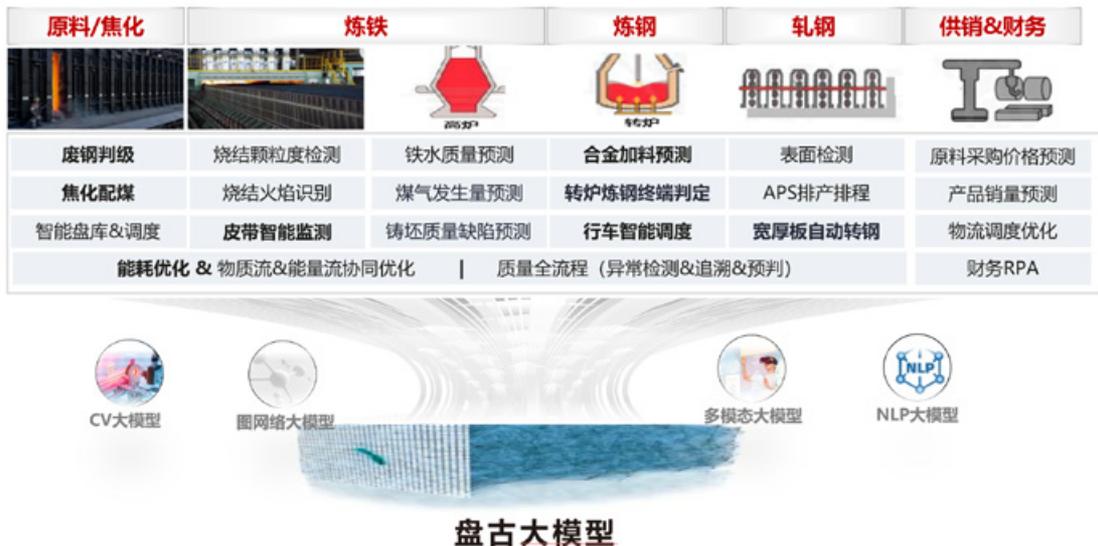


图 6-11 盘古大模型钢铁行业应用场景

6.2.2.2 应用场景

华菱湘钢在焦化、炼铁、炼钢和轧制环节进行了行车智能调度、智慧配煤、转钢自动识别控制等应用场景进行了部署，效果显著，同时试点孵化智能精炼、能耗预测等新场景应用。

一、智能行车调度

行车集装卸、搬运、运输功能于一身，是钢铁生产最主要的设备之一，也是决定钢厂高效运转的关键因素。湘钢炼钢厂转炉炉后精炼跨是当前行车作业最繁忙的区域，由于转炉出钢 & 炉后精炼 & 连铸钢水上台接收均在一个跨内，所有钢水包的吊运，精炼跨区域的大型物料设备吊运由五台冶金吊行车进行吊运。五台行车的调度由人工通过视频监控各工位情况，基于计划进行调度决策，通过对讲机、座机等指挥各岗位工作人员作业，实现各个环节的串联调

度。人工调度严重依赖调度员的经验和能力，由于生产过程高度复杂，人工无法全面感知各个环节中的情况，调度决策难以全面考虑后续各个工艺环节，会造成工艺环节衔接不顺导致的生产效率下降。人工调度制约了企业的生产管理水平和制造自动化水平的提高，从而影响了生产节奏、设备效率、生产成本等生产指标的改善提高。

华菱湘钢炼钢厂上线的行车智能调度系统，基于 AI 大模型的决策优化技术，综合考虑炼钢各环节生产计划、各冶炼工位生产实绩、行车和钢包实时位置信息、行车的检修和设备异常、各类业务规则等情况，整体统筹安排行车运行规划，智能生成行车调度计划，对于生产的动态变化，可以在 1 分钟内完成未来 30 分钟的决策并完成指令下发，有效提升钢包周转率，减少每炉次等待时间，降低过程温降，提升行车运行效率，减少炼钢环节的吨钢成本。

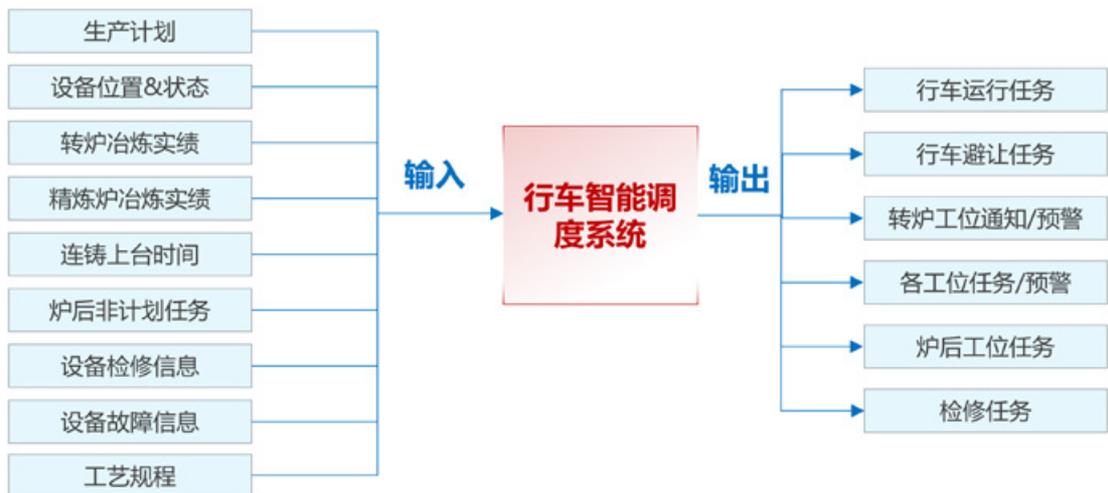


图 6-12 行车智能调度系统输入输出

二、转钢自动识别控制

钢厂在炼钢过程中，有一项主要的工序是成形轧制工序，也就是沿板坯长度方向或宽度方向进行 4 ~ 7 道轧制，把坯料轧至所要求的厚度。每块钢坯的轧制过程会涉及 2~4 次不等的转钢操作。转钢操作是通过控制转钢辊道上的锥形辊道的转向和转速，带动钢坯旋转 90 度之后再送入轧机，从

而达到在不同方向上展宽的工艺要求。这道工序是原来完全由人工通过监测画面，手动控制手柄控制操作完成。操作岗位需要工人 24 小时轮班值守操作，转钢操作的速度和效率依赖操作工人的技术熟练程度。



图 6-13 宽厚板粗轧转钢监测画面 & 人工转钢操作

基于 AI 的自动转钢应用，能够通过 AI 摄像头，自动采集转钢辊道区域和钢坯的视频数据，经过云端训练模型之后，可以将模型部署在昇腾服务器上。通过部署在计算节点上的应用和算法实现自动纪录钢坯、识别钢坯的方向和角度，并根据视觉识别输入结果对接轧机 PLC 控制系统，自动控制转钢辊道的转向和转速，实现板坯的 90 度自动旋转，从而节省人力，提高钢坯的轧制效率。

AI 转钢自动识别应用系统主要由 AI 摄像头、边缘计算 & 推理节点、视觉识别 & 角度计算系统、云端 AI 训练平台和轧钢 PLC 控制系统组成。通过计算服务器和推理节点支撑转钢动作、角度精确计算，经过实时计算的角度反馈对接轧机控制系统完成转钢操作。大模型 + 智能 + 云的结合实现了对企业实际作业自动化和效率的提升，也启发了更多计算机视觉的应用场景。



图 6-14 自动转钢原理 (左图) & 实时纪录角度计算 (右图)

三、智慧焦化配煤

配煤的决策需要考虑多种原料煤（质量、价格、库存）、生产工况（炉型、炉温、结焦时间等）以及焦炭质量要求等多维度因素，可选的求解空间非常大，当前焦化行业配煤主要依赖技术工程师的

人工经验，难以兼顾配煤成本与焦炭质量，很难做到精准计算。而 AI 智能可以很好地应对多参数的寻优问题，综合考虑多方因素给出多个较优的配煤决策方案。

华菱湘钢部署的智慧焦化配煤系统，依托 AI 大模型深度挖掘原料煤之间的配伍性和特征相关性，将配煤大数据与机理结合，并融入专家经验，结合业界先进配煤理论科学配煤。借助天筹求解器提升决策优化能力，输出一个配比仅需要约 30 秒，同时通过优化配比，平均每吨配合煤成本可降低 5 元 / 吨，实现降本增效目的。



图 6-15 AI 智慧配煤原理示意

6.2.2.3 方案价值

华为云和湘钢数智云两级架构和盘古大模型协同的人工智能方案，结合钢铁冶炼行业场景的不断创新，实现了人工智能在湘钢的高效开发与快速部署，加速了湘钢集团的智能化升级步伐。

实现由“人工调度”向“智能调度”转变。通过对生产计划、设备状态、钢包位置和业务规则等数学建模，基于大模型的天筹求解器，对生产动态调度进行全局优化决策，进一步提升生产效率，降低吨钢能耗。

实现由“人工操作”向“无人操控”转变。基于大模型的视觉 AI 分析技术，自动精准识别钢坯位置和角度，联动设备控制，实现无人操控。

实现由“专家经验”向“智能决策”转变。通过对多种原料煤、设备工况、产品质量等进行数学建模，集合行业机理，进行多参数寻优，平衡生产质量与成本，提供持续优化的生产参数，有效实现降本增效。

6.2.3 总结与展望

从 2016 年开始，湘钢开始推进“信息化、数字化、智能化”，建设智慧湘钢。湘钢从 2019 年开始推进 5G 智慧工厂建设，2021 年构建湘钢数智云，2022 年人工智能战略推进，落地实践了 21 个智能化场景，建设了 3 条智慧产线，在降本增效上取得了良好效果。人均吨钢从 2017 年的 700 吨 / 年，快速增加到 2022 年的 1600 吨 / 年，增长了 125%；吨钢能耗下降超过 10%。

展望未来，华菱湘钢常务副总经理喻维纲表示，当前生产过程的数字化、智能化转型发展主要集中在生产工艺优化、能耗降低、技术创新等方面，部分嵌入流程型智能制造、大规模个性化定制、远程运维服务等领域，总体而言对整个钢铁生产、经营、管理的全流程渗透率仍然较低，数智化转型的广度和深度都有待拓展。在今天所有巨大的不确定中，只有数字化和智能化是确定的。以前，数智化转型还是企业的选择题；现在，早已成了企业生存的必答题。未来 10 年，是传统行业推进数智化的关键 10 年。如果到了现在还不准备变革，相信最后一定会成为数字脱贫的对象。因为高效的商业模式和低效的商业模式之间，不存在竞争，只会“逐步取代”。湘钢将充分利用 AI 大模型、大数据、云计算等 ICT 技术进行全流程、全业务的数字化融合，打造钢铁行业的智能工业互联网平台。一方面推动湘钢的高质量发展，另一方面也推动钢铁行业向智慧钢铁、绿色钢铁方向发展。湘钢将持续加大与华为等伙伴的合作力度，打造智慧钢铁垂直行业应用标杆。确保公司在数字化智能化转型中成为行业的领跑者。





第七章 金融

▶ 7.1 上海浦东发展银行股份有限公司

7.1.1 案例综述

上海浦东发展银行股份有限公司（以下简称：浦发银行）是1992年8月28日经中国人民银行批准设立、1993年1月9日开业的全国性股份制商业银行，总行设在上海。目前，注册资本金293.52亿元，至2022年，公司总资产规模达8.7万亿元。

2020年8月，浦发银行董事长、行长与华为轮值董事长胡厚崑签署全面战略合作协议，2020年9月24日双方在华为全联接大会联合发布《“物的银行”白皮书》，这是业内首次提出面向“智能物”的全新金融服务模式和设计体系。在此基础上，浦发银行以科技赋能，推进产业数字金融创新，进一步丰富“物的银行”内涵，推进产业数字金融创新。

7.1.2 解决方案和价值

7.1.2.1 方案架构

浦慧云仓通过在仓库端布设的 RFID，称重和区块链 PDA 等设备，记录客观信息，货物进仓与融资放款匹配、还款与货品出库联动，实现有效控货，方便监管。通过贸易信息、感知信息和市场信息的融合校验，厘清权属，精准估值，高效处置，实现物流、商流、资金流、信息流的“四流合一”，确保动产获得应有的估值，让产业链上下游企业享受应有的金融服务，让金融服务覆盖到更多产业供应链场景。

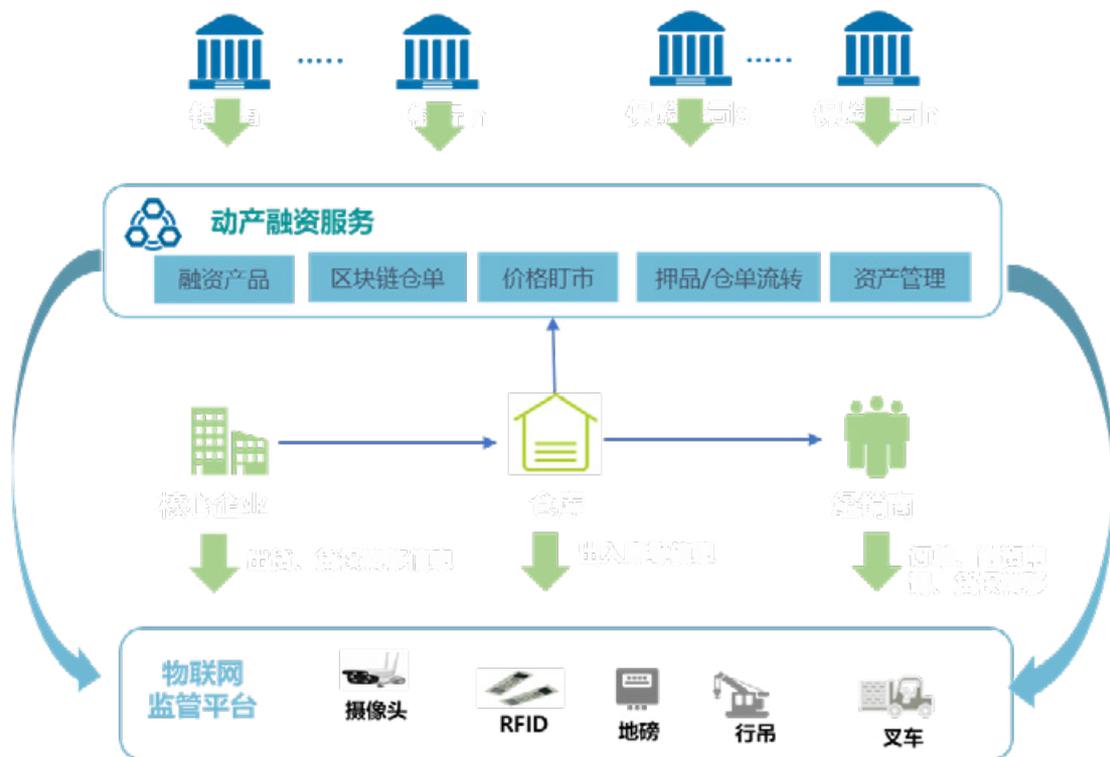


图 7-1 动产融资金融仓方案架构

1. 智能物联网升级，打造可信资产池

- 在物流仓储数字化的趋势下，双方通过前端部署智能物联网设备，引入 AI 能力和多维度数据的校验，将普通仓库转变为“金融仓”，将金融服务的触角延伸到仓库的现场管理与业务流程中。实现贷前确权共识互信、资产评估真实客观，资产数量多方校验；贷款管理高效、易操作，全天候资产监测，随时可查，出入库管理风险管控，实时预警。

- 依托华为技术底座，浦慧云仓让每个物品可被识别，让每条告警可被纪录，让每件货物可被信任：
- 基于盘古 CV 大模型，具有很强的泛化能力，可以实现模型的高度重用，1 个模型涵盖 9 个物流场景，监测收货，入库，在库，出库的全过程。通过融合多个摄像头数据，利用大模型的时空匹配能力，实现对目标轨道的精确估计。
- 通过分布式 RFID，实现出入库实时感知，实时纪录。
- 通过华为 IoT 平台，可以实现图形化建模，灵活定义、构建模型和统一管理，通过数字孪生技术对数据交叉验证，对货物资产、仓储作业过程、人员和环境进行全面的风险分析。

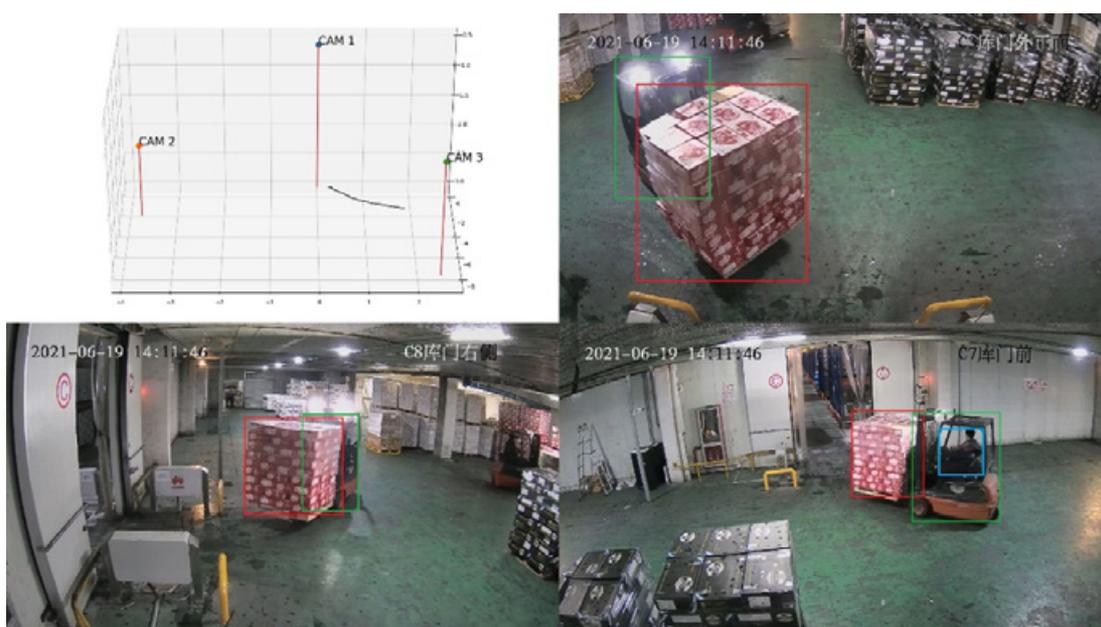


图 7-2 金融仓叉车出入库轨道及货物状态智能分析

2. 浦慧云仓监测中心，确保货物入库真实，保障动产投资服务

在浦慧云仓监测中心，操作人员可通过盘古大模型进行货堆计数，并且与前端入库的扫码和重量数据及贸易背景数据进行校验，确保货物入库的真实性。

银行通过部署智能运营平台，可视化透明监管，建成全数据、全系统、全网络的一体化指挥中心，实现跨部门、跨层级一体化指挥，实现出入库管理风险管控，实时预警。实现贷前确权共识互信、资产评估真实客观，资产数量多方校验；贷款管理高效、易操作，全天候资产监测，随时可查，出入库管理风险管控，实时预警。



图 7-3 浦慧云仓智能动产监管平台

7.1.2.2 应用场景

传统动产质押业务中有很多痛点：因货权不清晰，导致“多次质押”；因仓储监管公司道德风险，导致“黄金变铜”；因缺乏数字化手段，导致过程监管成本高，流于形式；因缺乏处置平台，导致货物变现难。通过智能化手段帮助传统金融仓提升风控管理能力，盘活百万亿动产，支撑中小企业融资需求。

7.1.2.3 方案价值

通过华为产业金融动产融资金融仓解决方案，华为助力浦发银行打造了“物的银行”首个应用——“浦慧云仓”，在业界率先推出线上信贷产品“优浦 e 仓贷”，单仓可服务 300+ 小微企业，支撑 3 亿元 / 年融资额。从原来线下异地审批周期 15 天 +，变成线上申请 / 审批 / 签约，贷款资金实时到账；线上还款、实时解质押；支持“整进零出”。风控从原来监管员人工跑仓库抽查，现在变成交易 + 实物的 360 度风控，

7*24 小时资产监测及实时预警。

利用盘古大模型小样本学习和泛化的能力，突破复杂三维场景下小目标密集检测这一 AI 算法领域技术难题。通过技术赋能产业场景，打造“端边云 + 智能”协同方案，增强了金融仓的三维可信。

7.1.3 总结和展望

浦慧云仓，作为浦发银行产业数字金融落地的一个试点，To B 企业服务市场的特殊需求，驱使银行思考并变革业务逻辑、重构技术能力，运用 AI 能力赋能场景，持续优化风险管理机制，提高风控水平与权重，增加“交易信用”和“物的信用”的权重，进一步发挥数据的价值。通过数字技术和产业的深度融合，以数字化能力切入各个实体经济产业的供应链场景，更好地实现金融赋能实体经济愿景。

▶ 7.2 广发证券股份有限公司

7.2.1 案例综述

广发证券股份有限公司成立于 1991 年，是中国资本市场最具影响力和创新能力的综合性证券公司之一。该公司以投行业务和经纪业务为特色，曾长期占据行业前列。然而，由于 2019 年保荐的康美药业爆出财务造假事件，广发证券受到监管处罚，其投行业务遭受严重打击。为提升对所服务的企业的财务报表的分析和预警能力，广发证券于 2020 年启动了“企业财务智能预警”的项目

华为与广发证券有着紧密的长期合作关系，提供了涵盖云计算、大数据和人工智能等多个领域的全方位 ICT 解决方案。针对广发证券在项目中面临的挑战，华为参与了项目的联合攻关，全力支持广发证券的项目成功。

“企业财务智能预警”项目主要聚焦于解决广发证券投行业务的核心痛点，即提升财务报表审核的合规性和风险管控能力。通过构建基于 AI 技术的财务报表智能解析和风险预警系统，期望实现对企业财务数据的自动化审计和异常识别。这将不仅大大提高财务风险识别的效率，还能够在识别潜在的造假风险和舞弊动机方面发挥重要作用。



7.2.2 解决方案和价值

7.2.2.1 方案架构

广发证券原先依赖人工和基础机器学习模型进行财务报表分析，这些方法耗时且准确性有限。为应对这一挑战，广发证券与华为合作，构建基于盘古大模型的企业财务智能预警平台。该平台融合了专家经验和先进的机器学习算法，并利用人工智能的最前沿技术——超大规模预训练大模型来实现预警准确率的提升。



图 7-6 智能财务预警系统架构图

架构图中突出的 AI 部分包括盘古大模型首创的图网络融合技术。这项技术允许多个模型并行学习各自特定的数据特征，并从中选择最优模型组合。这种创新性的模型比赛机制显著提高了财务异常识别的准确度，达到了 90%，并将财务异常识别时间缩短到分钟级别，大幅提升了风险识别的频度和及时性。

7.2.2.2 应用场景

广发证券的企业财务智能预警系统经过基于华为云盘古大模型的全面升级，现已广泛应用于多个业务流程和用户场景。该系统覆盖了 4000 多家上市公司和 6000 多家发债企业，能够细致地揭示 4 大类财务异常、160 多小类预警信号、6 大舞弊动机和 10 余种常见舞

弊手段。系统支持多条线用户，包括但不限于投资银行、风控部门，能够灵活应用于项目审核、尽职调查、业务督导等多个业务流程。以 2019 年的数据为例，该系统能够预测出被监管处罚、问询或 ST 的企业共有 496 家，其中成功预测出 439 家，覆盖度接近 90%。在监管处罚的企业中，盘古大模型无一漏网，展示了出色的准确性和广泛的应用场景。

7.2.2.3 方案价值

广发证券企业财务智能预警系统不仅在业内设置了新的高标准，还获得了多个重要奖项和认可。该系统成功将模型精度从最初的 79% 提升到近 90%，大幅增强了企业财务风险的识别精度。这一突破性成果已在广州市地方金融监

