视觉大模型将是下一个风口

本报记者 宋婧

计算机视觉正在迈入“大模型时代”。前不久，来自UC伯克利计算机视觉领域的三位知名学者联手推出了第一个无自然语言的纯视觉大模型(Large Vision Models)，并第一次证明了纯视觉模型本身也是可扩展的(scalability)。随后，谷歌、微软等国际大厂公布了对视觉大模型的探索，国内百度、华为、商汤、智源、中国电信、美图等多家公司也都展示了相关布局。继自然语言大模型之后，视觉大模型会成为下一个风口吗?

视觉大模型的训练难度更高

也许很多人会有疑问：在遍地都是大模型的今天，训练出一个优质的视觉大模型很难吗?随着大语言模型的爆发，不管是学术界还是业界，都开始尝试使用“文本”来扩大视觉模型的规模。以“苹果”为例，在训练时只需给模型看“苹果”的照片，并配上描述性的文字告诉模型“这是一个苹果”。然而，在面对更加复杂的图片时，就很容易忽略其中大量的信息，造成错误理解。比如，一面镜子中倒映的车辆可能会被模型误判为真实的车辆。

“相比大语言模型，视觉信息一般都是二维(图像)、三维(立体图像)或者四维(立体视频)信息，比语言信息多了1~3个维度，难度等级呈指数级增长。”一位深耕AI深度学习的专家说道。

业内人士普遍认为，视觉并非自然语言，作为基本视觉单元的像素距离高层语义更远，找不到像“单词”这样离散化、符号化的基本语义单元，因此，简单地借鉴预训练语言模型的实现方法恐难奏效。

赛迪顾问人工智能产业研究中心常务副总经理邹德宝向记者介绍说道：“视觉大模型是一种利用深度学习技术来进行图像或视频处理的算法模型。它的基本原理是基于神经网络，通过大规模的训练数据集和高性能的计算硬件，实现对图像信息的高效处理和理解。”

与语言模型相比，视觉模型的数据更难获取。对于语言模型来说，语言数据对所有数据都有一个自然的、统一的一维结构——文本流，获得大量的、多样化的大数据集相对而言是件很容易的事，所以目前市面上的语言大模型动辄就是成千上万亿的参数规模。

然而在计算机视觉领域，不同的数据源都有不同的结构，想要拥有同样规模和多样性的数据源非常困难。华为云人工智能首席科学家田奇表示，视觉模型提取特征可能是百倍、千倍的工作量，信息密度非常低。花费的成本、数据存储等开销巨大。

除了数据获取难，视觉大模型的训练框架也是一大难点。大华股份先进技术研究院院长殷俊表示，在视觉领域没有可参照的模型架构，和现在大众认知的AGI、AIGC技术方向存在差异，在CV(Computer Vision计算机视觉)领域完全靠自己搭建。

“如何处理更复杂的图像信息，获取高分辨率的图像并让机器识别图像各要素，这些技术难点不解决，很难打造出一个优质的视觉大模型。”邹德宝坦言道。

或将在工业制造场景率先落地

近年来，伴随制造业加速转型升级，智能制造已经成为人工智能技术在工业领域中最典型的应用之一。据中国电子信息产业发展研究院信息化与软件产业研究所数字经济战略研究室主任高婴劢介绍，人工智能技术能够帮助机器通过图像学习、声音识别、感知监测等方式，快速、准确地检测产品，在减少人工质检成本、提升产品质量和生产效益方面的作用显著，在手机、家电、汽车等离散制造行业质检环节中的应用渐成规模。

视觉大模型或将率先落地在工业制造领域。宁德时代首席制造官倪军曾提出“极限制造”的概念。他表示，工业领域做到6σ(每百万个产品里头有一两个不良品)远远不够，而是需要做到9σ~12σ，即对不良品的要求上升到十亿级，每十亿个产品当中，只允许出现1~3个不良品，这对机器视觉厂商是个极大的挑战。而如今有了视觉大模型的助力，“极限制造”或许有机会走进现实。

事实上，很多工业制造场景都为视觉大模型提供了落地的可能。试想一下，如果给智能网联汽车配备上视觉大模型，困扰自动驾驶多年的视觉感知问题可能会迎刃而解。具备强大的图像处理和识别能力的视觉大模型，可以更为精准地识别车辆、行人、车道线等道路元素，并处理城市道路、高速公路、雨雪天气等更加复杂的场景和环境，从而提高自动驾驶车辆在各种场景下的适应性和鲁棒性。智能网联汽车将会迎来更多可能性。