督管理局主办的 2021 年首届金羊“点数成金”数字金融创新案例示范活动中获得认可，并被评选为“金融科技创新应用示范案例”。更值得一提的是，该平台在 2023 年 4 月荣获了中国人民银行“金融科技发展奖”二等奖，凸显了企业财务智能预警系统在提高行业合规性和风险管理效能方面的巨大价值。

7.2.3 总结和展望

广发证券的企业财务智能预警系统是金融科技与传统证券业务融合的典型示例，展示了大数据和人工智能如何赋能传统金融机构，以更高的效率和准确性应对日益复杂的市场环境。经过与华为的深度合作和技术升级，该系统已达到行业领先水平，不仅在提高财务风险识别准确率方面取得了突破性进展，而且在金融科技领域获得了多个重要奖项和广泛的认可，为证券行业提升风险管理水平提供了参照标杆。

该系统成功的经验为金融行业提供了宝贵的实践参考。展望未来，随着资本市场的不断发展和监管环境的日趋严格，企业财务风险管理将面临更多的挑战。广发证券企业财务智能预警系统有望继续拓展其应用场景，将其运用到更多的业务场景中，以实现更全面的风险控制。





第八章 电力

▶ 8.1 国网江苏省电力有限公司

8.1.1 案例概述

国网江苏省电力有限公司（以下简称“国网江苏电力”）是国家电网有限公司系统规模最大的省级电网公司之一。江苏是用能大省，却是资源小省，一次性能源较为缺乏，主要从省外引入电力能源破解华东地区资源约束瓶颈。国网江苏电力拥有 35 千伏及以上变电站 3 千多座、输电线路近 11 万公里，初步形成以“一交四直”特高压混联电网为骨干网架、各级电网协调发展的坚强智能电网。为了保障外送电力能源的接受能力，国网江苏电力牢固树立“一切事故皆可预防”理念，始终将大电网安全放在首位。

为切实保障能源“大动脉”安全稳定运行，国网江苏电力积极应用无人机自主巡检、可视化远程监测等智能巡检手段，开展重要线路巡检与隐患排查，重点关注输电线路基础、金具、导地线弧垂和山火隐患等，确保线路安全运行。

国网江苏省电力有限公司已配置 135 座固定机场，配置输电线路巡检无人机近 2800 架，平均每百公里配置 2.5 架，产生输电机巡图片约 100 万张 / 月，已实现重点输电线路无人机自主巡检全覆盖。

国网江苏电力输电线路巡检业务，当前主要通过“杆塔可视化远程监测 + 无人机自主巡检”两种手段，其中杆塔可视化远程监测主要应用于三跨点（铁路、公路、河流），禁飞区（机场等）等，采用本地采集，远程后台监测的方式。无人机自主巡检则是通过巡检班组现场进行无人机拍照，后台通过人工智能识别算法对无人机拍摄图像进行分析方式。

近年来，国网江苏电力与华为技术有限公司，积极探讨 AI 智能在电网的实践，双方共同推进在江苏省进行试点应用。2022 年开始，相继在泰州等地开展了输电无人机巡检大模型试点。



图 8-1 国网江苏省电力项目 AI 智能合作进展

8.1.2 解决方案和价值

8.1.2.1 方案架构

针对输电线路巡检业务和智能化技术发展所面临的问题和困境，双方携手推出 AI 平台 + 大模型联合解决方案。基于云边协同架构，华为人工智能平台 ModelArts 和盘古电力行业大规模预训练模型，开展输电巡检图像智能识别工作，充分发挥盘古电力大模型在无人机巡检中的价值作用，提升输电线路巡检业务智能化水平，探索人工智能在电力行业的创新应用。

云边协同架构

基于国网江苏电力构建的多云部署能力，创新采用省电力公司中心云和地市边缘(变电站、配电房等)，云边协同两级架构。省公司中心云建设基于盘古电力大模型的 AI 开发中心、训练中心和运维中心，地市公司建设边缘基础设施。通过省公司统投统管，提供统一集约建设和服务，避免各地市公司烟囱式重复建设。

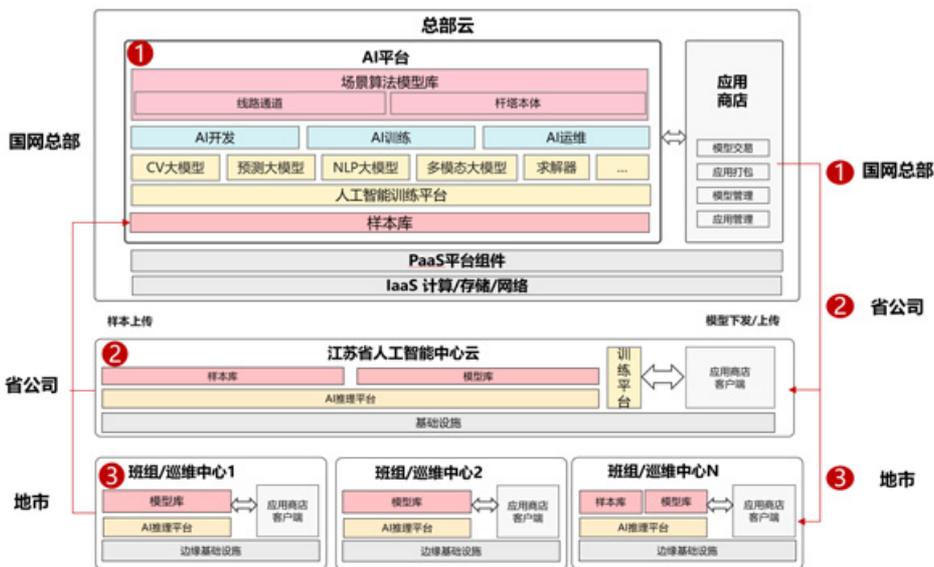


图 8-2 国网江苏省电力人工智能平台云边协同架构

部署于省公司中心云的盘古电力大模型，使用各地市边缘所采集的巡检图片构建的样本库，进行算法开发和训练，训练好的 AI 业务模型统一下发边缘站点，进行远程部署，并基于标准化 AI 框架平台，与应用系统进行集成。模型的推理信息与应用系统、生产系统实现联动完成业务闭环。

盘古电力大模型，实现 AI 开发由“作坊式”向“工厂式”转变

大模型即大规模预训练模型的简称。预训练模型，即首先在一个原始任务上预先训练一个初始模型，然后在下游任务上继续对该模型进行精调，从而达到提高下游任务准确率。而大规模预训练模型是模型参数发展到千亿、万亿级别的预训练模型，这种大规模预训练模型的泛化能力变强，能够沉淀行业经验，具有更高效准确地从文本 / 图片中获取信息的能力。

盘古大模型团队设计了具有超过 30 亿参数的图像预训练模型——盘古电力 CV 大模型，相对于小模型参数量更多，网络更深、更宽，吸收海量的知识，提高模型的泛化能力，减少对领域数据标注的依赖。盘古大 CV 模型结合 ModelArts 工作流工程化能力，可解决传统 AI 作坊式开发模式下不能解决的 AI 规模化、产业化的难题。

低门槛 AI 开发，开发效率高：传统的无人机智能巡检 AI 模型开发面临缺陷种类多达上百种，需要数十个 AI 识别模型，开发成本高的问题。盘古电力大模型提供自动化 workflow，自动数据处理，自动化调参，自动生成模型，减少对 AI 开发工程师的专业依赖。同时，可以根据更新数据快速进行模型迭代，做到“边学边用”，开发效率高。培训周期短，较传统 AI 人员培养，

大模型开发培训从数月缩短至 2 周，工程师快速上岗。模型开发快，学员 2 周共计完成 6 个真实场景，平均 1 个场景 2~3 天完成，较传统开周期平均 30 天，缩短 90%。

样本标注更高效，小样本识别更精准：盘古能够从海量数据中，高效筛选缺陷样本，减少 80% 以上人力标注代价。同时，输电巡检存在部分缺陷类别数据量不足或是已标注数据不足的问题，盘古大模型天生具有更高的小样本学习能力，其独有的自动化数据增强，数据挖掘等算法在小样本上能够取得明显优势，能够更好应对缺陷样本中长尾分布，在一些场景下，甚至能够实现部分零训练样本下缺陷样本自动识别。此外，同等样本，大模型精准度更高；同样小本样情况下，大模型较传统小模型优化效率优势明显。持续增加样本后，精准度持续提升。

泛化性能高，实现模型由“作坊式”开发，向“工厂式”生产升级：大模型可以更深层挖掘数据背后的逻辑，同时经过海量数据预训练，使大模型拥有强大的拟合能力，其泛化性远远优于小模型。原来需要多个模型覆盖的视觉场景，现在大模型可以用一个模型覆盖多个场景，解决了模型碎片化问题。比如，针对输电线路巡检中的 9 大类缺陷（杆塔、绝缘子、大金具、小金具、导地线、基础、通道环境、附属设施、接地装置），传统基本采用多个小模型适配不同的缺陷。盘古大模型针对输电场景下的前四大类仅使用一个模型，即适配上百种小类别缺陷，解决了模型碎片化问题，同时也大大提升了算法开发和应用效率，降低了开发和交付成本，从而实现模型开发产线化，推动人工智能应用从“作坊式”到“工厂式”升级，可规模化复制。

8.1.2.2 应用场景

一、通道环境缺陷识别

输电通道入侵识别类型主要包括建（构）筑物、施工作业、山火、线路悬浮物等。这些缺陷与杆塔本体上的缺陷差异比较大，图像的背景也通常存在较大差异。

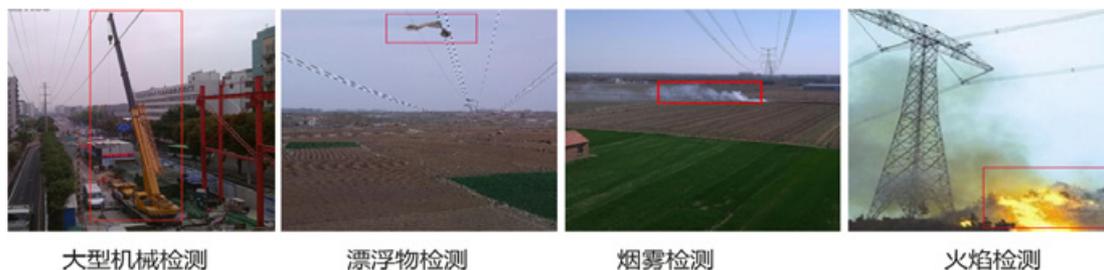


图 8-3 通道入侵识别示意图

实践证明，大模型借助其强大的持续学习能力、特征识别能力，在防外破及通道可视化场景性能方面更精准。基于大模型训练得到的模型，召回率、误报率优于在运模型。

在输电防外破场景，江苏盘古电力大模型仅用 1.8 万张图片训练，识别率超过传统小模型 50 万张样本训练。其中，大模型 1.8 万张图片样本库，召回率和误报率分别为 89.7% 和 18.4%。而在运小模型，在训练样本库为 50 万张图片样本库，发现率和误报率分别为 86.8% 和 16.7%。

在通道可视化场景，仅用两千余张缺陷样本库图片情况下，大模型发现率和误报率分别为 89.2% 和 41%，超过原定预期 85%。在增加样本规模持续训练后，效果持续提升，当训练样本缺陷图片规模达到 1.8 万张时，通道可视化场景，发现率和误报率分别为提升至 90.2% 和 18.7%。而在运小模型，在训练样本库为 50 万张图片时，召回率和误报率分别

为 89.2% 和 21%。

二、杆塔本体缺陷识别

在杆塔本体缺陷识别场景中，根据人工巡检经验，可能出现的缺陷类型主要包括塔身缺陷、部件缺陷、异物等。不同缺陷类型具有不同的特性，如“异物”和“塔身缺陷”只要发生，就说明图片中存在缺陷；而部件可能存在“正常”、“锈蚀”、“破损”、“丢失”等各种不同的状态，并且“正常”的部件和“有缺陷”的部件在外观上通常存在一定的相似性。因此，在识别“异物”和“塔身缺陷”时只需要判断“有”或者“没有”即可；而在识别部件缺陷时，不仅要判断部件“有”或者“没有”，还需要判断部件的状态是否正常。

实际采集到的数据中，不同故障类型的数据数量不同、难度不同，在设计识别算法时，需要考虑这些因素，针对不同的情况选择最合适的识别算法，保证最终的识别效果。



图 8-4 杆塔本体缺陷识别示意图

实践证明，同等小样本训练，大模型优化效率明显高于小模型。在训练样本库为 1000 张图片时，针对杆塔异物大模型召回率为 81.04%，误报率为 18.22%，而业界小模型召回率及误报率分别为 78.42% 和 59.2%。悬垂线夹船体偏移、防振锤锈蚀和绝缘子瓷质破损等场景大模型的召回率及误报率均优于业界小模型。持续增加样本，精准度持续提升，大样本情况下，召回率、误报率皆得到优化。

8.1.2.3 方案价值

国网江苏电力与华为在人工智能在电力行业的创新实践，坚定了国网“人工智能第一省”的战略投入。华为公司持续投入顶级算法专家，

江苏电力人工智能开发平台提供的强大算力支撑，算法专家和开发人员提供算法设计和优化方案，并结合输电巡检行业知识和数据，大大提升了盘古电力大模型训练算法针对输电巡检图像缺陷的 AI 识别准确率。针对输电防外破场景，在运模型召回率为 86.8%，大模型召回率达到 89.7%，比在运模型高 3 个百分点，基本做到了杆塔线路防外破场景故障的全识别。

盘古电力大模型在输电线路巡检中的创新应用，通过助力通道监拍设备和无人机巡检逐步减少人工巡检作业，提升了输电巡检工作效率和准确度，降低了人工巡检工作量，减少了输电巡检安全隐患，有力支撑了输电线路安全稳定运行。

8.1.3 总结与展望

国网江苏电力创新将大模型应用于无人机输电线路缺陷智能识别，是人工智能技术与电力业务场景深度融合的一次成功探索实践。无人机巡检和大模型等新技术在输电线路巡检中的广泛应用与成功实践，有力推动了输电线路巡检提效率、减成本和降风险，助力电网安全稳定运行。同时，盘古电力大模型在输电线路巡检的实践落地和成功经验，亦可推广至电力其他业务场景，加强人工智能技术与变电、配电、营销、安监、基建等多个业务场景的深度融合，训练开发更多的应用场景，提升电力安全生产保障水平和企业效益。

持续扩展到多模态大模型和 NLP 大模型，赋能电力行业高质量发展。当前聚焦电力 CV 大模型在输电领域的深化应用，未来通过对齐电力运维检修业务，扩展多模态大模型的应用落地，基于多种模态数据学习推理，助力设备故障健康性分析。基于客户营销客户业务需求，探索电力 NLP 大模型在电力知识问答，内容生成方面的应用，助力营销客服、设备知识积累等。通过各类大模型及人工智能技术与电力智能化建设的深度融合，推动电力行业实现安全、高效、绿色、可持续发展。





第九章 油气

▶ 9.1 中国石油集团

9.1.1 案例概述

中国石油天然气集团有限公司（简称中国石油集团）是一家集油气业务、油田技术服务、石油工程建设、石油装备制造、金融服务和新能源开发于一体的综合性国际能源公司，是中国油气行业占主导地位的最大的油气生产和销售商，是中国销售收入最大的公司之一，也是世界最大的石油公司之一。

中国石油集团抓住人工智能在石油行业刚刚应用的时机，在“十三五”期间规划设立了人工智能统一建设项目：认知计算平台（统建编号 E8）。项目于 2016 年开始可行性研究，全面开启了人工智能在勘探开发领域的应用实践。E8 是一个开放、可扩展的人工智能计算平台，按照数据、知识、算法、算力、场景五个关键因素进行设计，其目标是为油气勘探开发科研、生产、管理提供智能化分析手段，支撑油气勘探开发增储上产、降本增效。该项目由中国石油勘探开发研究院与华为公司联合建设，2021 年 7 月 1 日实现正式上线运行。认知计算平台基于不同的业务场景，通过 AI 的方式改进传统的工作模式。认知计算平台整体架构如图 9-1 所示。

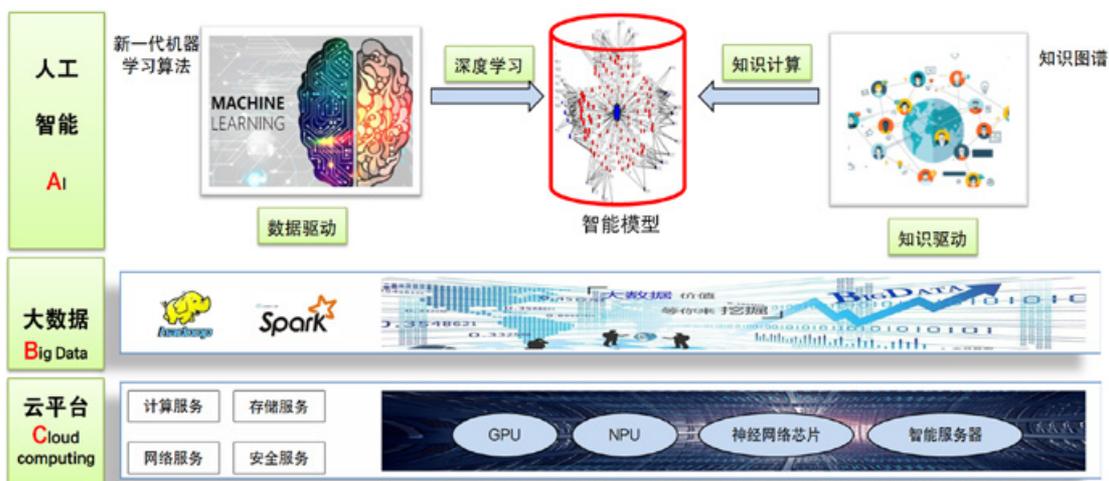


图 9-1 中国石油认知计算平台（E8）

项目建设目标包括：

- (1) 构建认知计算平台，实现数据管理、认知分析、业务应用、系统管理等功能；
- (2) 围绕稀油砂岩油藏梳理数据，构建并不断完善知识图谱库；
- (3) 建设勘探开发认知计算分析系统，支撑初至波自动拾取、地震层位解释、测井油气层识别、抽油机井工况诊断、单井产量递减和含水预测等业务应用需求。

随着平台功能的逐步完善，以及支撑业务范围的不断扩充，认知计算平台更名为人工智能平台，平台包括 AI 计算、知识图谱、油气 AI 社区、系统管理、业务应用 5 个一级功能模块，33 个二级功能模块。在认知计算平台试点项目基础上新增地震大模型、神经网络可视化建模、自动机器学习等 15 个二级功能模块，沿用数据管理、数据处理与标注、特征工程等 18 个二级功能模块。人工智能平台功能规划如图 9-2 所示。

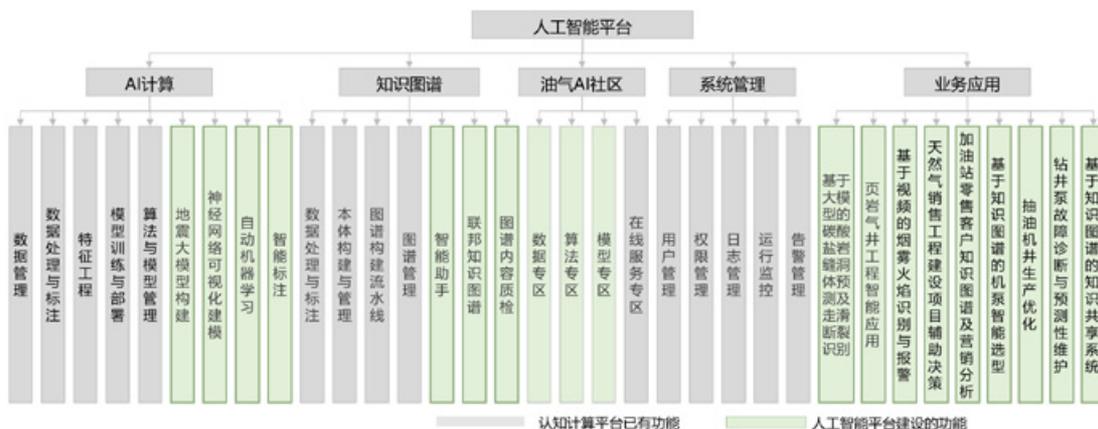


图 9-2 中国石油集团人工智能平台功能规划

中国石油集团人工智能平台采用云端训练 + 云端推理的模式，各二级单位可按需建设训练和推理分中心进行迭代优化，云边协同部署架构如图 9-3 所示。大模型的多级架构支持二级单位行业增量差异数据反馈、模型迭代、服务部署，实现 AI 模型的全生命周期管理，降低开发运维门槛。华为 AI 大模型方案助力中国石油集团打造属于自己的人工智能工厂，并以云边协同、边用边学的技术架构，满足各油田公司在各自生产经营领域的个性化需求。



图 9-3 云边协同部署架构

勘探开发研究院作为中国石油集团的二级单位，同时也作为国内科研创新排头兵，遵循集团人工智能建设发展体系，全面推进油田数字化转型与智能化发展。如下将主要以抽油机井工况诊断及测井油气层识别为例，介绍人工智能技术在油田勘探与生产现场的应用及价值。

9.1.2 解决方案和价值

9.1.2.1 行业挑战

现阶段油井工况的定量识别、趋势预测工作主要依赖于采油工程专家，专家通过分析各种参数的变化并结合物理模型来判断油井是否会出现故障、是否需要维护作业，这类方法完全依赖专家经验，做不到 24 小时实时识别，不能够及时发现潜在的油井故障，通常都是故障发生后才采取维护作业，往往造成延迟作业，增加停机时间，影响油井生产和油田整体效益。

油气田另外一个可以使用 AI 大幅提升效率的场景是测井油气层识别，虽然测井仪器的种类和测量信息在不断的丰富，但是传统的测井解释方法对多源测井信息相关性分析及储层评价和应用的研究不足，而且测井评价非常依赖专家经验，开发人员短时间内很难掌握，测井识别亟需进行智能化升级，并要立足解决油田勘探开发过程中面临的测井数据规模大、处理解释难度大、数据匹配性和平面覆盖性不足带来的问题。

9.1.2.2 解决方案

· 工况诊断方案：功图智能分析实时检测

随着油井的开采，任何人工举升方式下的生产环境、井筒条件和油藏条件都会发生变化，油井会出现如杆管磨损、油井出砂造成磨损加快、原油出蜡影响产量、液体中含 H₂S（硫化氢）腐蚀油管等状况，这些均能造成油井故障。

油井典型的示功图包括有气体影响、充不满、漏失、抽油杆断脱、出砂、结蜡、带喷井、活塞脱出工作筒、活塞下行碰泵、油稠等。典型示功图可作为人工智能对抽油机井泵况的参考依据，综合分析实测示功图，工况示例如图 9-4 所示。