如果让工业机器人搭载上视觉大模型，它可以模仿人类视觉系统的工作原理，对视觉数据进行更为精准的处理和分析，进一步破解工业机器人操作和控制“精度”的难题，还能不断进行学习成长，让工业机器人加速走向“具身智能”，也让智能制造迈上新的台阶。

再比如在工业质检领域，产品质检涉及大量数据和复杂的图像、声音、视频等信息，要求模型能够准确地识别和分析各种缺陷和问题，甚至是极为微小的变化。用工业知识和工业数据训练出来的视觉大模型可以捕捉到产线上工艺流程和设备运行等细节，大幅提升工业质检的效率与水平。

不过，现阶段来看，视觉大模型在工业领域的应用仍处于早期阶段，落地还面临很多挑战。例如，视觉大模型的训练和部署需要大量的数据和计算资源，而这在某些工业环境中可能难以满足。工业数据敏感性高、特殊性强，对大量实时数据、多种类型数据、异常数据的处理和分析，是一项异常复杂的工作，同时数据安全保护也是一大难题。此外，视觉大模型的准确性、稳定性和可靠性也需要在实际应用中不断验证和优化。

视觉大模型要追求“大一统”?

“在视觉模型设计领域中，追求大和统一，已经成为当下公认的重要研究方向。”旷视研究院主任研究员张祥雨表示。在他看来，构建大而通用的模型的好处在于“大反而简单”。

所谓的“大”，不仅指模型更深（越深的神经网络具有越多的参数、越强的非线性，可以建模更加复杂的任务），更加强调利用更多的数据和算力帮助人们解决通用问题，期望用统一的模型解决多个模态、多个任务。

除了“大”，“统一”也是一个研究方向。追求通用框架的意义在于可以大规模地从数据中学习知识，无须针对每个任务单独设计一套系统，避免引入大量人工，可扩展性强。从宏观的角度来说，想要解决通用人工智能问题，首先需要实现模型的规模化扩展。

“尽管‘大’是未来模型发展的趋势，但我们并不片面地追求更大的参数量和计算量。同样，我们也不应该简单追求形式上的‘统一’，更应关注统一带来的性能收益。我们希望设计出更强大的模型，以创新算法充分发挥大数据、大算力的威力，随着参数量、训练算法的提升在某个时间点上获得性能的质变，即‘涌现’。”张祥雨表示。

虽然视觉大模型能为机器学习带来更广泛的应用场景和更高的表现能力已成共识，但广东工业大学教授蔡念表示，这需要海量的数据支撑，企业如果想用好大模型，就要不断挖掘工业制造数据，为大模型提供足够的训练数据来源，此外还需要考虑参数和算力的问题。这些要求对于大企业而言问题不大，但巨大的成本是很多中小型企业难以负担的。

蔡念认为，此时不如考虑小模型和轻量化大模型。智能制造场景化、碎片化明显，这就需要专注于特定领域进行训练，进行不同场景的模型定制化，最终形成某一领域的通用模型。这意味着，机器视觉的理想未来，是在复杂的应用场景中打造标准化的应用方案。

工业大模型守好“安全门”

本报记者 宋婧

近日，一股“大模型热”正在延伸到工业领域。中工互联智工·工业大模型、思谋科技工业多模态大模型IndustryGPT V1.0、海尔卡奥斯工业大模型COSMO-GPT、科大讯飞羚羊工业大模型等大模型创新应用为智能制造提供了新的方向和思路。

记者在采访时了解到，工业企业利用分布式云将大模型引入生产管理流程中是未来趋势，但也会带来越来越多的安全问题。工业级大模型应用可解决生产力短缺问题，但大模型技术是一把“双刃剑”，需要以安全为前提开展应用。

工业云已成网络攻击主要对象

《2023上半年云安全态势报告》显示，由于各行业用户群体和业务需求不同，行业遭受攻击的次数呈现出明显差异，其中互联网(通用工具、技术服务、通用SaaS)、工业云、金融是遭受攻击最多的三大行业。据统计，2022年公开披露的工业信息安全事件共312起，行业分布覆盖十几个工业细分领域，制造、能源领域成为网络攻击重点目标。攻击方式包括恶意软件、分布式拒绝服务(DDos)攻击、网络钓鱼等，呈现出目标多元化、手段复杂化、影响扩大化的攻击趋势。值得关注的是，制造业已成为勒索攻击的主要目标。其中，电子制造行业遭勒索攻击最多，占比约23%。汽车制造行业成为仅次于电子制造行业的重点攻击目标，占比约13%。