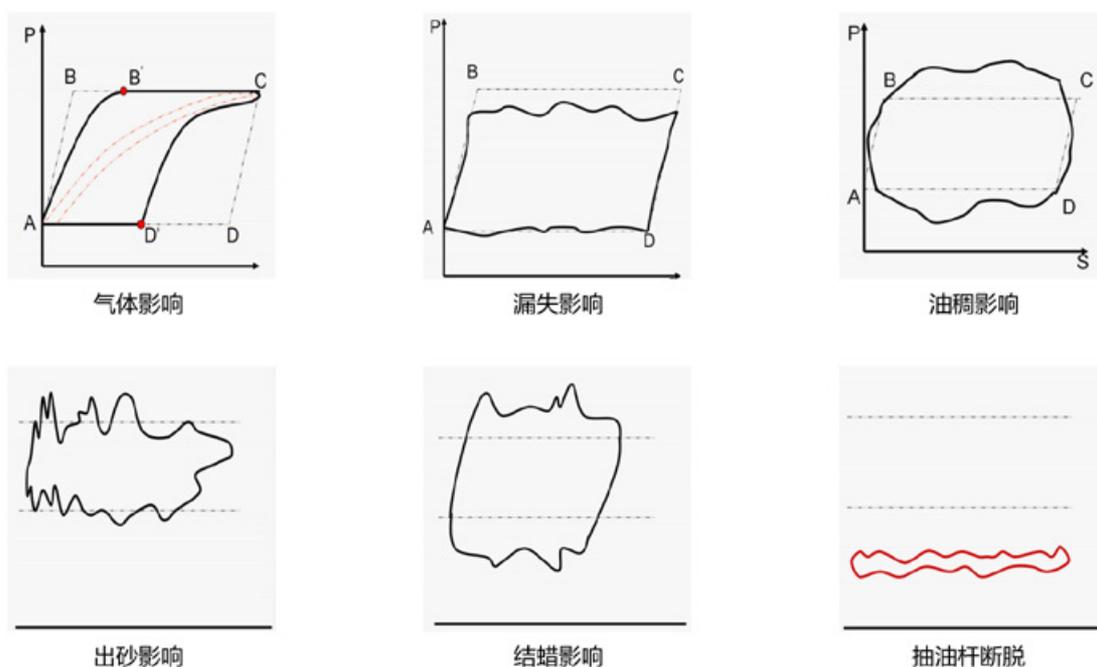


图 9-4 油井典型示功图

华为油气智慧井场方案中，在井场部署工况诊断应用，基于现场稳定高质量的功图数据进行智能分析，匹配典型功图进行工况诊断，当出现异常工况时，及时上报告警给采油工进行处理，提高井场智能化，减少现场作业的工作量，抽油机工况诊断方案流程图如图 9-5。

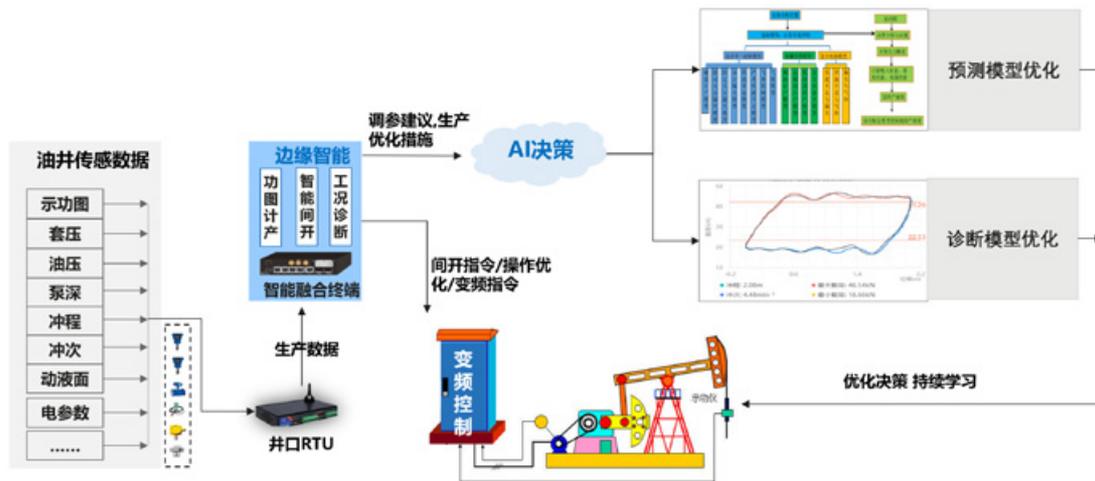


图 9-5 抽油机工况诊断业务流程图

井场部署支持边缘计算的智能融合终端，在融合终端中安装数据总线应用，用于接入和采集井场及井口数据。安装工况诊断智能分析应用，用于工况的实时监测、诊断及异常上报。在采油厂或作业区部署物联网管理平台，集中管理井场，查看各井场状态及告警，全局一览，实现作业区整体的生产管控。

· 测井油气层识别解决方案：油气水层智能识别

测井油气层解释是利用岩层的电化学特性、导电特性、声学特性、放射性等地球物理特性测量地球物理参数，从而求得地层的厚度、孔隙度、含油饱和度、渗透率，并划分出油、气、水层，给出地下油气藏分布范围和空间形态的重要手段，是油气上游业务中数据挖掘极为重要的信息来源。但测井油气层识别工作涉及的资料繁多、工作量大、解释周期长、所需的专业性与经验性强，使得油气层测井解释的效果

不太理想。

人工智能技术可以为测井油气层识别场景构建知识图谱，提高测井评价工作的效率与符合率，并且可以应用多维相关数据，提高解释精度，降低工作人员在测井油气层识别工作中的业务难度。

中石油人工智能平台对油气层识别场景进行了应用和实践，完成了区块优选、数据准备、分析处理和标定，确立了智能化油气层识别的相应流程，建立了油气层识别场景的知识图谱、实体数据库，建立了油气层人工智能认知的识别模型，完成了在油气层识别场景中的测试和智能应用，帮助测井解释专家和地质专家自动化地分析相关数据，识别潜力油层，并提供支撑解释结论的相关数据与证据。测井油气层场景化解决方案架构如图 9-6 所示。

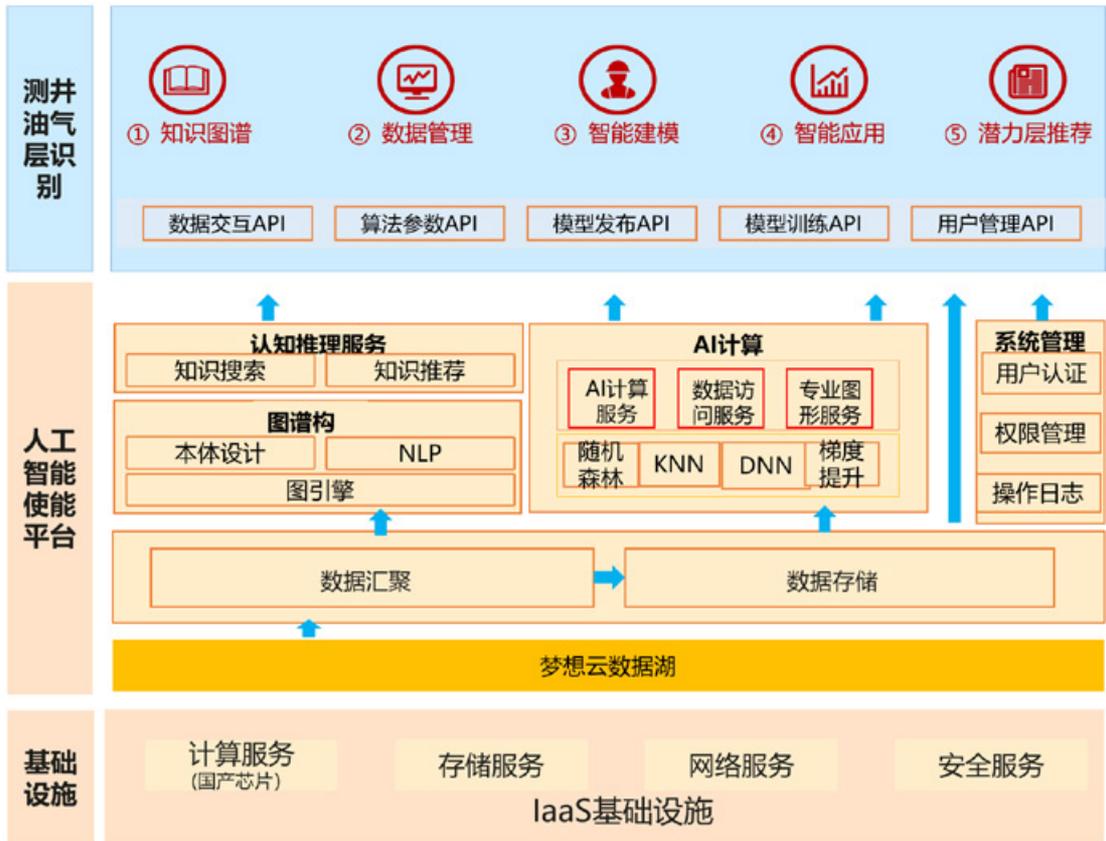


图 9-6 测井油气层场景化解决方案

方案通过建立智能参数预测模型和油气层识别模型，有效解决岩性复杂多样、非均质性强的储集层的参数计算和流体识别问题，并且预测准确率高，统一的知识图谱库，实现专家知识沉淀、专业知识融合以及知识挖掘和应用，大幅降低了测井的评价周期。

9.1.2.3 方案价值

工况诊断通过 AI 的加持，24 小时进行实时识别并提高了准确率，异常工况诊断准确率超过 90%，异常工况会及时告警通知作业人员上井实施措施，让工况处理从事后诊断升级为事前预警，大大减少停机时间，预计可提升机采效率 10%，减少作业维护费用 20%。工况诊断解决方案的推行加快了油田的智能化进程，助

力井场无人化、少人化的实现。

测井油气层识别通过人工智能在智能参数预测模型、油气层识别模型、知识图谱等方面的应用，在大港油田港东二区块的油气层识别中，准确率达到 90%，评价时间缩短了 70%，业务人员快速对相似区块进行人工智能识别，节约了项目资金，助力了对油井数据的深度挖掘，对油气藏分布范围和空间形态分析更加准确。

9.1.3 总结与展望

人工智能技术应用发展前景广阔，加速勘探开发产业业务与人工智能技术的深度融合发展，推动 AI 大模型的创新应用，利用机器学习、机器视觉、数据挖掘等算法提高专业软件的智

能化分析水平，促进人工智能在初至波自动拾取、地震层位解释、测井油气层识别和抽油机井工况诊断，储层参数预测等领域的成熟应用，能够大幅度提升研究效率、提升解释精度、提升预测能力，降低综合研究与管理成本。将物联网、机器视觉、深度学习等技术嵌入油田智能生产装备，促进工业机器人，无人机等装备成熟化并代替人工进行管道巡检、高危作业、物资投放、工程救援等工作，将大大降低生产成本，提高生产效率，保障生产安全。推动人工智能技术在传统油气行业的融合应用，是全面驱动油田勘探开发生产业务转型发展的最佳选择，对实现石油工业提质增效和创新发展具有非常重要的战略意义。

▶ 9.2 国家管网

9.2.1 案例概述

国家管网集团于 2019 年 12 月 9 日在北京正式成立，是国家深化油气体制改革和国资国企改革的重要举措，主要职责是负责全国油气干线管道、部分储气调峰设施的投资建设、管道联通、管道输送，并统一负责全国油气干线管网运行调度，有利于提高油气资源配置效率，促进油气行业高质量发展，促进管网互联互通，进而提高基础设施运行效率，降低终端用户用气成本，保障国家能源安全。

在 2021 年，基于国家管网追求跨越式发展的自身需求，同时也是落实国家部委三年行动计划关于“工业互联网+危化安全生产”的要求，国家管网集团深圳天然气有限公司（下称深圳 LNG）在安全作业管理平台速赢项目中与华为深度合作，聚焦安全生产领域，遵循集团安全标准、规范流程架构，在集团流程专班、数字化部、生产部及液化天然气管理公司的指导下，践行数字化转型的“流程”、“数据”、“IT”的试点，并取得了一些成效。通过 IT 系统规范作业人员的标准动作，提升了本质安全水平和工作效率。

9.2.2 解决方案和价值

9.2.2.1 行业挑战

当前在 LNG 接收站作业过程依赖现场监护人监督，巡检人员巡查，后方人员无法实时远程了解现场情况，风险隐患无法及时识别提醒，而且每天 LNG 接收站会有 10 到 30 项不等的检维修作业，现场监督人员不足，无法保证一对一全程监测，当前主要采用委托施工方主管监督或者一对多监管的方式，违规可能无法及时发现提醒。另外场站早期规划安防摄像头数量少，对作业地点的覆盖不足、拍摄角度和高度无法灵活适配现场采集视频图像、作业合规性的监督要求，且视频点位普遍不具备 AI 能力。

9.2.2.2 解决方案

整体方案基于国家管网的管理层级及各层级的业务职责，以“云边端”协同体系进行分层部署，实现数据、业务、运维的上下协同，整体架构如图 9-7 所示。

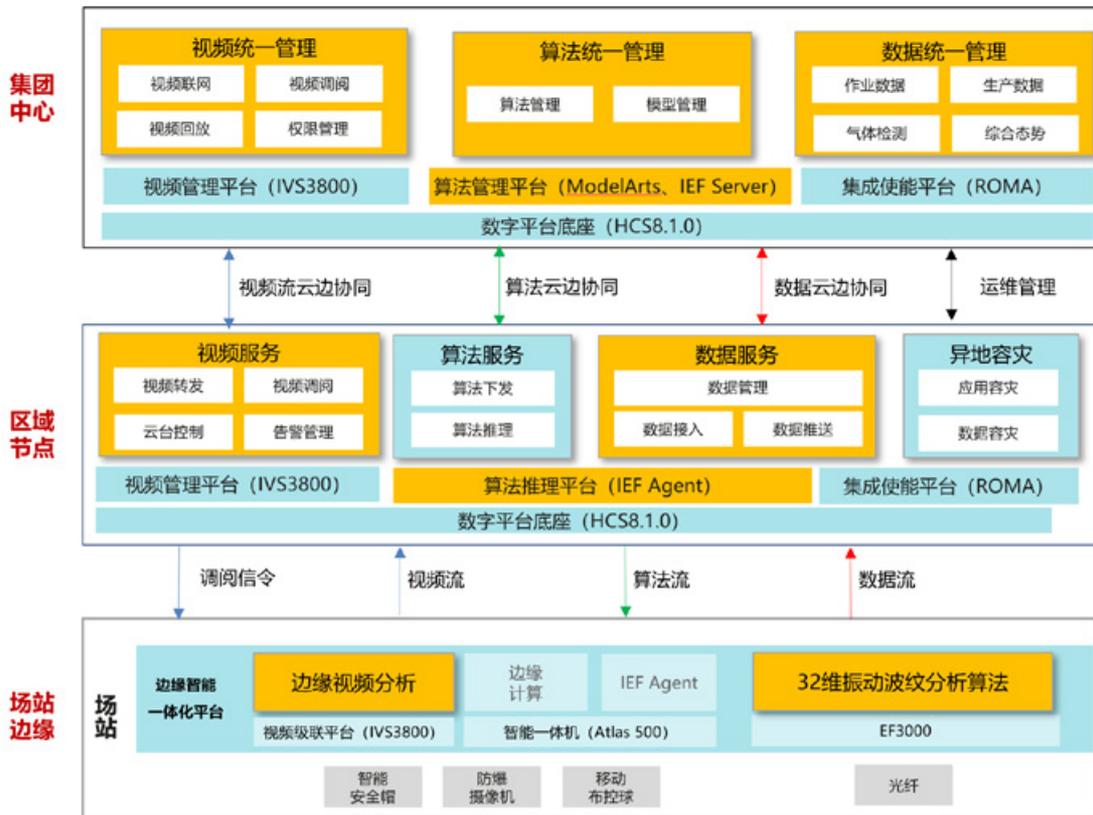


图 9-7 智慧管网云边端协同架构

为达到全面监测 LNG 接收站作业过程，通过在边缘侧场站部署视频 AI 分析算法，实现对现场生产安全风险、作业过程的智能监测监督，并通过云边协同技术实现集团和区域公司对现场画面的实时调阅监测、云端训练算法向现场的自动推送更新。

· 安全作业智能实时督导解决方案：作业全程实时可视，安全风险智能识别

方案集成场站内固定摄像机、移动布控球、智能安全帽等视频感知设备，可以基于作业维度对现场进行实时监测。后端监督中心/专家可与作业现场连线，实时了解现场状态，进行违规提醒，并可以对问题求助答复。

在场站边缘增加智能计算平台，引入作业前、作业过程的 AI 监督算法，有效弥补因人力不足带来的违规监督死角。在作业前对于需要特种作业资质的作业，作业申请人或监护人在作业现场通过算法关联作业资质，对特种作业人员进行人员、资质审核，审核通过后进行许可证审批。在作业过程中，使用 AI 算法对各类风险事件、违规行为进行自动检测，并及时上报告警，用于辅助监测并发现作业异常情况，现场作业监护人员根据告警提醒，在作业现场及时纠正，对告警问题进行回应，加快现场风险感知和响应速度，做到作业过程风险的可靠可控，作业中 AI 辅助监督的算法包括安全帽佩戴检测、灭火器识别、人员汇集检测、监护人离岗检测、火焰检测等。

方案基于云边协同的能力，可将最新算法模型统一下发各场站，持续提升算法识别的准确率，作业风险识别算法云边协同架构如图 9-8 所示。

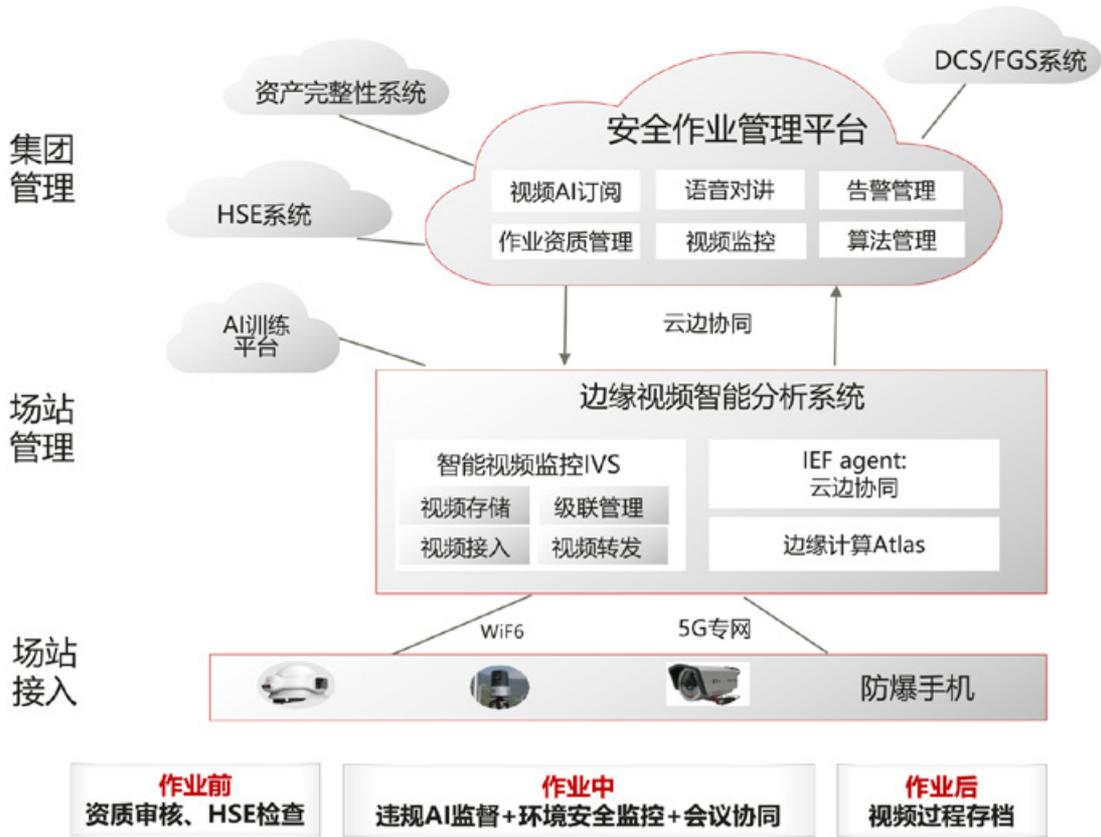


图 9-8 作业风险识别算法云边协同架构

9.2.2.3 方案价值

安全作业管理智能实时督导方案基于风险精细化管控、实时预警闭环的作业流程及管理需求，将 AI 算法、视频语音等 ICT 技术与流程融合，实现对风险作业过程的可视可管，监督覆盖率达到 100%，同时提升风险事件的感知和处理效率，减少了 70% 的作业安全风险事件，提升场站本质安全。

9.2.3 总结与展望

安全作业智能实时督导解决方案利用 AI 算法填补了油气基层风险作业中的部分监督盲区，人的不安全行为、物的不安全状态、环境的不安全因素进行全方位的管控，提升了场站的本质安全。现实作业中存在复杂繁多的风险事件，而目前的已有算法很难完全满足并且多数算法是具有行业特色的长尾算法，存在训练素材少、训练周期长、推广范围窄等现实问题，未来利用 AI 大模型，可以基于已有模型高效适配油气场景的算法需求，实现更多风险类型的 AI 算法识别，进一步提升监管精细度，减少风险事件发生。



第十章

机场和轨道交通

▶ 10.1 深圳机场

10.1.1 案例综述

深圳机场是中国境内集海、陆、空联运为一体的现代化国际机场，2019年旅客量突破5000万，迈入全球最繁忙机场行列，是中国内地唯一入选国际航协“未来机场”的试点机场，在2021及2022年度被评为Skytrax“五星机场”，2023年上半年旅客吞吐量2436万人次，国内排名第2，货邮吞吐量74.13万吨，国内排名第3。机场集团立足于粤港澳大湾区规划，立足于深圳“先行示范区，强国城市范例”建设要求，以及奋力打造世界一流的机场产业集团的目标，确立“客、货、城、人、智”五大发展战略，以智慧化为基础，打造重要国际航空枢纽，期望通过数字化转型，为旅客创造更好的出行体验，提升运行效率和安全，以及提升企业经营管理能力，打造智慧机场建设标杆，树立高品质发展的深圳样本。

深圳机场于2017年6月，与华为签署战略合作协议，正式启动数字化转型进程。2018年5月携手华为推进数字化转型，全面开启了数字化转型之路，2019年正式将智慧化确立为集团五大战略之一。深圳机场坚持规划先行，加大投入，积极探索云计算、大数据、物联网、AI等新技术应用，全面打造互联互通的ICT基础设施和平台，围绕安全，航班，旅客，建设“大运控”、“大安全”、“大服务”、“大服务”四大业务体系，分批建设、有效落地智慧应用项目，全面提升航班运行效率、安全监管能力和旅客服务体验。

10.1.2 解决方案和价值

运控是机场核心业务，在航空业务繁忙且运行及保障资源紧张的现状下，如何通过大数据、AI等技术提升机场运行效率是机场指挥面临的

难点问题，前期深圳机场面临的主要挑战为：在机位资源分配方面，深圳机场传统手工机位分配要专人进行 3 ~ 4 小时的分配作业，24 小时进行动态的机位分配和调整，机位调整耗时较长，效率较为低下，难以迅速满足计划外的分配需求。且人工机位分配效果不能用实时数据量化，单一机位的分配效果和全局分配效果无法整体评估，也很难快速准确地适应分配规则的动态变化，制约了机场运行效率的提升。

10.1.2.1 方案架构

双方基于“平台 + 生态”的理念构建起了“未来机场数字平台”，这个平台以华为 ICT 基础设施为底数字座，并联合生态伙伴构建行业生态系统，围绕“大运控”“大安全”“大服务”“大管理”四大业务体系，逐渐形成了运行“一张图”、安防“一张网”、服务“一条线”的新模式，总体架构如 10-1 所示。



图 10-1 深圳机场总体架构图

在大运控方面，围绕深圳机场核心业务流的航班流，以机场“全流程、全场景、全要素”的运行管理为目标，基于统一的数字平台及 AI 技术打造机场“运行一张图”，以构建运行更高效，保障更安全，协同更顺畅的机场运行整体解决方案。打造智能机位分配、航班保障节点采集、主动运控等场景化解决方案，实现机场运营管理中心对空侧、陆侧的全局精准可视、智能精准预测及多域高效协同，保障航班正常运行，提升运行效率与智能化调度。

智能机位分配场景化解决方案

作为业界创新型项目，华为联合深圳机场打造了智能机位分配场景化解决方案，这是 AI 技术在机场核心生产系统中的首次应用。智能机位分配系统基于深圳机场统一的数字平台底座，通过大数据对来自不同信息系统、不同部门的大量信息进行有效的数据融合处理，以 AI 算法为核心，打造智能机位分配解决方案，实现机位分配的自动化、智能化，同步构建易用、可靠的应用系统提升操作效率。方案架构如图 10-2 所示。



图 10-2 智能机位分配解决方案架构

项目在全面梳理深圳机场现有有机位资源分配业务流程的基础上，对当前机位分配特有的业务规则（9大维度,2800+细则）进行详细梳理、抽象归纳及数字化，构建灵活可配置的智能 AI 规则引擎。并基于靠桥率、冲突率、旅客体验等多目标综合考虑，利用华为“天筹”求解器快速求解的性能，打造智能机位分配 AI 算法，快速高效输出机位分配结果，最终实现机位“机器为主、人工为辅”智能分配的目标，有效提升机场运行效率。

10.1.2.2 方案价值

通过智能机位分配系统的建设，有效支持了深圳大运控领域的智能化升级及转型，提升了靠桥率、廊桥转率等核心指标，降低了场面冲突风险，助力机场机位资源分配的全局最优化，

在保证地面运行安全的基础上提高地面运行效率和旅客满意度，取得了成效如下：

第一、以目前深圳机场每日 1000 多架航班量级测算，通过智能机位分配系统建设，可使批量分配时间从原有耗时约 4 小时缩短至不到 1 分钟，动态调整耗时约 10 秒钟，极大提升了人工操作效率。

第二、深圳机场的靠桥率提升 5%，每年使约 260 万人次的旅客登机免坐摆渡车。廊桥周转率从 10.24 个班次提高到 11 个班次，相当于每个廊桥每天可多保障 1 个航班，卫星厅建成投运后，随着近机位资源的补充，该系统的效能进一步释放，当前靠桥率为 86.2%，有效提升了机位资源的使用效率。

第三、靠桥率的提升可带来远机位保障车辆平均每天节省 300 辆，降低了场面车辆与航空器在滑行道交叉点的冲突风险，有效降低了场面冲突风险。

10.1.3 总结与展望

在智慧机场建设的过程中，将大量采用人工智能等新技术，作为五千万级机场，IT 技术对业务的支撑作用必须有可持续性，在探索和拥抱新技术应用时，需控制新技术应用带来的风险。深圳机场与华为成立联合创新中心，不断探索新技术，并将最新的技术引入到机场业务中，目前 AI、大数据、5G 等新技术已逐步引入，未来通过联合创新的模式，更多的新技术将引入机场业务。

智能化是机场数字化应用创新的主要技术手段，围绕运行效率、旅客体验、安全保障，基于视频 AI、神经网络、运筹优化等 AI 技术的业务应用已经逐步嵌入机场日常运行的每一个环节，但单一场景、单一算法仍然具有其局限性，未来基于大模型的多场景多模态算法将会发挥更大的作用。如在面对机场日常运行中产生的海量数据、表格、文本、文档，利用大模型的理解能力能够从中抽取相应的信息，包

括文档智能信息抽取、OCR、延误事件分析预测等，提高决策能力；利用大模型的对话能力，可以打造全新的智能客服交互体验，提供专业建议和航班业务指引，更精准、人性化地响应旅客需求；利用大模型的图像识别能力，进一步提高图像分类和目标检测的准确率，对安检物品快速精准研判等，从而实现对违禁物品的实时监测和管理，保障运行安全等。通过 AI 及大模型逐步重塑机场业务，助力破解机场当前日益突出的保障资源瓶颈，全面提升机场安全保障、运行效率和旅客体验。



▶ 10.2 中国铁路郑州局集团有限公司

10.2.1 案例综述

中国铁路郑州局集团有限公司成立于 1949 年 3 月 11 日，位于全国路网中心，素有“中国铁路心脏”之称。它是国有大型铁路运输企业，也是承东启西、连接南北的重要交通枢纽，在中国铁路交通史上具有重要地位。郑州北车辆段管内京广线、陇海线、京九线，为全路运输最繁忙的区段，5T 检测车间共有 80 个检测工位，每天要完成 4 万多辆货车、280 多万张图片的检查任务，每人每天平均要处理 1.5W 张。随着货车车辆不断迎来提速重载，这对车辆运行安全构成严峻复杂的挑战。因此，郑州局迫切需要智能分析手段来提高铁路故障识别准确率，提升动态检车员的工作效率。

2021 年 11 月，国铁集团货车事业部把 TFDS 故障图像智能识别项目作为国铁集团第一批科研计划“揭榜挂帅”课题，指定郑州局集团公司郑州北车辆段 5T 检测车间和华为公司等参研单位共同研究，联手推进 TFDS 故障图像智能识别项目。课题成立后，华为组建了包括多名算法博士、AI 领域首席科学家在内的顶级技

术团队。在国铁集团及郑州局相关部门的指导和支持下，联合伙伴，多次深入作业现场，锁定了 AI 技术应用在车辆故障图像识别中的几个关键难题：

- **标准不统一**：不少图像介于故障和非故障之间，故障定义和判定标准不统一，影响智能识别系统准确性和一致性。

- **故障类型多**：铁路货车常见车型有 60 多种，涉及 400 多种故障，故障形态多样，难以穷举。同时，故障之间还存在复杂的因果和关联关系，增加开发难度。

- **样本分布不均衡**：实际中重大故障发生概率低、样本数量较少，轻微或者无故障的样本数量较多，样本不均衡，影响图像识别模型的学习效果。

- **干扰因素多**：铁路货车运行时会受到光照、雨雪、背景等环境干扰因素的影响，会导致图像中出现噪声或模糊，容易造成误报或漏报。

为解决上述问题，基于郑州北车辆段 5T 车间职工团队在故障分类、判断方式等方面的业务指导，华为联合伙伴探索出完整的可复制的 TFDS 货车故障图像智能识别解决方案，并在实际应用中取得良好的效果。



10.2.2 解决方案和价值

10.2.2.1 方案架构

华为 TFDS 货车故障图像智能识别解决方案概括为：中心训练、边缘推理、集中部署，总体架构如图 10-3 所示：

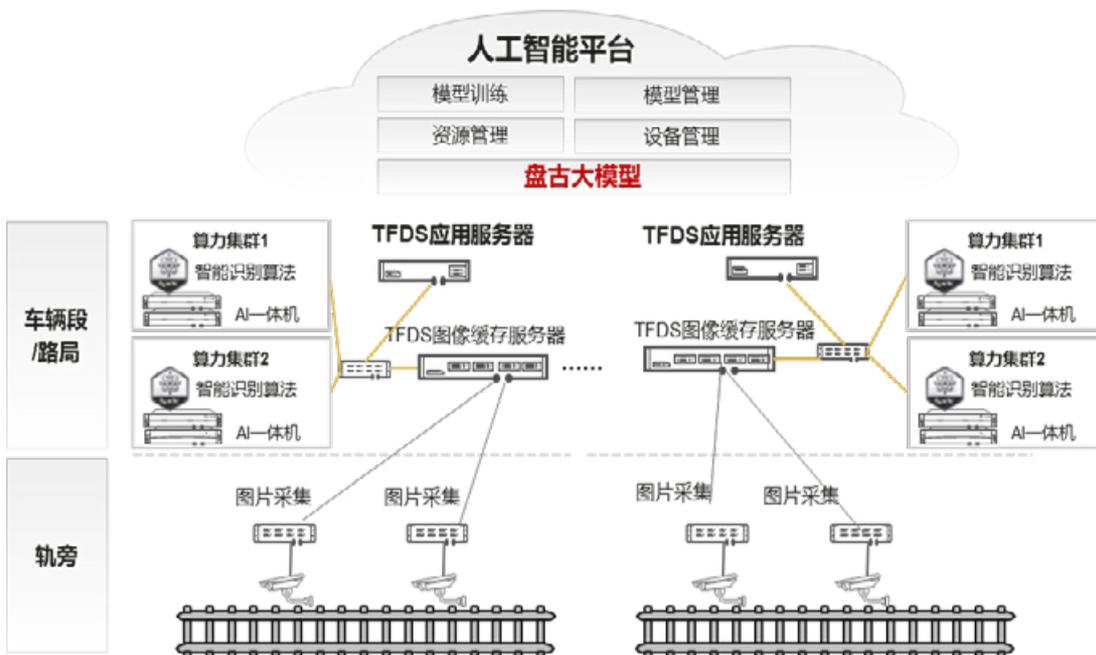


图 10-3 TFDS 解决方案架构

在人工智能平台，基于盘古铁路大模型进行算法训练以及调优。盘古铁路大模型为 TFDS 解决方案提供技术先进性的保障和支撑，基于 30 亿级参数的 CV 训练模型，可以极大降低算法训练周期，提高算法的迭代速度和准确度。TFDS 解决方案通过深度学习网络和大量数据样本，自动总结部件特征，自动寻找故障规律，并在实际试用中持续改善分析效果。针对车型多、故障类型多、样本分布不均衡、干扰因素多等问题，方案采用自适应增强检测算法、数据增强、图像重构、不均衡识别器等技术，实现从整体到局部、再到故障特征全量精细识别，如图 10-4 所示：

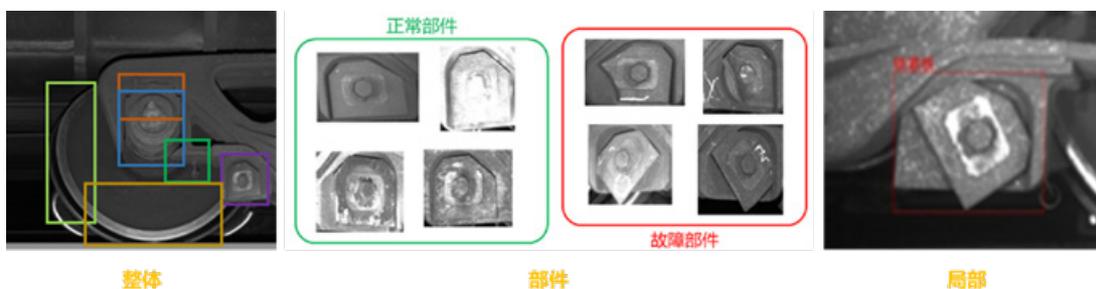


图 10-4 全量精细化识别过程示意图

先进行整图分析，定位和识别所有部件，分析大部件丢失（如摇枕弹簧丢失）、异物等故障；再根据故障部件样本，算法自动抽象总结异常特征，进行异常样本的初步筛查；最后识别关键特征（如部件间角度、间距等），结合故障标准输出故障（如锁紧板移位、手把关闭等）。

边缘侧，在车辆段机房部署 AI 一体机，通过铁路内部局域网连接到 TFDS 图像缓存服务器，获取过车图像数据，进行模型推理，识别故障图片。并将识别结果上传到作业平台，供人工确认复核。

另外，AI 一体机采用集中方式部署，可支持两个算力集群，根据过车任务实现动态资源分配，以解决铁路货车过车不均匀、列车高峰密集到达作业量突发等问题。

10.2.2.2 方案价值

TFDS 解决方案支持 -2/-3 两种类型的探测站设备拍摄的图像，车型适用率 95% 以上，覆盖 307 类《运规》故障以及 100 多类 TFDS 可视范围的其他故障，覆盖更全面，关键故障近“0”漏报。

郑州北车辆段已完成集中部署小规模试点，相比独立部署，资源利用率提升 30%。系统具备一定容灾能力，在部分算力故障的情况下，TFDS 智能识别业务可正常开展，系统稳定可靠。

通过 AI 识别，无故障图片剔除率 95%，辆均误报故障数小于 4，大幅减少人工看图工作量，

极大地提升了工作效率。TFDS 解决方案可以 7*24 小时高精度不间断工作，大大减轻了动态检车员的工作压力。2022 年底疫情期间人员到岗大幅减少，借助该方案，郑州局依然高效识别故障图片，保证车辆段正常运转。

10.2.3 总结和展望

如今的中国铁路，不管是在路网规模，还是在装备水平上，都处于全球领先地位。华为 TFDS 解决方案的使用推动铁路从传统作业向数字化转型，郑州局立足中原，服务四方，极大地提高了列车检修的智能化水平，节约数千万运营成本，促进从“肉眼看图”向“智能识别”转变。借助华为 TFDS 解决方案，效率提升的同时，大幅减轻了动态检车员的工作强度，也带来了全新的生产组织方式，实现了智能化和信息化作业形式，取得了一系列实践成果。

展望未来，除了 TFDS 货车故障图像智能识别场景外，AI 图像智能分析技术可在更多领域发挥重要价值。例如，在动车 TEDS、客车 TVDS 等相似场景，同样可利用 AI 代替人工看图；在动车运用所，AI 图像智能分析结合巡检机器人，可实现对动车的智能机检；在编组站、货场等运输站段，利用 AI 图像智能分析技术可实现远程货检。AI 图像智能分析技术未来将广泛应用于铁路各个场景，提升铁路智能化水平。



▶ 10.3 武汉地铁集团

10.3.1 案例综述

武汉地铁集团负责武汉轨道交通建设、运营、经营，截止 22 年 12 月，已建成运营 11 条线路，总运营里程达 460 公里，车站总数 291 座。在 2012 年 -2022 年 10 年间，武汉地铁实现了从无到有、从线到网的历史性跨越，从建设“城市地铁”走向建设“地铁城市”。2012 年至 2021 年，武汉轨道交通已累计安全运送旅客 68.19 亿乘次，2021 年武汉市轨道交通占全市公共交通客流比例达到 52.51%。面对日益增长的出行需求，武汉地铁一直致力于探索智慧轨道交通建设，为乘客提供更好的出行体验。

武汉地铁的智慧城轨建设不仅有着清晰的发展路径，而且也有着强大的技术支撑。2020 年 5 月 14 日，武汉地铁集团与华为公司签署战略合作协议，双方共同推进 5G、大数据、物

联网、AI、云计算等新 ICT 技术在武汉轨道交通上的应用和实践，打造安全、可靠的数字底座，实现地铁数字化转型目标，进一步促进武汉轨道交通高质量发展。

在双方联合开展的一系列转型探索与实践中，如何通过大数据、AI 等提升运营效率和服务质量是其中的一个重点工作。地铁运营高度依赖客流预测，而当前客流预测缺少足够的数据来源输入和精准的 AI 算法，预测精度较低；对于突发大客流及事件等，无法提前 30 分钟或者 1 小时预警，依赖人工进行调度及决策。因此，双方基于武汉地铁的城轨云平台，重点开展大数据治理共享及客流预测、预警等工作。

10.3.2 解决方案与价值

10.3.2.1 方案架构

面对线网化精细运营的挑战和诉求，智慧客运解决方案联合行业生态伙伴，通过业务流程分析和客运大数据资产，建立客流画像、线路画

像，使用机器学习和深度学习算法进行精准的客流分析、预测和预警。

基于城市轨道交通线网智慧运营调度对于客流大数据实时、精准、共享和全过程评估分析的要求，构建面向数据分析、标准统一的混合型数据资源仓库，汇聚多源异构数据，通过数据治理后形成不同颗粒度的数据资源层级，涉及贴源库、主题库和专题库等客流大数据资产，实现数据“统一采集、统一存储、统一管理、统一运营和统一服务”，形成线网级的数据和服务共享能力。其中，数据仓库采用高性能数据仓库 GaussDB（DWS），利用其 delta 及列模式存储模式，地铁客流亿级规模的 OD 量等客流统计分析效率可提升 10%；客流大数据除了传统的 AFC 刷卡数据，还通过视频 AI 分析、手机信令大数据等补充客流断面数据，支撑客流预测精度提升，方案架构如图 10-5 所示：

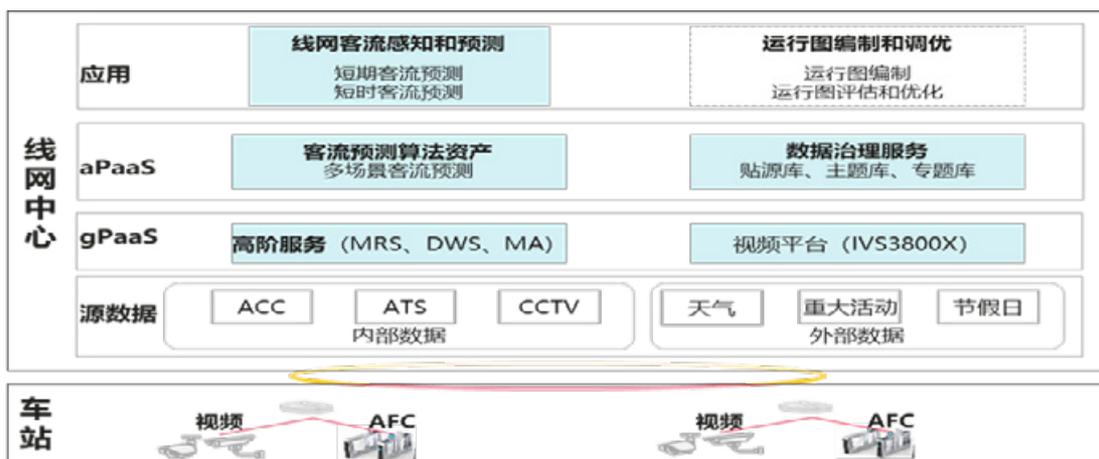


图 10-5 智慧客运方案架构

基于地铁客流大数据资产，综合考虑几十种客流影响因素，如早晚高峰、节假日、天气及大型活动等，构建客流 AI 预测推演模型与算法，经过不断的参数优化及算法训练，实现车站、线路和线网多场景、多状态的客流预测，包含客流量、客流服务能力和客流时空分布等，提升客运管理分析和客运服务质量。

通过客流实时监控、短时预测等手段，实现客流的状态和预测可视化，帮助调度人员、站务人员准确掌握客流实时状况和短时变化趋势，提前准备大客流采取应对措施和引导方案，提前时间从 30 分钟到数小时不等，提升客流管控效率并保障运营安全。

10.3.2.2 方案价值

武汉地铁基于大数据底座建设完善的地铁大数据资产，从业务上数据共享平台实现对武汉地铁集团全业务流程的覆盖，从能力上支持“智慧地铁”的构建，促进全国地铁全面数字化转型。通过智慧客运提升突发大客流、行车故障等场景下的客流应急疏导能力。针对不同维度和时长进行精准预测，精度提升到 90% 以上，实现提前 30 分钟以上的提前大客流应对准备。

10.3.3 总结与展望

武汉地铁智慧客运当前已实现精准的客流预

测，并初步实现客流的监察和预警。在智慧运营方向，后续将基于线网客流预测数据和运力资源统筹编制线网运行图，通过 AI 求解器的算法模型建立多目标多约束的运力调度模型，实现秒级求解，将运力精准投放到客流汇集面，解决线路间运能不匹配导致的站台、通道乘客积压问题，实现经济性和乘客体验的最佳平衡。

同时，在加速行业智能化和武汉智慧城轨建设方面，将继续围绕规划、建设、运营和经营等城轨业务，持续探索 AI 大模型与行业业务场景的深度融合和应用。





第十一章 公路水运口岸

▶ 11.1 天津港

11.1.1 案例概述

天津港是全球领先的大港，连续多年跻身世界港口 Top10，码头等级达 30 万吨级，航道水深 -22 米，拥有各类泊位 192 个，其中万吨级以上泊位 128 个，是中国北方第一个 5 亿吨港口，集团资产总额超过 1400 亿元。华为与天津港

集团的合作在 2020 年前主要在 ICT 设备层面。2020 年 6 月 24 日，双方在第四届世界智能大会上签署了《战略合作协议》，之后双方合作打造的水平运输、生产智能安防、集装箱智能计划等解决方案陆续在天津港落地。2021 年 11 月后，双方进一步在生产经营协同、平台资源整合方面深化合作，并将业务范围从集装箱码头拓展到散货、物流等领域。



图 11-1 华为与天津港合作历程

天津港北疆港区 C 段智能化集装箱码头项目是双方合作的一个典范，2019 年 12 月 28 日，该项目正式启动建设，设计年吞吐量 250 万标准箱、岸线总长 1100 米、泊位数量 3 个、码头等级 20 万吨、用地面积 75 万平米、前沿水深 -18 米、岸桥数量 12 台、场桥 42 台，ART76 台。2020 年 7 月，天津港集团公司联合华为、西井、主线等科技企业启动了智能水平运输系统建设工作，经过系统开发、实验室测试、和实船测试后，水平运输系统于 2021

年 10 月 18 日正式运营投产并在 2022 年 3 月达到 6 线作业，其后现场持续调优，于 2023 年 4 月实现 10 线同时作业。2021 年 10 月，天津港集团公司与华为以及北京天睿启动了生产智能安防系统建设工作，部署后经持续调优，该系统于 2022 年 3 月份启动试运行，通过持续采集数据优化算法模型，2022 年 5 月份系统正式上线。天津港北疆 C 段码头作为全球首个“智慧零碳”码头，真正实现了全自动化、少人、无人。



图 11-2 华为智慧港口解决方案总览

智慧港口解决方案整体示意如上，涵盖了一系列子场景解决方案，其中智能闸口、智能理货、综合安防、生产智能安防解决方案通过融合计算机视觉技术，有效减少人工进行实时识别及处理；智能计划通过融合运筹优化技术及调用天筹求解器，减少人工快速生成高效的生产作业计划；智能水平运输系统通过与自动驾驶、计算机视觉、运筹优化技术有效融合，达到了码头场景下超 L 4 级无人驾驶。

2021 年 10 月 17 日，天津港 C 段码头作为全球首个“智慧零碳”码头正式建成并投入运营，成为智慧程度高、建设周期短、运营效果优、综合投资低、适用范围广、绿色发展佳的自动化集装箱码头 2.0 标杆，同时也是全国港口行业首个获得权威机构认证的“碳中和”港口。

11.1.2 解决方案和价值

11.1.2.1 行业挑战

目前港口行业主要面临效率、成本、安全三方面的挑战：港口的运营，效率是核心，全年 24 小时不间断作业，而在单设备效率达到极限的情况下，因为码头生产作业“多因素、多目标、多约束、多时序、多变化、多工况”的特性，导致计划制定慢、作业协同难，资源难以最大化利用；成本方面，每台桥吊/龙门吊/集卡等设备，24 小时作业配备至少 3 名司机轮换，码头不是所有龙门吊同时作业，部分现场司机会闲置，而龙门吊司机是特殊工种，要求高培训时间长，招工难；安全方面，码头现场作业条件艰苦，除现场恶劣的噪音和震动外，几个小时不能下港机，禁食禁水，司机室 30 米高，长期低头作业，极易得颈椎病，司机视野有限、安全隐患大，一般 45 岁转岗。

11.1.2.2 解决方案

C 段码头的建设目标是要通过一个“稳定、安全、高效、智能”的信息化系统，实现对整个码头的现场操作人员、装卸设备、集卡等生产元素的自动化、智能化指挥、调度和监测，合理调配人力和机械资源，提高作业效率，达到运作高效、生产安全，实现对码头的生产作业的全面科学管理，满足智慧港口的实际业务需求，建成“全球智慧程度最高、运营效率最优、绿色发展最佳”的智能化集装箱码头。C 段码头整体架构示意如下，其中最核心的是智能水平运输系统、生产智能安防系统、以及智能计划系统。前两者已在 C 段码头上线，后者仍在测试中。



图 11-3 天津港 C 段码头整体架构

1. 智能水平运输，实现高效无人化运输

智能水平运输系统的目标是实现高效无人化运输，该系统在 C 段码头落地面临着—系列挑战：基于传统人工码头布局，地面交通态势错综复杂，水平运输系统需要具备强大的运筹学能力；支持变拓扑路径规划，需要强大的“智慧大脑”配合 ART “单车智能”实现拟人化的交通控制和车车协同；串行作业工艺 + 地面集中解锁，在进出堆场与上下岸的关键区域形成多向交通流，需要基于多任务优先级的路权分配策略，实现多向和多级路口连续通行，快速消除交通拥塞；自动化水平运输设备与外集卡在堆场出入口存在路

线交叉，需要具备高效的车路协同、车车协同能力，达到航陆运作业效率最优；码头布局中，自动化区域与非自动化区域交叉、常态化生产过程中，需要确保人机交互的本质安全。

为了有效应对以上挑战，华为和伙伴联合打造的智能水平运输系统，通过 AI 赋能智能水平运输系统统一协管控：通过监督学习分析历史作业数据，优化调度模型，提升系统整体调度效率；通过强化学习方法预测多智能体交互模型，多智能体协同规划减少冲突，提升安全；通过基于深度学习的自动驾驶态势感知，实现全港区全场景实时状态智能识别。



图 11-4 车路云网使能港口自动驾驶场景高效、智能

通过运动学仿真与 AI 模型训练提升协同效率与本质安全：基于 ART 运动学仿真的动态路径规划，是实现车 + 云协同自动驾驶高效、安全运行的核心技术。通过启发式算法实现自动驾驶路径的动态计算，通过 AI 训练优化车辆运动学模型，使能 ART 自动驾驶更流畅、更安全。

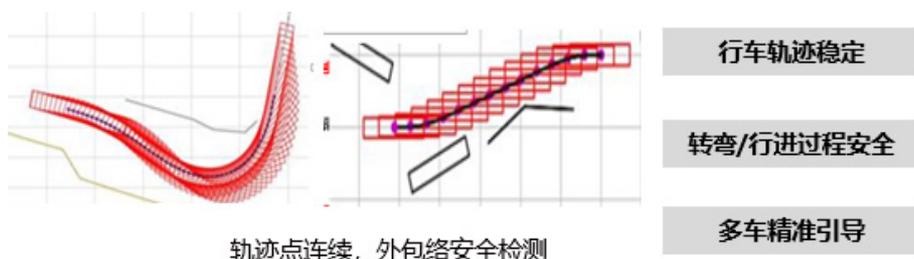


图 11-5 基于车辆运动学的全局路径设计

通过使用华为移动数据中心计算MDC(Mobile Data Center)平台,提高设备控制精度和可靠性。基于双轴控技术实现灵活转向控制,通过局部精细化引导提高ART停车对位一次成功率,根据路径曲率自动调节行驶速度、提升转弯效率。通过采用基于深度学习的人工智能技术驱动多传感器的感知、定位、融合,实现全场景、全时段、全天候条件下360度、100米半径、50米高度范围内的有效感知和全域定姿定位;控制算法引入强化学习与模仿学习,利用感知、定位信息提高路径规划的实效性和可用性。

华为车队管理系统FMS(Fleet Management System)作为水平运输的核心系统,通过对码头内自动驾驶车辆的统一调度,实现了岸边与堆场间集装箱运输车辆自动化作业。水平运输系统作为一个系统工程,生产流程的自动化打通并不是TOS系统业务时序+单车智能就能实现的,需要全业务流跨系统对接和基于云端的统一调度。车队管理系统通过统一标准接口,同时支持多种不同类型的协议,提供了对码头场岸桥自动化控制系统、交通系统、充电管理等系统的快速集成,其整体架构示意如下:



图 11-6 标准化构建水平运输系统外层集成能力

华为FMS通过全局路径规划的AI算法,基于车辆运动学特征,实现多车协同驾驶安全、高效的自动化。同时,云端的动态短路径规划,可解决码头作业场景多变、任务需实时调整的难题,实现了全局效率最优。通过采用北斗+5G+高精地图,结合路侧辅助感知,实现厘米级高精度定位,极大提升作业效率。同时,华为通过开放生态模式,实现云和车的解耦,目前已经兼容主流单车自动驾驶厂商(主线科技、西井、仓擎等)和徐工、国唐等多家单车厂商,不仅成本降低,而且可推广性更强。另外,华为FMS系统已具备大车队(96辆车)的常态化运营经验,且运营同时基于AI持续调优,实现了动态靠泊,有效提高靠泊效率和岸线资源利用率。

2. 生产智能安防,提升港口可视化管理能力

华为和天睿联合开发的生产智能安防系统的目标是提升港口可视化管理能力,该系统包含全景拼接、事件中心、综合态势、智能监测四个板块,如下图所示。

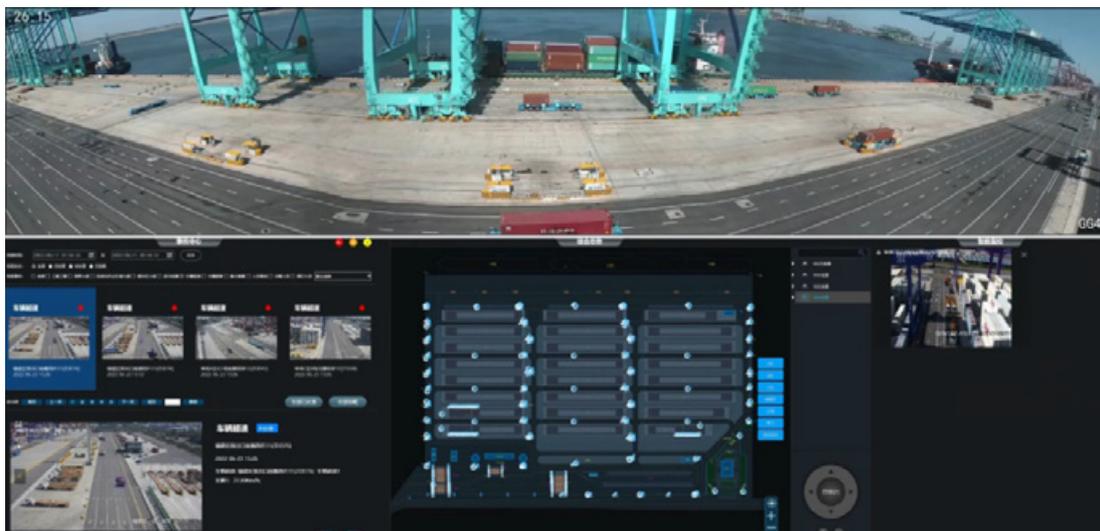


图 11-7 生产智能安防系统

该系统通过全景拼接技术，把单一的摄像机所显示的视频图像，通过坐标维度的再整合，以宏观角度呈现大场景大视角，通过一点即视、自动追视功能实现宏观与微观联动，做到大场景安全生产保障；通过 9 类算法模型解析码头监测摄像头视频流，实时发现自动化区域周界入侵、烟火感知、车辆逆行、车辆超速等事件，并联合伙伴开发匹配港口安全管理需求的应用系统，系统通过融合 AI、地理信息系统 GIS（Geographic Information System）、视频技术，实现 GIS 一张图，事件以及人、机、车、货、物信息通过 GIS 地图按位置呈现，系统支持事件与视频快速联动，管理人员可快速识别和处置事件，大幅提升安全管理效率，有效的保障港口企业的安全生产。

11.1.2.3 方案价值

天津港 C 段码头项目建设周期仅用时 1 年零 9 个月，人员配置降低 60%，集装箱倒运环节减少 50%，综合投资较同等岸线自动化集装箱码头投资减少 30%，并实现绿电功能 100% 自给自足。同时也荣获了一系列荣誉，

如 2022 年天津港、华为、中国移动共同荣获了世界移动通信大会 MWC（Mobile World Congress）的全球移动大奖（Global Mobile Awards）—“互联经济最佳移动创新奖”（全球首个获此殊荣的智慧港口项目）；天津港、华为、主线、西井等科技企业共同申报的“智慧绿色集装箱码头关键技术研究与应用”项目也荣获了 2022 年度“中国港口协会科学技术奖”特等奖。



图 11-8 天津港“5G+ 智慧港口”项目荣获 MWC2022 大奖

华为智慧港口解决方案在天津港 C 段集装箱码头的落地，代表着全球首创的自主研发的封

闭场景下的车路协同 L4 (Mind Off) 级无人驾驶在港口规模化商用落地；也是真正基于 AI 的“智能水平运输管理系统”首次成功，系统同时连接了智能闸口、自动化场桥、智能加解锁站、自动化岸桥全作业链，率先实现水平布局自动化集装箱码头全流程自动化作业“完整版”。迄今 92 辆无人驾驶集卡车满负荷常态化运营已超 18 个月，行驶百万公里，接管率低于 1‰。而生产智能安防系统在 C 段码头的落地，协助事件处置效率有效提高，结合 AI 智能分析，事件秒级提示，安保人员可快速到达事故现场，有效减少日常值守人员监测时间。

11.1.3 总结与展望

天津港已初步实现智慧港口 2.0，将各种人工智能技术融入了生产、管理、服务、安全等领域。C 段码头项目的成功，充分验证了自动驾驶、运筹优化、计算机视觉等 AI 技术在码头生产作业各环节的价值。除此之外，华为智能计划系统也已帮助天津港联盟码头多条航线船时效率提升 6% 以上，与聚时科技联合打造的基于深度强化学习的智能操控系统近期也将在天津港展开试点。

基于 ViT (Vision Transformer) 架构的计算机视觉大模型的出现，不仅使得智能安防、智能闸口、智能理货这些视频监测系统在算法模型层面统一成为可能，也使得集装箱残损识别、锁孔破损识别等小样本场景的识别准确率有望进一步提升。基于 Transformer 架构的语言大模型的出现，不仅使得集团和码头智能客服的智力水平有望得到显著提升，也使得未来通过大语言模型协调多个内部生产系统完成复杂任务成为可能。随着大模型领域技术的逐步成熟，港口行业的相关部门已考虑基于大模型训推一体机使用大模型服务，进一步提升港口智能化水平，提高港口运营效率，减少人员工作

强度。部分港口为了早日达成“建设世界一流港口”的目标，也已开始筹划智慧港口 3.0 的准备工作：在装卸生产领域，智能化从单一码头走向多个码头群、机器从自动化走向类人操作、水平运输从有限混行走向完全混行；在物流贸易领域，自动驾驶从码头内拓展到物流园区、物流服务从分散式走向一站式、物流周边生态数据从孤岛走向共享；在运营管理领域，报表从人工生成走向 AI 自动生成、决策从经验判断到数据驱动、经营与生产从割裂到联动；在生态服务领域，从单一运输港 / 工业港走向港产城融合、服务窗口从多个集中到一个、数据也逐渐走向规范化，标准化。

港口作为国家战略基础设施，未来智能化应用必须要建设在安全可控的数字化底座上，目前各港口集团均已发起针对 TOS 安全可控的攻关，基于安全可控的计算、存储、中间件、数据库、操作系统、以及 AI 引擎建立云化、微服务化的能力。在这个过程中，随着大模型为代表的 AI 技术的逐步融入与赋能，码头操作系统 TOS (Terminal Operating System) 的智能水平也将持续提升，为智慧港口新业态、新模式的发展提供更好的支撑。得益于这些 AI 技术驱动的科技创新，港口客户正携手华为及合作伙伴共同向智慧港口 3.0 稳步迈进。

▶ 11.2 深圳市交通运输局

11.2.1 案例概述

深圳市交通运输局负责交通运输（道路、枢纽、场站、港口、航道、空港、道路交通、道路运输、水路运输、城市公共汽电车、城市轨道交通、出租小汽车等）管理工作，协调铁路、民航、邮政、海事等涉地管理事务，深入推进全市综合交通运输建设，为城市交通综合治理、运输行业管理、交通出行服务提供了强大支撑。进入“十四五”期间，随着《数字交通“十四五”发展规划》、《交通领域科技创新中长期发展规划纲要（2021—2035年）》等指引智慧交通、交通科技的顶层指导文件陆续印发，城市交通数字化进程加快，深圳市面临着先行示范区、交通强国试点城市等现代化城市标杆创建的建设要求。与此同时，随着深圳市经济快速发展，机动化出行率持续提升，城市交通拥堵持续加剧、营运车辆安全问题突出、信息化壁垒和数据烟囱等严峻挑战，传统的交通信息化业务系统建设已难以满足当下城市交通快速发展的数字化业务需求。为此，深圳市交通运输局基于深圳市政府提出“便民利民、互通共享、自主可靠、开放创新”等方面的指导意见启动了统一底座、统一平台、统一能力的《深圳市交通运输智慧平台建设项目（一期）》的立项与建设。

深圳市交通运输智慧平台项目建设目标是围绕“安全、高效、便捷、绿色”四大主题，依托科技信息化技术手段，打造更加安全可靠、绿

色便捷、智能先进的城市交通运输管理和完整出行服务体系。

深圳市交通运输智慧平台建设项目（一期）是基于华为云计算、大数据、AI等技术，通过汇聚海、陆、空、铁各类交通数据，以“互联网+”思维重塑运输管理模式，再造行政服务流程，打造智慧交通管理服务体系，面向智慧设施、智慧公交、智慧运输、智慧执法及综合交通治理等五大方面，基于数据的“监测-预警-处置-指挥-发布”实现交通运输日常稳定运营和重大节假日安全应急智慧管控，构建“全息感知、一体监测、精准预警、调度指挥、全程服务”的面向平战结合的新一代交通运输智慧管控体系，打造智慧交通建设的“深圳样板”。

深圳交通运输一体化智慧平台项目是以数据技术一体化，促进交通运输内外部高质量协同一体化，深圳市交通运输一体化智慧平台将围绕“感知-调控-服务”的主线开展建设。数据技术一体化是以人、车、企业、设施等跨行业动静数据融合，实现复杂交通系统演变机理和规律深层次认知；内部应用一体化是以行业全链条治理一体化、设施全周期管理一体化及跨方式综合调控一体化，推动数字时代交通业务流程再造；外部协同一体化是以城市跨部门、跨区域协同治理和政企合作服务新生态，促进城市治理与服务高质量协同发展。

项目建设内容如下图所示：



图 11-9 深圳交通运输一体化智慧平台总体架构

深圳交通运输一体化智慧平台以华为云计算平台为载体，大数据平台为支撑，汇聚交通、数政、交管、应急等行业数据通过数据治理构建交通大数据平台为基础，使用人工智能、数据治理、数据分析、BIM、GIS、视频分析、物联网、融合通信等技术，基于局内各业务处室职能与业务，面向设施安全管养、运输安全管理、综合执法管理、公交资源管理、枢纽出行服务、行业信用管理等 6 类交通业务进行业务需求分析与功能需求分析，通过整合业务流、数据流，最终建设综合交通一体化监测、交通规划建设、交通综合治理、设施安全管养、应急指挥调度、交通综合执法、公交资源调控、枢纽出行服务、运输安全监测 9 大应用体系，打造交通运输业务监管平台。功能架构如下图所示：



图 11-10 深圳市智慧交通业务架构

11.2.2 解决方案和价值

11.2.2.1 行业挑战

近年来，随着深圳高速发展及湾区一体化建设加快推进，深圳智慧交通已经取得了显著的成果，但是在交通事故预防、交通拥堵治理、保障供需平衡等方面还面临着诸多挑战。

1、交通事故预防

在交通安全方面，深圳市约有公交、危险品运输、泥头车、长途客车等营运车辆 120095 辆，占机动车保有量的 2~3%，但每年发生的伤亡人交通事故却占交通事故总量的 50%，万车死亡率也是一般车辆的数倍，交通安全行驶十分严峻，希望通过本次项目建设，全市平均万车死亡率降至 0.75 人以下。

2、交通拥堵治理

随着深圳机动车保有量的进一步增加，道路交通流量的制约，加速了城市交通的拥堵，根据互联网数据统计分析 2023 年 Q2 路网高峰行程延时指数为 1.724，85% 的通勤出行时间在 1 小时以上，高峰干线公交运行速度将由现状的 22.3 公里 / 小时，绿色出行流量分担率低于 60%，因此在交通拥堵治理、高效绿色出行方面也面临着严峻挑战。

3、公交运营与服务

传统公交运营模式存在几大行业挑战：一是驾驶员、车辆等生产计划制定不准确，或各种原因导致计划变更，造成发车计划不稳定，关键运力资源无法保障；二是营运排班采用传统信息录入的作业模式，部分城市甚至采用手工作

业模式，整体耗时长，方案不灵活；三是全面新能源化，公交企业的运营支出由燃油成本逐步转换为电力成本，无需用电易导致成本浪费。

随着新能源公交车辆的持续推广，公交企业面临着从传统运营模式到数字化转型的难得机遇。如何通过人工智能、大数据等技术能力，在现有公交系统的基础上提升公共交通供给侧效率，提升自我造血机制，成为缓解城市交通拥堵、提升居民出行服务水平，公交可持续发展的重要环节。

表 11-1 某典型城市的阶梯电价

| 类别 | 时段 | 电价 (元 / 度) |
|----|----------------|------------|
| 高峰 | 9:00-11:30 | 1.027 |
| | 14:00-16:30 | |
| | 19:00-21:00 | |
| 平时 | 7:00-9:00 | 0.675 |
| | 11:30-14:00 | |
| | 16:30-19:00 | |
| | 21:00-23:00 | |
| 低谷 | 23:00- 次日 7:00 | 0.231 |

4、客运枢纽供需匹配

随着中国抗击新冠疫情取得决定性胜利，各省市之间的人员出行流动性在迅速恢复，2023 年 1-6 月全国营业性客运量 43.2 亿人次，同比增长 56.3%。庞大的客运出行需求带来重要枢纽节点的高质量、高稳定服务与管理需求的挑战。

一是客运量激增，带来枢纽内部统一的管理指挥与运营调度压力加大；二是出租车辆、公交车辆的服务调度与客流不匹配，易导致驾驶员

或乘客等候时间过长；三是停车位的周转效率不足容易影响枢纽配套车库内部交通流，导致私家车进、出枢纽不便；四是海量视频覆盖枢纽各个位置，依托传统人工巡查的方式，发现潜在风险与隐患效率有所不足。

如何实现整个枢纽统一的平台支撑、统一的运营调度、统一的综合管理、统一的旅客服务、统一的安全保障等，是智慧枢纽建设的主要挑战。

5、设施管养缺乏监测预警体系

在重大交通基础设施运行安全管控方面，缺乏基础设施监测预警体系，不满足城市级大规模交通基础设施安全管控的需求。目前市交通局目前负责的设施管养内容包括道路 9184 条、桥梁 3285 座、隧道 100 座、边坡 2998 座等，现阶段交通运输局对全市 7 座桥梁，9 座边坡开展了实时监测，基础设施管理仍主要采取人工巡查的方式和定期检测的方式开展工作，导致基础设施类灾害预警评估所依据的数据来源不全面、不准确且滞后，基础设施类灾害的预

警预报缺乏及时性、有效性和准确性，另一方面，安全智能监测体系尚未建立，在只有片面实时监测信息和传统数据的基础上，无法开展管养决策准确性研判，无法满足城市级大规模设施运维的需求。

11.2.2.2 解决方案

目前深圳交通感知设备建设已基本完善，构建覆盖较全的数据感知体系，但数据感知尚未突破人力巡查无法覆盖的“盲区”，各类采集设备功能发挥有限，对人、车、路、场、环境等交通安全要素状态的感知没有实时化，智能化，精准度不高，城市交通综合体征检测和全息感知能力不足，平战结合的一体化管控与多业务协同有待加强。因此需要通过建立深圳交通运输一体化智慧平台项目，实现跨委办局、跨区县的数据汇聚，通过 AI 的 CV 能力、预测能力、多模态能力、NLP 能力，助力实现交通事故预防、交通拥堵治理、枢纽出行服务、信息接报服务等业务场景。

整体逻辑架构如下图所示：



图 11-11 AI 解决方案架构图



深圳交通运输一体化智慧平台项目以华为云计算平台为载体，大数据平台为支撑，跨委办局、跨区县融合汇聚交通、数政、交管、应急等行业数据，通过数据治理构建大数据资源中心，利用 CV 能力、预测能力、多模态能力、NLP 能力，基于视频实现对人、车、路、事件的快速智能识别，支撑交通事故预防、交通拥堵治理、枢纽出行服务、设施智慧管养等应用场景。

当前已经接入“雪亮”工程等视频监测数据，包括高速公路收费站视频、隧道视频、公交首末站视频、城市干道视频、港口码头视频等共计约 31000 路，使用大模型 CV、NLP 等能力基于 3131 路进行视频智能分析，包括机动车特征识别、交通事故检测、交通事件检测、交通违法检测、交通流量检测、营运车辆管控、班车违法上下客识别算法、车辆违规运输算法、占道施工算法、路面病害识别算法、隐患二次识别算法以及对道路、桥梁、隧道、边坡等重点设施健康状态监测算法，构建一个全天候、全时空交通事件监测、交通安全运营的交通空间。

同时还跨委办局、跨区县汇聚局内外 20 个业务部门、9 个企业单位，66 个业务系统、930 种数据，共计约 6000 亿条数据。基于这些海量的结构化数据使用大模型预测、多模态等能力进行模型分析，包括公交线网优化、客流量

分析预测、供需匹配优化、信号控制优化、交通流量分析与预测、交通 OD 分析等，针对不同的预测结果采取不同的处置手段，保障交通高效、便捷运行。

在跨委办局、区县数据共享层面实现与市交警局、规资局、城管局、区县交通局等部门间的信息共享共用；在业务层面通过系统对接，与市应急局、交警局、住建局、区县交通局等实现应急协同指挥、运输安全管理等业务联动；在信息资源层面充分复用市政政务云资源，避免重复建设；并依托运行指挥中心，纵向与市政府管理服务指挥中心联动，实现市级层面的日常运行监测和事件处置联动；横向与区智慧城市运营中心联动，服务区级交通数据共享、辖区特色交通监测和相关业务处置。

1. 通过交通事件智能检测助力交通事故预防

交通事件智能检测依托 AI 技术结合安装在公路、桥隧等道路上的视频监控流自动检测公路、隧道发生的交通事件，包括交通事故、交通拥堵、逆行 / 倒车、异常停车、路面抛洒物、路面积水、烟火等。检测出来交通事件后可以自动推送到指挥中心，指挥中心接收到预警信息后可以快速处置，避免发生次生事故。



图 11-12 交通事件检测 I

注：如图所示从左上到右下分别是交通事件检测、交通拥堵检测、倒车检测、异常停车检测。



图 11-13 交通事件检测 II

注：如图所示从左到右下分别是机动车逆行检测、路面抛洒物检测、路面积水检测、烟火检测。

通过针对机动车违法检测与预警，降低交通事故发生概率。

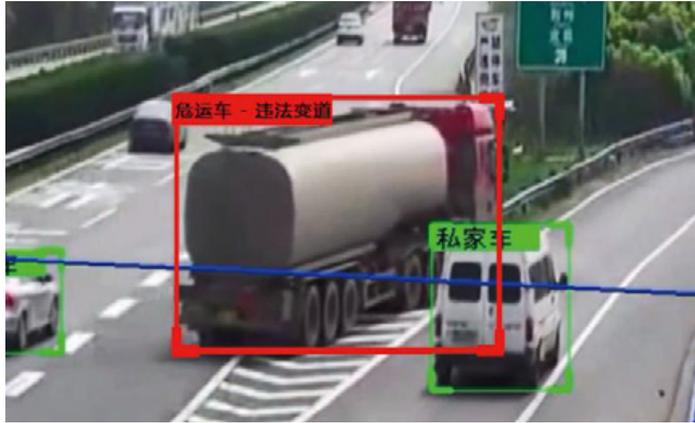


图 11-14 机动车违法检测

2. 通过供需匹配优化助力交通拥堵治理

基于视频交通流量检测、客流量、卡口过车数据进行交通流量和 OD 分析，掌握出行规律，识别并设置 HOV 车道、公交车道、潮汐车道的路段以及拉链式通行交叉口，提升交通运输效率，实现 85% 通勤出行在 45min 内可达。



图 11-15 交通拥堵治理

3. 智慧公共交通

公交智能排班通过汇聚公交企业源业务系统的 IC 卡刷卡数据、车辆 GPS 数据、线路及停靠站数据、首末站数据、车辆属性及驾驶员等相关数据资源，基于公有云服务提供 ModleArts 算法引擎及智慧排班算法，形成标准化服务接口供公交企业或第三方应用系统开放厂商调用，输入相关参数（开收班时间、车辆数、站场埋车数等）后即可调用 AI 算法引擎。

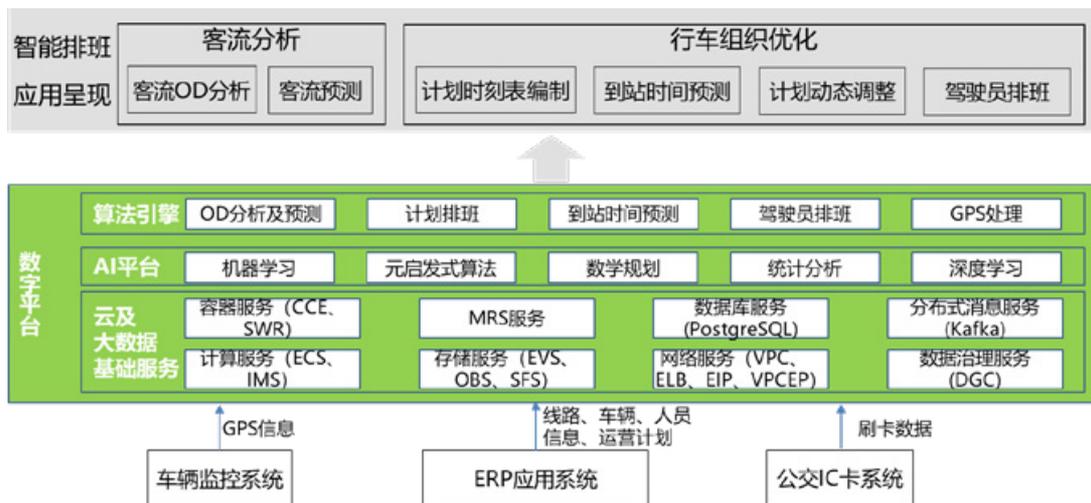


图 11-16 智慧公交排班流程

公交智能充电，通过基于华为云的算法引擎，合理安排公交车辆班次和充电的计划，让公交车辆尽量利用低谷时段充电，减少平峰和高峰充电时间，从而降低新能源公交企业整体用电成本。

一是夜间充电，根据车辆的剩余 SOC，通过 AI 充电模型预测出每个车辆在不同充电桩充满电需要的时间，然后根据充电时间把充电任务合理分配给各充电桩，尽可能让充电桩利用率达到最大。

二是白天补电，电量不足以支持后续班次计划时，可以通过调整班次计划，来减少白天补电。如有的班次车辆电量不足时，利用其他电量充裕的空闲车辆完成后续班次。一般不增加或减少班次，只是改变班次对应的车辆或者对应的驾驶员。

4. 智慧交通枢纽

智慧交通枢纽通过枢纽场站内部的感知与传输等信息化基础设施，以华为云计算平台为承载，建立“感知 - 分析 - 预警 - 决策”全链条闭环管理的调度组织与运营管理业务流程。

枢纽交通运行感知系统，基于部署于枢纽周边道路与枢纽内部的智能摄像头，可以实时监测到周边

道路的车流画面及枢纽内部的客流画面，通过基于 Atlas 服务器与昇腾计算卡的 AI 视频解析能力，将相关数据结构化并形成动态数据推送至 ModleArts 智能分析平台。

智能分析平台聚合枢纽内部运行状态、出行配套交通资源、周边道路拥堵情况等关键要素，基于交通客流分析能力，结合列车准点情况，出行需求匹配算法对当前枢纽运行情况进行分析，动态识别枢纽的客流滞留、拥堵位置，对高于阈值的分析结果进行预警，并通过 NLP 大模型生成预警信息报告，自动向管理者与公众通过不同渠道进行及时提醒与发布。

5. 通过交通基础设施智能监测自动巡检设施健康状态

通过安装在桥梁、隧道、边坡、护栏、道路上的视频监测流，可以依托 AI 技术通过视频监控对桥梁路面状态、隧道路面状态、护栏破损、路面坑洼、标线不清晰进行自动监测，当发现隧道路面积水等异常情况后第一时间发送预警信息，交通管理人员可以进行快速处置，降低交通安全隐患。



图 11-17 护栏状态监测和路面健康检测

11.2.2.3 方案价值

AI 在城市综合交通运输一体化项目中建设的应用，为城市综合交通运输智能监测提供核心支撑，通过 AI 技术在各业务流程的应用，实现综合交通运输全流程提升，带来多方面的 AI 价值体验。交通事件智能监测，减少 50% 的人力资源投入，突发事件 AI 算法主动识别、自动上报预警，让事件响应时间缩短 30%，避免发生二次交通事故；通过危化品车辆智能检测、营运车辆违法检测、营运车辆驾驶员状态检测有效的预防交通事故，为交通违法治理提供智能化的技术分析手段，营运车辆发生交通事故数量降低 30%。

通过公交线网及设施优化、公交运营服务考核、卡口过车数据分析，深圳 85% 的通勤出行时间在 45min 内，需要通过个体出行链需求分析、公交潜在客流识别与大公交多方式结构调控，实现绿色出行分担率达到 85% 以上，大幅提升城市竞争力；通过公交线网及设施优化，实现高峰干线公交

运行速度将由现状的 22.3 公里 / 小时提升至 25 公里 / 小时，高峰单次出行公交站点停靠延误将由现状的 8 分钟下降至 6 分钟以内，出行服务更加高效快捷。

依托 AI 技术通过视频监测对桥梁路面状态、隧道路面积水、护栏破损、路面坑洼、标线不清晰进行自动监测，自动预警，防范于未然，提升 60% 巡逻效率。

11.2.3 总结与展望

深圳市交通运输局在综合交通运输一体化项目中，主要还是通过使用小模型来实现交通事件智能监测、危化品车辆智能检测、营运车辆违法检测、营运车辆驾驶员状态检测、基础设施监测等场景，虽然极大的节省了人力成本、降低了人工识别的误差，但是在准确率方面还有提升空间，另外还存在超重、超高等小模型不能识别的场景方面还有扩展空间。未来，可以充分利用综合交通运输行业大模型使得模型算法更加精准、场景更广泛，能够高效的实现交通事件发现、交通事件处置、交通事故预防、运输能力提升、民众出行服务。





第十二章 智慧城市

▶ 12.1 深圳市福田区政数局

12.1.1 案例综述

深圳市建设鹏城智能体致力于民生服务、城市治理和数字经济的高质量发展。通过建立多级联动的协同机制，实现统筹与分布紧密结合，解决复杂城市问题。福田区政务服务数据管理局贯彻落实党中央、省委、市委关于政务服务、电子政务、数据管理、网格管理工作的方针政策和决策部署，按照区委工作要求在履行职责过程中坚持和加强党对政务服务、电子政务、数据管理、网格管理工作的集中统一领导。福田区政务服务数据管理局根据市级的统一规划，强化区级系统集成、场景赋能，构建横向全覆盖、纵向全联通的城区运行管理平台，打造数字政府全域治理现代化典范。

目前福田区在城市管理中只有街道级巡查，事件的自动分拨过程存在分拨不准确，日均退单

率高，统计分析数据不准确的问题。福田区需要搭建城市治理一体化联动处置系统，实现事件上报、分拨、处置与效能评估，能够将城市治理事件进行统一管理。通过对福田区现状的分析，制定统筹规划、分步实施的建设原则，第一阶段在全区街道部门全面上线“民意速办”系统，并且通过视频AI算法实现事件智能发现，推动“全域治理”的工作任务落实；第二阶段实现工单自动分拨，并且在香蜜湖街道、福保街道试点，区分拨系统上线，智能分拨准确率达到90%；第三阶段实现效能分析，通过对民意事件数据的主动分析研判，为领导决策提供数据支撑；第四阶段持续推进福田区物理实体空间数字化，推进城市基础空间数据与政务数据深度融合，基于CIM底板展示“民意速办”区域工单处理情况，提供区域画像、区域管理能力评估等功能，提升城市精细化治理水平。

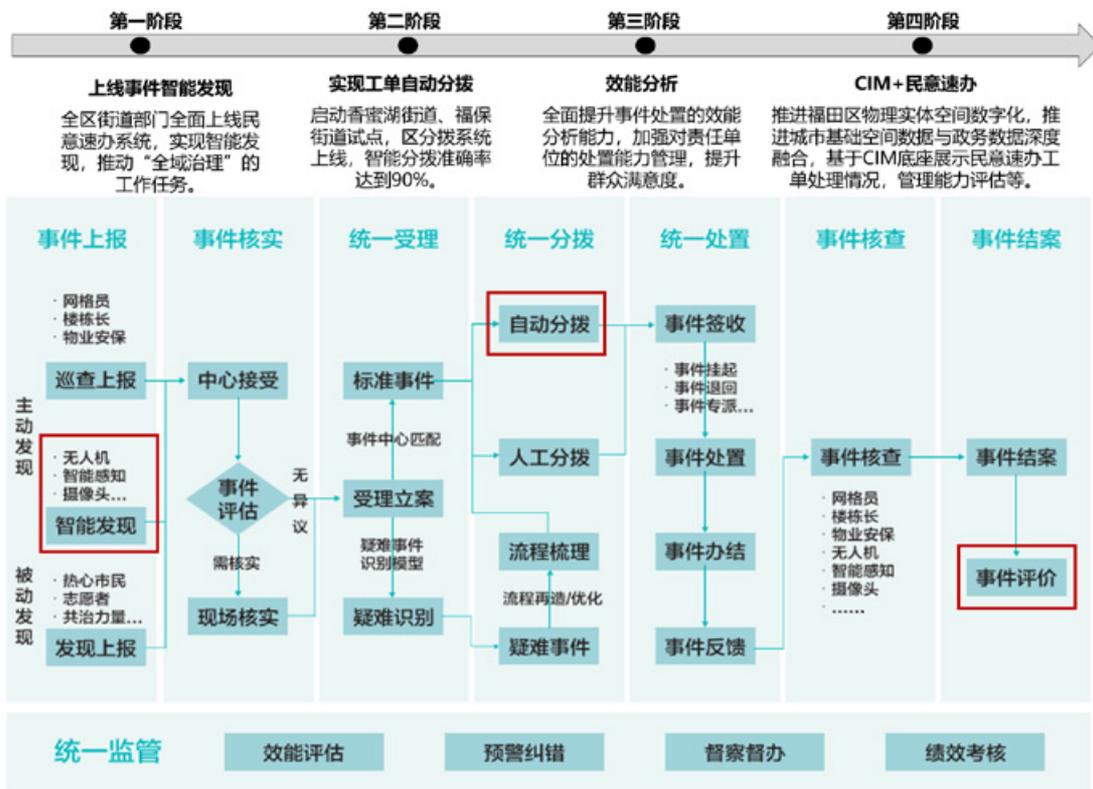


图 12-1 “民意速办”一网统管业务流程

福田区“民意速办”是福田区数字化转型的典型应用，面对群众诉求反映渠道多、多个单位分头管理，办理标准不一致等问题，福田区通过搭建全覆盖的立体运行网络，创新全流程的闭环运行机制，灵活运用 AI 技术，建设全智能的“民意速办”平台，通过智能发现、自动分拨、事件评价等，打造“民意速办”一网统管业务系统。

12.1.2 解决方案与价值

12.1.2.1 解决方案

针对城市治理过程中事件发现时间晚、处理周期长以及效果难纪录等问题，设计交互层、联结层、中枢层、应用层的四层总体架构，实现数据归集、智能分析，并高效服务于福田区城市治理业务。

(一) 总体架构

福田区“民意速办”解决方案的总体架构分为 4 层，分别为智能交互、智能联接、智能中枢、以及智慧应用，全方位支撑“民意速办”业务高效处置城市治理事件。



图 12-2 “民意速办”一网统管总体架构

(1) 智能交互层通过前端感知设备、视频监控设备等终端的交互，面向城市地上+地下的立体环境展开城市感知体系建设，全面感知城市治理的业务数据，实现城市事件的动态感知，精准控制。

(2) 智能联接层面向城市治理关键场景，提升电子政务外网支撑能力，打造 IP+ 光一体化政务外网，构建最优网络架构，将电子政务外网向街镇、居村延伸；构建城市级物联感知网络，为城市治理全面感知、实时互联和数据共享开放奠定基础。

(3) 智能中枢层是实现城市治理数字化的基础，将城市多云进行统一管理，协同市区政务云资源，构建逻辑集中的城市治理数据湖。依靠业务使能、数据使能和 AI 使能平台，提供包含视频 AI 智能发现算法、事件智能查重算法、事件智能分类算法等，为智能应用提供集约化的平台资源和能力。

(4) 智慧应用层依托城市治理需求，通过慧眼识事、智能调度、联动处置、效能分析等行业应用，精准服务民意渠道统一、民意汇聚、民意速办、民意处置、民意分析等多个场景，提升城市事件管理水平和部门协同效率。

(二) 应用场景

(1) 事件智能发现场景

在进行城市精细化管理过程中，传统方式依赖人工进行巡查效率低下，指派到基层的巡查人员发现事件的实效性较差。部分项目通过视频 AI 方式识别城市运行事件，但视频分析任务的作业主要依靠人工启停，算法 24 小时运行，消耗较大算力。

福田“民意速办”项目使用视频 AI 智能分析技术，全面覆盖城市治理自动化事件上报场景，精准识别事件。各算法模型支持按时间、按业务潮汐灵活调度，根据治理场景设置 AI 算法

运行的时间策略，有效提升计算资源利用率。



图 12-3 视频 AI 智能发现场景

(2) 工单智能分拨场景

通过事件发现模型与事件中心联动，30% 以上的城市治理事件自动发现、自动立案，项目中福田区城运工单的自动分拨率和分拨准确率均超过 90%。同时，基于 GIS 地理信息系统和事件中心权责清单等信息自动派遣工单，提升城市事件多环节处理效率。



图 12-4 城运工单自动分拨流程

(3) 效能智能分析场景

事件处置完成后，对事件处置结果进行智能分析，通过构建诉求感受指数及办理感受指数，形成事件处置指标评价体系，准确反映市民评价情况，以评促改，提升各责任单位的城市治理能力。

12.1.2.2 方案价值

福田区通过构建民意速办平台，不断深挖 AI 技术在城市治理方面的价值，有效提升福田区城区治理体系和治理能力现代化水平，逐步打造科学化、精细化、智能化的城市治理品牌。

AI 智能技术的应用，为事件高效处置的智能化、自动化提供了可能，相比传统的城市治理事件处置模式，AI 技术能够更好地实现事件全时响应，群众诉求快速办理。福田区通过智能发现 + 工单自动分拨等功能，整合全区 30 个以上的民生诉求反映渠道，建立“7×24 小时”全

天候服务，工单自动分拨率和分拨准确率均超 90%，快速类事件在 2 个小时内分拨、处置完毕。深圳福田区“民意速办”平台自 2022 年 5 月上线至 2023 年 6 月，共受理 46.55 万件民生诉求，按时办结率达 99.99%，平均办理时长比之前压缩了 65%，总体满意率达 98.8%。

12.1.3 总结与展望

深圳市福田区政府在《深圳市福田区国民经济和社会发展第十四个五年规划和二〇三五年远景目标纲要》中明确提出推进城区精细化治理，

加快城区空间治理、安全治理、生态治理能力现代化，努力走出一条符合超大城市中心城区特点和规律的治理新路子。福田区第八次党代会、第八届人代会明确提出建设“民意速办”新品牌。

目前，“民意速办”行动已经成为深圳市福田区的城市治理名片，构建了高效组织运行体系、整合各类民意信息的平台，实现了“多口归一”民意信息汇聚，“智慧一体”民意事件速办，“分头推送”民意情况反馈，“主动分析”民意研判分析等实际管用的城市治理应用体系。项目得到中央电视台“焦点访谈”栏目的专项报道。

下一步，福田区将持续推进 AI 技术主动赋能各类城市治理应用场景，深化政务数据融合创新应用，以新一代通用人工智能技术为支撑，构建政务大模型应用生态，积极尝试 AI 大模型与民生诉求、营商环境、政务办公等场景结合，基于对大模型数据的预训练，为群众办事提供咨询引导，不断提升市民办事满意度和生活幸福感。





第十三章 政府及公共事业

▶ 13.1 国家气象中心

13.1.1 案例综述

中国气象局是国务院下属机构，负责拟定气象工作的方针政策、法律法规、发展战略和长远规划；管理全国陆地、江河湖泊及海上气象情报预报警报、短期气候预测、空间天气灾害监测预报预警、城市环境气象预报、火险气象等级预报和气候影响评价的发布等职责。随着 AI 在气象领域应用的深入，中国气象局也制定了相应的人工智能发展规划，如图 13-1 所示。

目标任务 (2023-2030年)

- ✓ 到**2025年**，确定人工智能气象应用发展路线图，**形成“542”整体框架布局**。初步建立人工智能大数据库、算力环境、算法模型、开放平台和检验评估的“五大基础”支撑；**启动气象预报大模型等新兴技术研发，开展人工智能新兴技术与监测预警、预报预测、数值预报和专业服务“四大领域”融合**；
- ✓ 到**2030年**，**人工智能气象应用发展水平进入世界前列**，业务能力建设取得重大进展。建立较为完善的人工智能气象应用“五大基础”支撑，**GPU算力规模达到30-200PFlops**；**建立与人工智能深度融合地球系统数值预报业务体系；打造数字孪生大气**；实现人工智能会话式气象服务网络机器人系统上线运行；实现人工智能技术与气象预报和服务业务系统全面融合，基本建立智慧气象服务系统；

人工智能气象应用发展路线图



图 13-1 中国气象人工智能气象应用发展规划图

国家气象中心（中央气象台），是中国气象局负责气象预报的业务部门，其职责其一是为机关提供支持保障的职能，如承担中国气象局天气预报、气象服务等有关发展规划的编制和起草；为重大活动举办提供气象保障服务，为重大灾害和重大突发事件提供应急气象保障服务等任务。其二是面向社会提供公益服务的职能，提供全国及全球所需范围内的天气监测预报服务产品，生态气象、农业气象、环境气象、水文气象、海洋气象等专业化气象监测预报服务产品等。当前阶段，气象中心主要通过传统数值预报系统的模式算法来开展短、中期气象预报工作。

2022年11月，华为发表了“盘古气象大模型，中长期气象预报精度首次超过传统数值方法，速度提升10000倍以上”的论文，引起了气象圈子专家的广泛关注。2022年12月，国家气象中心与华为盘古团队在中国气象局进行了盘古气象大模型交流，会后双方建立了工作组，

共同推动模型在气象中心开展业务验证。鉴于盘古气象大模型在台风路径预报方面能力比较突出，双方验证的重点就放在了台风路径预报。验证的时间安排在2023年汛期，4到10月份，台风高发的时间段，通过对所有场次台风路径预报准确度的统计分析来评估盘古气象大模型的有效性。AI模型在单台服务器上高效率的生成了和传统数值模式相媲美的预报结果，并和气象业务系统实现对接，提供预报员参考，并在会商中使用。

13.1.2 解决方案和价值

13.1.2.1 方案架构

在国家气象局 PoC 项目中，AI 气象预报方案高效率、低能耗实现气温、气压、风速和湿度等气象要素的高质量预报，为短、中期气象预报决策提供有力的补充。AI 气象预报方案总体架构如图 13-2 所示。

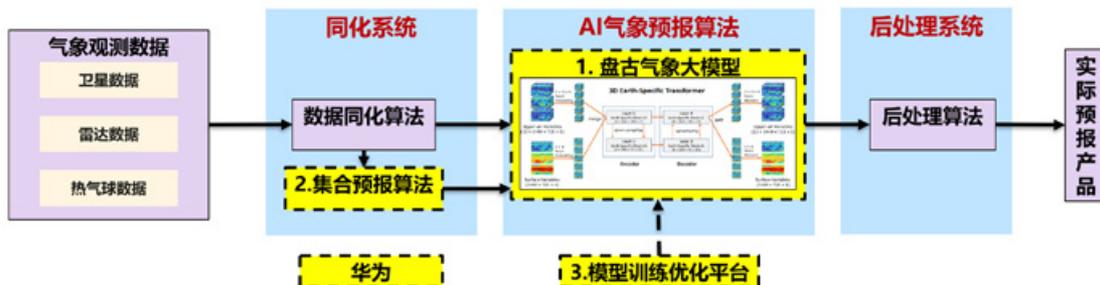


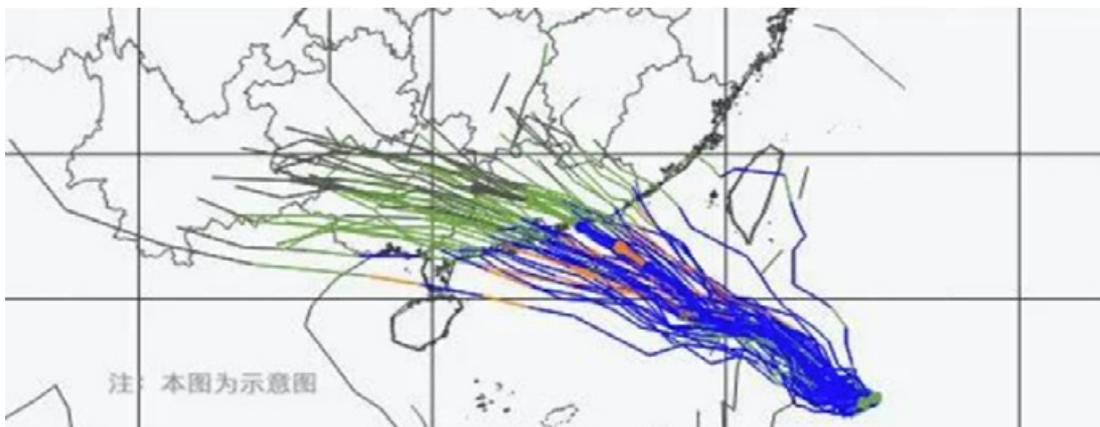
图 13-2 AI 气象预报方案

整体方案以盘古气象大模型为核心，可基于昇腾 AI 算力平台实现秒级中期预报成果。同时构建同 AI 预报大模型匹配的集合预报能力，实现上千集成员规模的集合预报，助力预报员对未来气象形势进行综合而全面的评估。为支撑气象大模型在运行过程中持续优化预报精度，以云服务的形式建立了模型训练优化平台能力。

1. 盘古气象大模型：

盘古气象大模型可用单块 AI 卡在 10 秒左右时间提供未来十天 25km 分辨率的全球地表及高空 13 个等压面气温、气压、风速和湿度等气象要素的预报信息，其技术创新点主要有两个，一个是 3D Earth-Specific Transformer 神经网络，即在每一个 transformer 模块中引入和纬度、高度相关的绝对位置编码来学习每一次空间运算的不规则分量，这样不仅更准确的学习了物理规律，而且大大加速了模型训练收敛的效率。另外一个为层次化时域聚合策略，即通过 4 个不同预报间隔的模型（间隔 1 小时、3 小时、6 小时、24 小时）的组合应用，使得预报的迭代次数最小，不仅减少了迭代误差，并且避免了由递归训练带来的训练资源消耗。这两点明显差异化于其他 AI 气象模型。

2. 基于 AI 大模型的集合预报：



注：本图为示意图

图 13-3 台风路径集合预报

集合预报，如图 13-3 所示，是天气预报业务系统的必备功能，预报员需要根据集合预报判断不同演变趋势发生的概率，并通过集合预报平均值提升预报精度。根据 AI 气象预报模型的特点，华为设计了非线性滤波和误差学习估计等多种集合预报的算法，在几十块昇腾卡算力支撑下可在一小时内完成上千集合成员的预报评估工作。此项成果待验证。

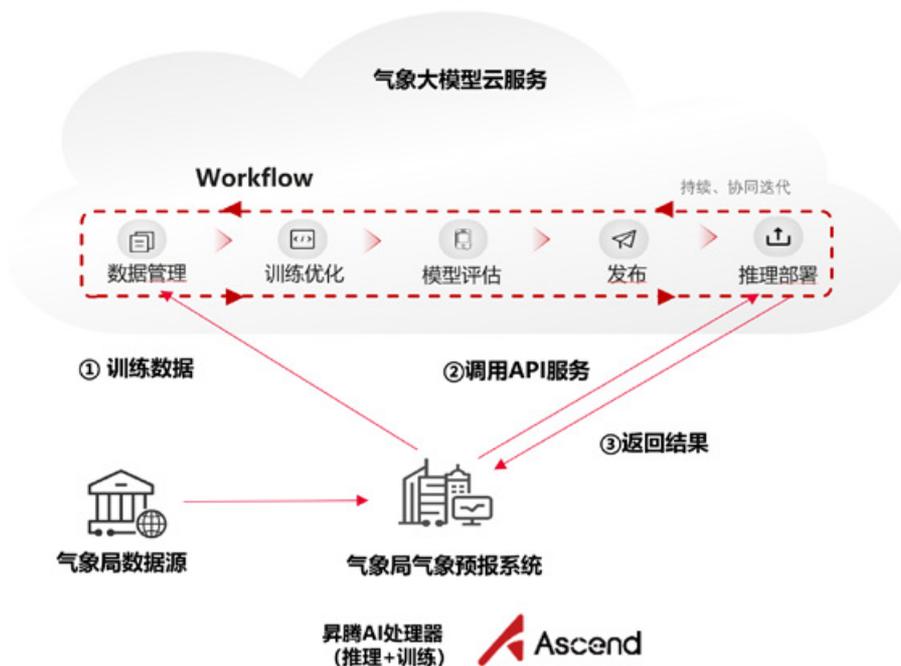


图 13-4 训练和推理一体化云服务

在前两个能力基础上，构建 AI 算力云平台，如图 13-4 所示，持续支持气象大模型的数据处理，模型框架选择，训练参数配置，模型评估以及模型更新部署等功能，实现模型在应用的同时通过新的数据持续训练，不断提高预报的精度。

3. 模型训练优化平台：

构建自主创新的国产化算力云平台，持续支持气象大模型的数据处理，模型框架选择，训练参数配置，模型评估以及模型更新部署等功能，实现模型在应用的同时通过新的数据持续训练，不断提高预报的精度。

13.1.2.2 应用场景

台风路径预报，根据预报的全球地表及高空 13 个等压面气温、气压、风速和湿度等气象要素，通过后处理计算海平面气压极小值，850hPa 旋度极值和 10 米风速等指标，可准确地预测出热带风暴和台风的路径。

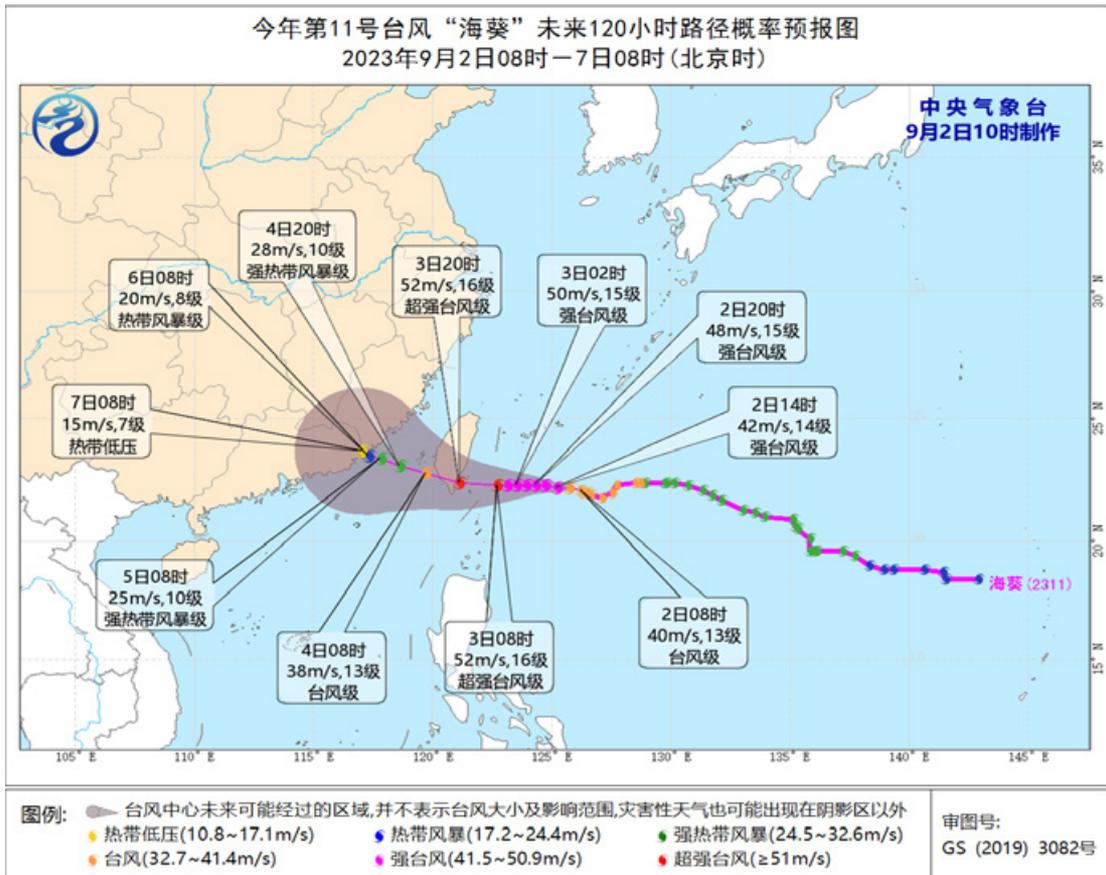


图 13-5 台风路径短、中期预报示例

国家气象中心运用 AI 气象预报方案，在单服务器算力支撑下，高效率开展了台风路径的预报验证工作，如图 13-5 示意，提供每个台风未来 1-10 天的路径预报，通过 2023 年 4 月以来所有场次台风的测试，其路径预报准确度符合预期，而路径趋势的把握相对于传统数值模式尤为稳定。随着上千集成员集合预报综合评估的应用，能帮助预报人员更全面的把握台风路径演进的各种可能。

13.1.2.3 方案价值

相对于当前传统数值天气预报遇到的计算速度慢（5-6 小时）、资源消耗大（上万 CPU 核）、预报精度提升不易等问题，AI 气象预报方案

为业务人员提供了有力的支撑，预测精度统计视角分析高于传统数值方法（欧洲气象中心的 operational IFS），同时预测速度提升 10000 倍，能够提供秒级的全球气象预报，计算资源由上万 CPU 核计算运算数小时下降到单 AI 卡运算数十秒。由于推理效率的指数级提升，使开展上千集成员的集合预报成为可能，大规模的集成员可更完整覆盖可能的天气演变，提前洞察各种可能的极端天气情况，为公众提供有效的预警信息，社会价值巨大。同时随着高质量再分析数据的持续积累，通过训练平台周期性的对 AI 模型进行迭代优化，可持续提升模型的预报准确度。

13.1.3 总结与展望

中国气象局 2023 年印发《人工智能气象应用工作方案（2023—2030 年）》。“方案”指出，人工智能已经在传统数值模式预报能力较弱的领域展现出一定的优势，通过进一步的生产实践，一定能探索出一条 AI 和传统数值预报模式协同发展，优势互补的路径。

到 2025 年，确定人工智能气象应用发展路线图，形成“542”整体框架布局。初步建立人工智能大数据库、算力环境、算法模型、开放平台和检验评估的“五大基础”支撑；启动气象预报大模型等新兴技术研发，开展人工智能新兴技术与监测预警、预报预测、数值预报和专业服务“四大领域”融合；优化人工智能创新合作和人才培养，成果转化和知识产权保护的“两大保障”环境。

▶ 13.2 301 医院

13.2.1 案例综述

301 医院创建于 1953 年，是集医疗、保健、教学、科研于一体的大型现代化综合性医院，医院的信息化建设一直走在全国前列，取得了对行业发展有重要影响的系列成果。经过多年的科研与临床知识沉淀，301 已成为国内规模最大、覆盖面最全、数据量最多、时间连续性最好的医院医疗数据资源库之一。

在医疗健康领域，人工智能技术面临医疗数据共享难、数据质量标准低、优质 AI 系统少、支撑平台资源缺等突出问题。

第一，经过多年信息化建设，国内医院普遍积累丰富的临床医疗数据，但由于隐私保护及信息安全等原因，数据资源难以出院进行有效共享，使得大量医疗 AI 企业缺少优质充足的数据，从而阻碍医疗 AI 模型算法的研发创新和推广应用。

第二，基于数据驱动的人工智能技术对数据标准、数据质量要求较高，但是多源异构的医疗

数据却在编码、格式、内容等方面存在较大差异，增大数据治理、数据标注、训练推理的难度，在一定程度上桎梏医疗健康 AI 行业的融合发展。

第三，针对常见病、多发病以及重要疾病的诊疗、防治等重大需求，缺乏高质量的样本数据、高性能的网络框架和高效率的模型算法，直接导致功能强大、性能优异的医疗健康 AI 系统非常稀缺，人工智能赋能反哺乏力。

第四，在医疗机构之间、以及医疗机构与 AI 企业之间难以有效共享算力、数据和模型，“信息孤岛”、“服务壁垒”等问题亟待解决，急需构建面向医疗健康行业，开放共享、高质高效的人工智能筛查和辅助诊断公共服务平台。

当前，以深度学习、卷积神经网络等技术为代表的新一代人工智能发展迅猛，前景广阔，有力推动人类社会从万众互联走向万众智联，从数字经济转向数智经济，从信息时代迈向智能时代。在此背景下，人工智能正在加速推进医疗健康行业的交叉融合与转型重塑。因此，建

设开源共享的新一代人工智能公共服务平台将有效破解行业痛点，加快培育智慧医疗生态，有力驱动面向医疗健康行业的人工智能产业新模式发展。

作为重要服务支撑和关键基础设施——新一代医疗健康人工智能公共服务平台，被工信部科技司纳入 2020 年“高质量发展”重大专项。301 医院牵头该项目，提出软硬系统设计、训推异构兼容、云网功能融合、存算资源优化等原创理念，通过联合攻关和协同创新，建成国内首个主要采用国产软硬件设备，面向医疗健康行业，提供集人工智能数据治理、模型训练、测试验证及推理应用等核心功能于一体的大型公共服务平台，探索总结出一套面向医疗健康 AI 系列产品研制的通用研发 workflow、团队分工协作链以及任务路径规划图等工作机制，极富中国特色、时代特征和技术特性，具有突出的研究、应用与推广价值，如图 13-6 所示。

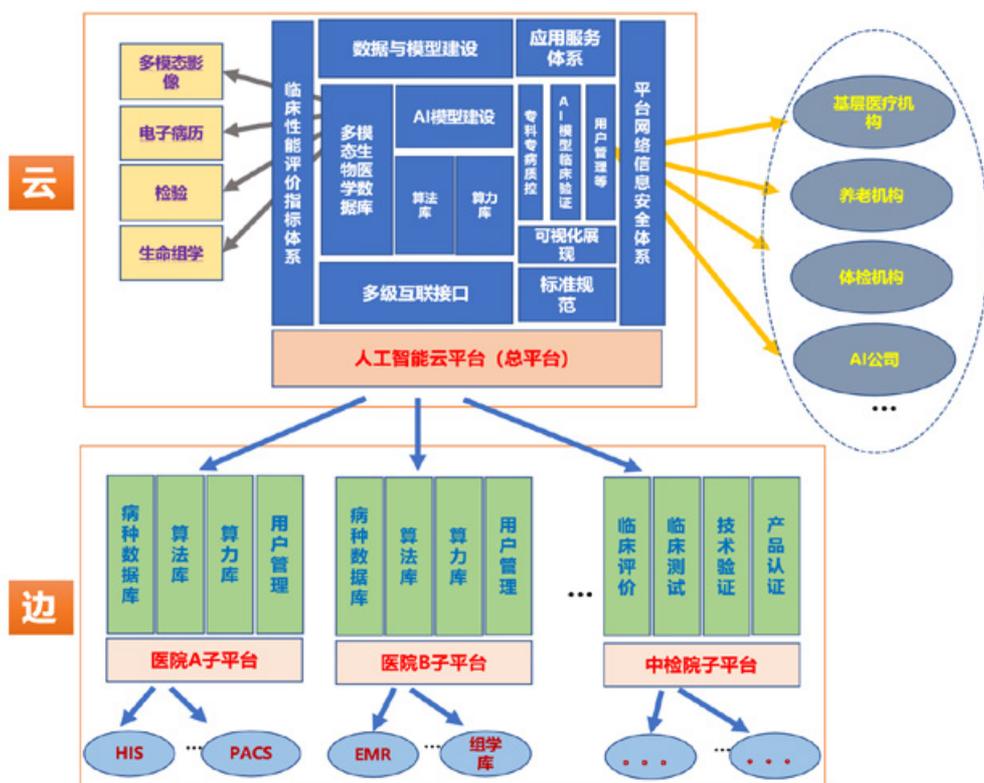


图 13-6 医疗辅助诊断公共服务平台建设路径

13.2.2 解决方案和价值

13.2.2.1 方案架构

在公共服务平台建设中，以 301 医院为牵头单位，首都医科大学附属北京天坛医院、首都医科大学附属北京同仁医院、北京大学第一医院、北京肿瘤医院、郑州大学第一附属医院、北京安德科技有限公司、华为技术有限公司、中国食品药品检定研究院、中国移动为成员单位，共同构建医疗健

康人工智能公共服务平台。平台采用云边协同架构，构建集“云、大数据、人工智能”于一体的一站式人工智能云平台，如图 13-7 所示。



图 13-7 人工智能公共服务平台总体架构

华为承接了项目构建一站式人工智能云平台的任务，支撑数据集成、数据治理、模型训练和应用部署等功能。云侧是人工智能总平台，边侧是人工智能子平台以及各基层医疗、养老、体检机构等，如图 13-8 所示。



图 13-8 一站式人工智能云平台

如上图所示，平台云侧提供医学人工智能训练和推理所需的存储、算力、基础算法，提供模型推送与联邦训练管理能力，对部署在分支的边缘节点进行管理，同时结合云端大规模的算力，满足全系统海量数据的计算、存储、服务分发等需求。边侧提供本地训练所需的存储、算力以及模型管理、模型推理能力。基于物理安全、网络安全、平台安全、数据存储安全及云边服务安全等，形成完备的网络安全体系，保障医疗大数据的信息安全及有效共享。

芯片采用神经网络处理器技术的昇腾 AI 加速处理芯片和昇腾 AI 加速处理芯片构建训练和推理的算力；提供业界单节点性能最强、功效比最优的高性能 AI 训练集群系统。

云平台采用华为云 Stack 建设方案，基于新 ICT 技术可提供对计算资源、存储资源、网络、数据库等资源的专属使用，同时提供高安全的网络隔离环境满足网络隔离要求，资源独享可以避免业务高发期资源被抢占造成的业务卡顿情况，从而满足性能、安全、可靠性、可扩展性等关键业务诉求。

医疗数据库（数据集）使用华为 DWS 和 MRS 来构建。DWS 是基于华为研发融合数据仓库 GaussDB 产品的云原生服务，兼容标准 ANSI SQL 99 和 SQL 2003，同时兼容 PostgreSQL/Oracle 数据库生态，为 PB 级海量大数据分析提供有竞争力的解决方案。大数据服务（MRS）基于 Shared-nothing 分布式架构，具备 MPP 大规模并行处理引擎，由众多拥有独立且互不共享的 CPU、内存、存储等系统资源的逻辑节点组成。在这样的系统架构中，业务数据被分散存储在多个节点上，数据分析任务被推送到数据所在位置就近执行，并行地完成大规模的数据处理工作，实现对数据处理的快速响应。

智能边缘平台部署在总平台，通过纳管子平台（边缘节点），提供将云上应用延伸到边缘的能力，联动边缘和云端的数据，提供医疗训练模型的推送和存储，同时在云端提供统一的设备 / 应用监测、日志采集等运维能力提供完整的边缘和云协同的一体化服务的边缘计算解决方案。

ModelArts 平台是面向开发者的一站式 AI 开发平台，为机器学习与深度学习提供海量数据预处理及半自动化标注、大规模分布式 Training、自动化模型生成，及端 - 边 - 云模型按需部署能力，帮助用户快速创建和部署模型，管理全周期 AI workflow。

平台围绕 5G 通信产业发展前沿，建立完整实用的云平台推广通信体系，促进 5G 通信在远程人工智能医疗共享，推广落实的应用，将为用户提供高质量的医疗服务网络。

医疗辅助诊断公共服务平台，不仅拥有独占、高质量、大规模的和完整的医疗数据，而且又有完善的医疗管理体系，承载了 10 种肺炎影像 AI 诊断、心脏核磁 AI 辅助诊断、心脏动态超声 AI 辅助诊断、乳腺疾病磁共振 AI 辅助诊断等功能，基于深度学习的人工智能影像分析技术改善了人工阅片的痛点，深度学习通过广泛的图像训练，从底层提取特征，能够实现对更加多样化的影像表现识别并不断自动优化。能够在提高影像工作效率同时，提供精准的疾病风险评估影像信息，真正满足临床需求。

13.2.2.2 应用场景

1. 临床辅助决策系统

临床辅助决策系统通过多模态数据库，采集 38 万例的心肺血管疾病住院患者主诉、现病史和检查检验等信息，利用系列算法模型进行病情评估、风险预测、治疗推荐、预后转归、共病管理、住院花费等，涵盖入院患者全周期。人工智能平台中的临床医疗辅助决策支持系统从标注、训练、测试、验证等软件，联合研发医疗 AI 芯片性能超业界 2 倍，算法网络创新，预测性能提升 23%，如图 13-9 所示。

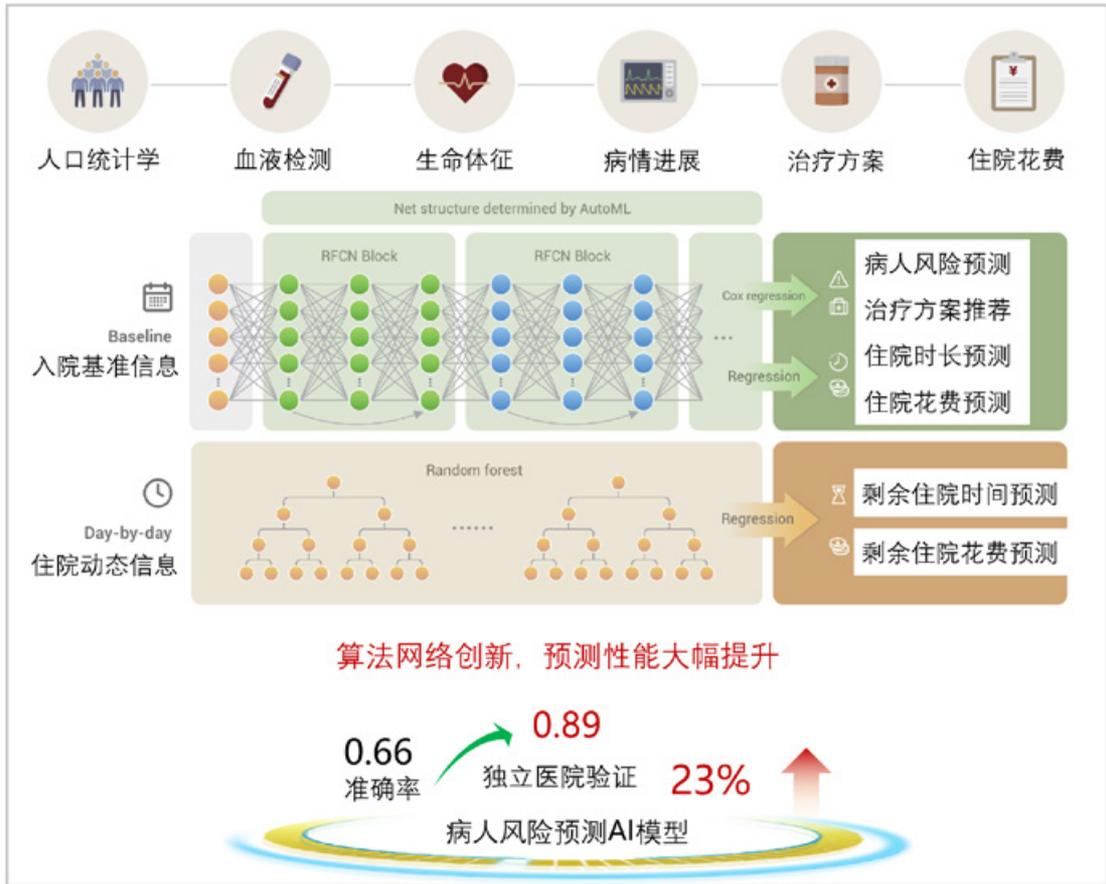


图 13-9 临床辅助决策支持系统

2. 肺炎影像诊断

新冠疫情肺炎期间，平台借助人工智能的图像分割与病灶发现，对传染性肺炎进行精准诊断、治疗方案推荐、进展评估、预后及风险预测，结合国际指南，提供标准化的临床诊疗决策支持。传染性肺炎临床辅助决策支持系统可以直观展现图像的参数分割及病灶发现，如图 13-10 所示。



图 13-10 传染性肺炎临床辅助决策支持系统

3. 心脏核磁辅助诊断

利用多序列的智能分割，分割心脏视频，自动提取结构和 16 节段功能参数；可以通过对心脏结构和节段功能参数曲线智能分类，诊断左心室心肌病。人工智能的介入，提高了诊断的准确率和效率，与人工测量和诊断相比，诊断时间大幅度缩减，诊断效率提升了 30 倍，如图 13-11 所示。

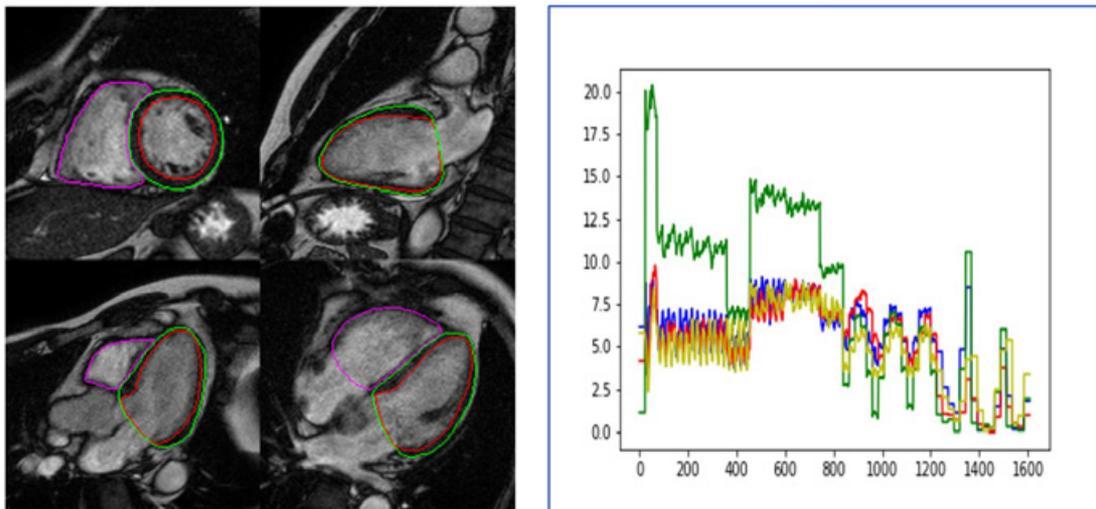


图 13-11 心脏结构和节段功能参数曲线智能分类诊断

4. 心脏动态超声辅助诊断

系统可以自动分类适配智能分类，分解 45 种超声心动图切面，包括二维动态、彩色多普勒、多普勒频谱、M 型超声等；以及 5 类 14 种心脏疾病智能辅助诊断；并通过对 61 种核心参数的自动测量，自动精准分割，数秒生成报告。

5. 乳腺疾病磁共振辅助诊断

根据乳腺疾病核磁 BI-RADS 1 到 5 级进行分级，并警示病灶的良恶性。对乳腺腺病、纤维腺瘤、乳腺炎、导管内乳头状瘤、原位癌、浸润性癌等 6 种乳腺疾病诊断。提升阅片效率，提高诊断准确率，提供大量定量分析信息。

13.2.2.3 方案价值

国家级医疗辅助诊断公共平台的建设，直接面向行业痛点，最大化提升劳动生产率和诊疗效率对于患者。辅助诊疗系统、AI 影像可以减少医生的诊断时间，得到更可靠的诊断结果，大幅提升整体诊疗水平，降低医疗机构的隐形运营成本。

平台聚焦医疗影像产业链的下游，可以有效的推进诊疗的一体化。优化诊疗流程，有效的提高医疗设备的使用效率和诊疗效率；打造影像智能会诊中心，突破地域限制，争取“空间红利”；创新服务模式，拓展疾病预防、健康管理等新型增值医疗服务领域。

平台以推动人工智能筛查和辅助诊断水平提升为核心，依托平台联合医院和标准制定单位共同制定专业、权威的标准规范认证体系，参与《人工智能医疗器械 质量要求和评价》系列标准的编制；并在平台建立模块化、体系化的人工智能筛查诊断模型，在服务医疗机构开展应用测试体系建设，形成标准化的人工智能筛查诊断流程，完善中国人工智能筛查诊断能力提升的关键环节布局，带动中国人工智能医疗产业发展，提升在国际的影响力和竞争力。

平台推动优质医疗服务资源前移下沉，实现分级诊疗模式，减少病人的无效流动和医疗资源

的浪费。本平台为至少 80 家医联体医院和基层医疗机构提供远程人工智能筛查和辅助诊断服务可以使牵头医院的优质资源能够及时下基层，向基层延伸，进一步优化服务体系，放大优势资源效应，使得群众在家门口也能够看好病。

平台成为中国医疗服务产业协同创新发展的重要支撑力量。一方面，平台通过人工智能医学模型共享，云推广等方式，推动医疗服务行业基层医疗机构进行技术、知识、标准、知识产权等方面的学习与交流，加快产业链间的交融互惠，促进中国不同区域医疗健康产业协同发展；另一方面，平台将通过线下高峰论坛、行业峰会、沙龙等形式，集聚人工智能医疗产业发展力量，推动本土医疗机构人工智能筛查诊断协同创新发展。

13.2.3 总结与展望

建立人工智能云平台将为后续医院进行大规模的科学研究、图像分析合作、疾病辅助诊断与管理提供便利。同时更深度的集成居民健康服务系统，提供更全面的医疗健康档案和区域统计数据。通过大数据分析，为临床提供更多决策支持，如用药分析、治疗效果相关性分析、制定个性化治疗方案等，为疾病诊断与预测、临床试验设计等提供统计工具和算法。同时也提供健康数据的监测与预警、风险评估、健康干预等服务。医疗辅助诊断公共服务平台建设是现代算法运行的燃料，也是精准医学未来的基础。

现阶段在人工智能、5G、大数据等新兴技术的支撑下，新医疗技术的转化应用、基于人工智能技术等干预手段介入传统的院内外医疗环节，应用在药物研发、临床决策、影像辅助诊断等各个医疗环节，不断优化管理及临床实

践，提升患者体验。大模型的应用，必将进一步改善单个业务场景限制、碎片化严重、模型维护成本高、模型参数量小，优化训练对业务效果提升有限等问题。可以“引经据典”，打造医学领域的GPT。基于盘古大模型与千万篇医学权威科技文献全文数据，进一步训练准确、有据可依的问答系统，为医疗行业从业人员提供专业的知识体系，为全民健康提供更准确的指导与支持。影像互联互通，促进结果互认。利用人工智能、大模型等技术进行医学影像分割、目标器官检测等，消除主观断定影响，促进真正的互认共享。

面向未来的医疗信息化，智能化是发展方向，大模型是大势所趋，加强人工智能、大模型在医疗行业内的深度应用，将会大幅提升医院竞争力，打造医院的特色学科，推动全民健康，为具有中国特色的数字化、智能化的医疗健康体系建设奠定基础。

▶ 13.3 南京市水务局

13.3.1 案例综述

南京市水务局为市政府工作部门，承担全市水行政管理工作责任，经过十多年的不断努力，南京水务信息化建设取得了较大的进步，智慧水务建设进行了积极的探索。建设完成水情、雨情、工情、水质、管网液位、压力等多种类型物联网感知系统，为水务工程调度、供排水

管理等业务提供全面详实的数据支撑。2018年，南京市水务局启动南京市智慧水务顶层规划编制工作，提出通过智慧水务的建设，将智能化技术与业务工作深度融合，基本实现智能感知、智能调度、智能决策和智能服务。完成南京水务从“数字化”向“智慧化”的提升，形成与“智慧南京”相适应的水务现代化管理新格局，实现“感知广覆盖、数据一体化、服务大平台、应用智能化、安全有保障”的智慧水务体系。南京智慧水务建设规划，如图13-12所示。



图 13-12 南京智慧水务建设规划

南京市智慧水务一期项目从 2020 年的 9 月启动正式建设，是南京市智慧水务建设和数字化转型的第一阶段，重点围绕当前水务工作迫切需要解决的防汛、供排水、河湖管理等方面问题展开，充分整合和提升水务信息化建设成果，基本完成智慧水务框架体系的搭建，深化信息化技术与水务业务融合，从而来推动水务业务工作的现代化，统筹推进水务调度管理（防汛防涝）、供水管理、排水管理、河湖管理等应用系统建设。

本次项目的建设内容主要包括立体感知体系、水务大数据中心、应用支撑平台、智慧水务应用和标准规范体系等内容。构建了“一网、一图、二中心、四平台、五+N 应用”的总体架构，初步形成“平台集约化、业务协同化、决策科学化”的南京市智慧水务平台。为巩固全市治水管水成果、提升行业监管和社会服务能力提供保障。建设成果能够成为领导决策的依据、水环境治理的预警先导、水调度的方向指引、涉水管理执法的支撑和管长片长制落实的基础。南京水务一期项目架构如图 13-13 所示。



图 13-13 南京智慧水务项目架构

针对人工智能技术在水务行业的应用，华为基于 AI 技术和计算机视觉技术相结合，通过对涉水视频图像进行特征提取、学习、识别和分析，实现城市道路积水识别、河道水位水尺识别，排口晴天排水识别、水源地人员入侵、泵站前池水面漂浮物等场景的智能识别。解决了金川河、秦淮河等主流河道的常见人为污染防治治理，汛期信息掌握不全面等问题，极大的提高了河湖管理和防汛防涝等场景的效率，支持业务监管快速闭环。践行了人工智能在水务行业的实践应用，进一步提升水务业务工作的智能化水平。

13.3.2 解决方案和价值

13.3.2.1 方案架构

视频智能分析平台利用先进机器学习、深度学习和计算机视觉等技术，结合智慧水务中的视频采集能力，面向水务提供智能视频分析和视觉作业任务管理等能力，实现基于通用视频 + 智能分析来克服传统的传感器和人工在处理涉水业务上的不足，实现技术创新，AI 方案的架构设计需要考虑如下内容：

1. 构建满足智慧水务业务的 AI 模型开发、优化、管理、部署的端到端能力；
2. AI 技术发展迅速，AI 平台需要保持其技术先进性，算法、算力升级，部署模型更新演进等能力；
3. 基于 AI 模型算法训练对算力的要求，原则

上使用云上提供丰富的算力资源、先进的算法资源、成熟的 AI 能力；

4. 基于视频的 AI 推理，考虑视频对传感链路的成本因素，推理部分的部署原则上部署在边侧，平台的部署满足业界的边云架构。

华为视频 AI 智能分析架构采用当前云计算、边缘计算主流技术架构，使用边云部署方式。主要分为云端和边缘端，将计算节点部署在客户侧（边缘端），进行算法推理，智能识别，通过公有云将计算节点纳管到平台中，实现管控面在公有云上。

云端可以实现升级算法服务版本、管理智能分析任务、查询任务状态、任务调度编排；边缘端充分利用边缘部署原则，节省视频上云的带宽，降低投资成本。通过云 + 边的模式，取得 AI 能力与效益的平衡。视频智能分析 AI 解决方案架构如图 13-14 所示。

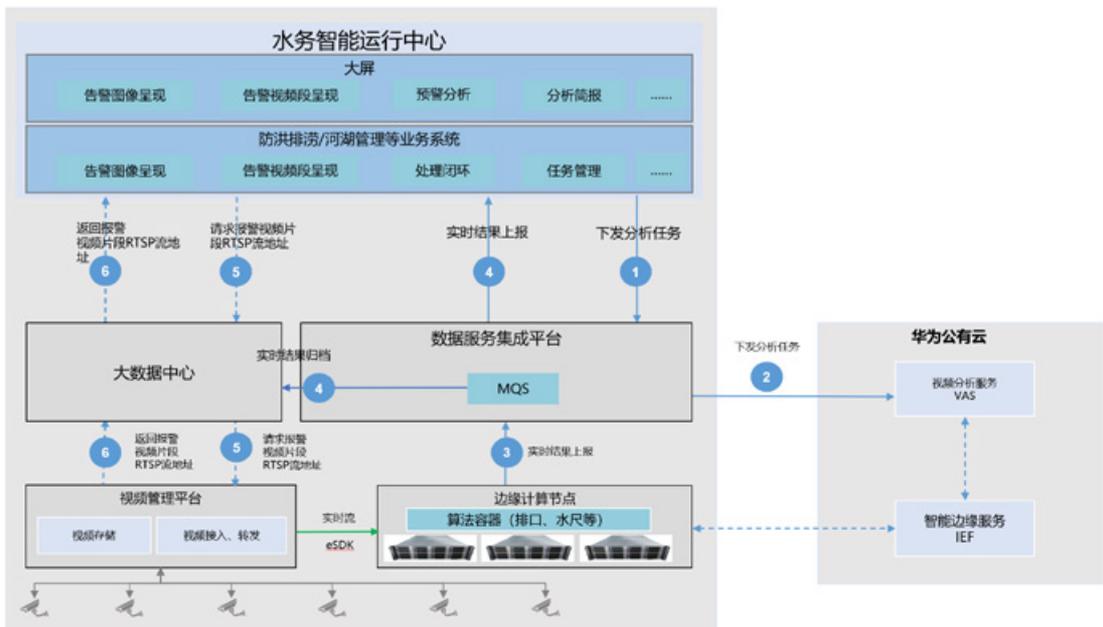


图 13-14 水务视频 AI 方案架构

业务流程如下：

1. 业务系统下发任务请求至数据服务集成平台；
2. 数据服务集成平台将该请求转发至 VAS（视频分析服务），VAS 通过 IEF 推送任务至边缘计算节点；
3. 边缘节点上视频解析算法推理将产生的告警信息发送至数据服务集成平台；
4. 数据服务集成平台将该告警信息推送至业务应用系统，同时把告警监测数据归档至大数据中心，实现监测数据全量入库操作；
5. 业务应用系统根据业务处理需要,选择向大数据中心获取告警相关信息和问题发生现场视频流地址；
6. 视频管理平台反馈现场视频流地址，业务应用系统选择现场视频流地址进行问题现场视频回看。

13.3.2.2 应用场景

一、城市道路积淹水识别

防汛期间，由于全市道路范围大，市重点积淹水位置和数量，也会随着每年积淹整治工程的实施，发生动态变化。大多数道路没有安装水尺，只能利用水务局、交通等部门建设的视频监控人工查看城市道路是否有积淹现象，这种检查汛情的方式及时性差，工作效率低。通过 AI 技术和机器视觉技术相结合，自动识别视频中存在的道路积淹水情况，实时报警，能够迅速了解全市道路积淹水情况，便于市长、水务局等领导快速做出决策，安排对应人力进行排涝，减少城市内涝。城市道路积淹水识别业务流程如图 13-15 所示。



图 13-15 道路积淹智能识别业务流程

1. 视频分析平台对视频流实时分析，智能识别城市道路内涝情况，上报业务系统；
2. 监管人员根据上报结果，结合视频就可以实时感知城市各道路积水内涝情况；
3. 根据道路积水情况有针对性的进行排涝工作调度，快速解决内涝。

二、河道水尺水位识别

接入视频监测系统河道视频流数据，通过 AI 视频识别技术，直接识别出图像中的水尺水位信息，同时可反馈直观的图像供人工二次确认。当系统探测到水位超过一定数值后，则发出告警信息。水位信息和现场图片通过实时传输网络传输，处理后存入系统数据库，水位处理模块快速抓取数据库数据，最终生成可向用户显示的每个监测点的水位的准实时信息。

河道水尺水位算法主要用于监测读取河道水尺的淹没水位，通过算法识别目前水尺的淹没程度，来输出当前淹没的水位深度。河道水尺水位识别场景如图 13-16 所示。



图 13-16 河道水尺识别场景

三、排口排水晴天识别

接入视频监测系统雨水排口视频流，通过 AI 视频识别技术，对特定排水口或溢流口在晴天时有无排水进行监测。排口流量有无检测算法，用于检测当前视频拍摄范围内的排口是否有偷排现象，如果有晴天排水，即时上报告警信息，并保留现场排水图像数据，结合排水时间、地点用于研判是否存在偷排现象，进行后续的业务处理。排水排口晴天识别场景如图 13-17 所示。



图 13-17 排口排水识别场景

13.3.2.3 方案价值

根据南京水务业务实际应用场景，通过 AI 技术和计算机视觉技术的结合，构建了城市道路积水识别、河道水位水尺识别，排口晴天排水识别、水源地人员入侵和泵站前池水面漂浮物 6 种智能化监测。实现易积水点、排口、水源地、泵站等监测视频变成了智慧的“眼睛”，让传统的视频信息变成了更有价值的水情感知信息，涉水事件信息自动上传，改变了原来靠人看、靠人巡的现状，实现预测预报更准确，监测预警更及时，运行调度和联动指挥更加科学有效，进一步提升南京水务的智慧化管理水平。

13.3.3 总结与展望

随着人工智能技术在水务核心业务场景中使用，利用水务大数据的大样本，不断的学习和训练过程，在保证智慧水务模拟和预测的准确性前提下，实现城市内涝智能预警、涉水事件自动识别、水情智能化监测等场景化应用，把工作人员中从单调、重复、机械的工作中解脱出来，有效实现水资源变化、水灾害防御主动监测，形成水、物、人全方位的监测，带来了被动监管到主动监管的转变。

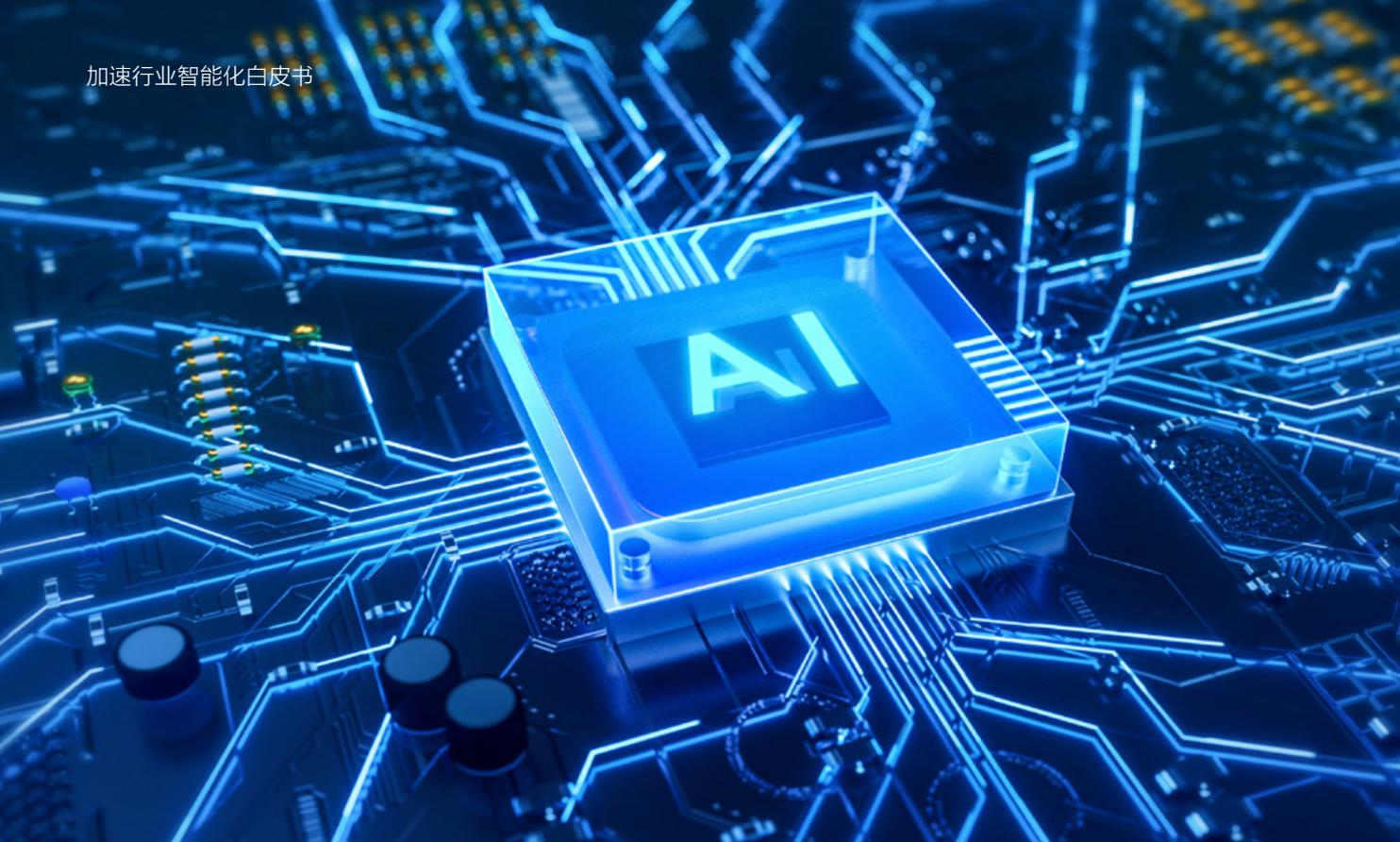
伴随着水务数据量的增长和积累、专业智能算法的不断完善，未来通过进一步推广应用，人工智能将更广泛地应用到水资源利用、水灾害防御、水污染防治、水环境治理等领域中，大大提升水务行业的智慧监管水平。





行业智能化 转型展望

第四篇



第十四章 总结与展望

人工智能在不断迭代升级演进的同时，应用范围不断扩大，正为全行业智能化转型拓展出无穷无尽的新空间，迸发出源源不断的新动能。人工智能在社会生产生活各个领域的广泛应用将激发全新体验，给人类带来翻天覆地的变化。为了更好地推动人工智能的创新应用和产业发展，需要产学研用多方参与和协同，把握好行业智能化发展趋势的前提下，不断追求技术创新，聚焦工程实践，确保人工智能安全可信，为人工智能造福人类保驾护航。

▶ 14.1 行业智能化发展趋势

大模型的强大泛化能力，算力基础设施的高效供给，使得 AI 从面向行业定制走向面向行业场景微调，大幅度降低了 AI 在行业应用的研发成本，在各个行业中得到应用。基于 AI 前沿技术研究方向和行业智能化应用实践经验进行总结，可将人工智能在行业的应用分成“减少人”、“超越人”、“辅

助人”、“理解人”和“不如人”五种模式。

“减少人”样本数据量大，验收标准相对明确，是 AI 应用的主战场，能够极大地提升作业的效率和质量，甚至做到完全无人化。主要适用于站点现场作业，财经票据识别，客户服务等应用场景。

“超越人”系统复杂，有相对清晰的整体优化方向，有大量数据、准确高效的仿真模拟器，可利用白盒优化、黑盒优化、强化学习、多目标优化等 AI 技术，超越人的能力极限。可适用于电力能源调度、港口物流调度和油气管道入侵检测等场景。

“辅助人”数据样本量小，依赖专家的经验，但人工智能可以起到关键辅助作用，极大地提高专家的效率。主要包括财经风险识别、风控内审和长途辅助驾驶等行业应用。

“理解人”数据样本量大，验收标准不明确，

最终效果取决于服务对象个体的喜好，AI 要用起来就要理解人，可能就涉及到隐私。如果不能理解人，人工智能就用不起来，需要解决“即能理解人，又不侵犯隐私”的挑战。就要利用联邦学习、知识图谱等技术，适用于推荐搜索、智能助手、情感对话等场景。

“不如人”数据样本量小，验收标准不明确，专家知识也不够，无法用数学语言进行描述。当前的 AI 技术难以介入，例如无边界的开放域对话、创意类、美学、股票预测、黑天鹅事件等。AI 算法只能“高仿”，不能创造并超越真正的艺术大师。

当前主要是“辅助人”的模式，少部分重复性的工作可以减少人，个别系统复杂有相对清晰优化目标的场景可以超越人。未来随着大模型等 AI 技术的快速发展，AI 从感知智能走向认知智能，越来越多的工作将会被 AI 取代，意味着辅助人、减少人、超越人的比例越来越高，而不如人的比例逐渐降低。

▶ 14.2 工程实践探索方向

人工智能技术和产业已经取得了长足的发展，人工智能技术创新将进一步加快，产业规模持续扩大，聚焦工程实践逐渐成为人工智能发展的重要方向。人工智能工程实践探索主要包括以下几个方向：

1) 人工智能工具链是工程实践能力核心，围绕着数据处理、模型训练、部署运营和安全监测等全生命周期的工具链要不断完善，大幅提升数据处理、模型开发效率，增强模型部署、运营监管能力，加速人工智能大规模落地应用。

2) 人工智能基础设施是人工智能的基石，为加速迈向智能社会提供强大牵引力。在全球各国多项政策和资金支持下，人工智能技术攻关和产业应用的发展势头迅猛，于此同时，人工智能基础设施建设要跟上，一方面要建设强大的计算集群，另一方面也需要高速可靠的网络让数据及时高效流动。获得先进的 AI 算力，也是每个国家、企业在发展中的重要权利。

3) 大模型轻量化后端侧部署是加速人工智能应用的必然趋势。随着大模型小型化、场景化需求增强,推理可逐步从云端扩展向端侧,用户可以更加经济、可靠、安全地使用 AI 服务,大模型端侧的应用布局 and 标准规范制定要不断加速。

4) 降低大模型训练和推理成本是当前人工智能工程实践的重要探索方向,是推动 AI 技术发展和应用的重要手段,可使得 AI 技术在更多领域得到应用,带来更大的社会效益。

5) 发展安全可信的人工智能已成为全球共识,是关系到人工智能长远发展的重要课题。未来围绕着人工智能稳定性、可解释性、隐私保护、公平性等方面的法规政策和技术研究将会形成重要的力量。例如,要建立健全的法律政策和监管机制,规范数据处理活动,保障数据安全,促进数据流通,保护个人、组织的合法权益,维护国家主权、安全和发展权益;行业用户数据具有隐私性和保密性的要求,存在自有数据不出本地问题。要加强相关技术研究创新,通过加密机制下的参数交换方式,在不违反数据隐私法规情况下,建立一个虚拟的共有模型。这样既可以不出露用户数据隐私,又可以达到高质量模型的效果。

人工智能可以提高经济效益,提升社会服务水平,更好地保障公众的安全,改善公众的生活。因此,应充分发挥人工智能技术的优势,加强人才培养、生态建设、政策法规标准规范制定,以更全面、更有效的力度推进行业智能化的发展。人工智能的未来,是全人类共同的未来,每个国家都有权利和义务参与到人工智能的发展进程中来,共同推动人工智能技术的研究创新和在千行万业的应用,带动全球经济和社会走向一个高质量、高水平的快速发展期,以造福全人类。



商标声明

 HUAWEI, HUAWEI, 是华为技术有限公司的商标或者注册商标, 在本手册中以及本手册描述的产品中, 出现的其他商标、产品名称、服务名称以及公司名称, 由其各自的所有人拥有。

免责声明

本文档可能含有预测信息, 包括但不限于有关未来的财务、运营、产品系列、新技术等信息。由于实践中存在很多不确定因素, 可能导致实际结果与预测信息有很大的差别。因此, 本文档信息仅供参考, 不构成任何要约或承诺, 华为不对您在本文档基础上做出的任何行为承担责任。华为可能不经通知修改上述信息, 恕不另行通知。

版权所有 © 华为技术有限公司 2023。保留一切权利。

非经华为技术有限公司书面同意, 任何单位和个人不得擅自摘抄、复制本手册内容的部分或全部, 并不得以任何形式传播。