此外，供应链也是网络攻击的一个重点。由于工业供应链攻击具有隐蔽性、威胁对象多、涉及维度广等特点，利用上游企业的安全薄弱环节实施攻击能达到“突破一点，伤及一片”的效果，工业供应链已成为网络攻击最佳切入点。丰田汽车就曾因主要零部件供应商遭受网络攻击，导致其不得不关闭在日本的全部14家工厂和28条生产线，影响约1万辆汽车的生产，经济损失严重。

腾讯安全副总经理董文辉表示：“上云是必然趋势，我们看到越来越多的工业企业是在上云的。他们会选择优先把自己原本就有的服务器等利用起来，进行本地化部署。像宁德时代、三一重工等比较典型的工业企业就采用的是本地云、专有云。”在他看来，这些工业企业利用分布式云将AI的能力逐渐引入到整个生产、管理流程中，这是一个数字化的过程，也是未来的趋势。不过也正因如此，他们面临的安全威胁也越来越多了。

“一方面，传统工业企业在专业技术人才储备方面存在短板。另一方面，工业制造业涉及的特殊工艺、技术以及工业知识，给攻防端都带来了很大的技术挑战。”董文辉说。

大模型与安全行业结合成为新趋势

在工业领域，制造企业的安全体系面临挑战。蔚来汽车信息安全基础设施负责人马磊表示，对于汽车制造业而言，传统的云上风险持续存在，而智能网联时代，整个汽车生产制造，包括服务都跑在云上，车载应用迭代和更新的速度不断加快，这样也会加大安全漏洞的风险。此外，汽车制造业还面临关键基础设施的合规风险及供应链风险，必须从被动防御转化成主动。

为了应对大模型时代的云安全挑战，腾讯云在通用模型基础上投喂安全知识语料库二次训练出了一个安全大模型，并基于安全大模型打造了一款腾讯云AI安全助手。不只是腾讯，微软Security Copilot、阿里云安全大模型、360安全大模型、深信服安全GPT、安恒科技AI恒脑、奇安信Q-GPT安全机器人等产品的涌现，足以说明大模型与安全行业的结合已经成为新趋势。

据奇安信集团总裁吴云坤介绍，奇安信Q-GPT安全大模型已经在安全运营、事件响应、攻防演练等场景中进行大量实践验证，大幅度提升了相关应用场景的安全能力和效率。

“安全大模型的不断进化将驱动安全运营从人工时代进入全自动运营时代。”腾讯云安全产品负责人周荃表示。不过从现阶段来看，工业领域安全交互产品本质上仍停留在信息交换层面。很多工业设计包括产品体验设计更多都是提高信息反馈的及时性、准确性、丰富度，帮助安全运营人员拿到信息以后能够快速准确地做出决策和判断。下一步，如果能把安全大模型应用在产品交互、安全能力建设、安全服务等各个方面，将大大提升安全运营效率。

工业大模型要遵循安全原则

“工业级大模型应用可以解决生产力短缺的问题，但大模型技术是一把‘双刃剑’，需要以大模型应用安全作为生产力输出的前提和基础。然而，当前大多数大模型应用无法真正解决网络安全生产力问题。”吴云坤表示。他建议，安全大模型达到工业级应用，需要满足三个关键条件。一是需要高质量知识数据、专家队伍、实战经验和场景支撑；二是必须基于多种安全任务的强化学习和顶尖专家的反馈训练；三是要面向安全生产场景中的任务和应用强化实战能力。

制造业是一个高度专业化的领域，且细分行业众多，不能简单地将一个通用大模型应用于生产环境。即便是引领技术潮流的GPT-4，也由于缺少专业知识、无法理解工业场景的具体需求，而无法准确回答制造行业的相关问题、识别常见的工业缺陷。在工业安全场景中，更是需要有全面的工业知识和安全知识的结合，这对安全大模型本身也提出了更高的要求。腾讯安全科恩实验室高级安全研究员唐祺壹表示：“要跟实际的业务场景去做结合，通过场景数据微调的方式，以更低的训练适配成本和插件化的灵活性，将大模型应用到不同的场景当中，发挥其真正的价值。”

360创始人周鸿祎认为，大模型很热，但绝不是风口和泡沫，而是代表了一次工业革命的机会，将大幅提高生产力和生产效率。在构建企业级大模型的时候一定要遵循“安全原则”。安全是大模型发展的底线，也是未来的核心竞争力。

激发数字“生产力” 工业互联网企业为新型工业化写下生动注脚

本报记者 齐旭

电子看板监测显示车间运行情况，实时数据为生产线任务派发和维护排期提供依据，“黑灯工厂”24小时不间歇运转……当下，这样的生产场景越来越多地在工业企业里出现。

工业互联网是数字经济和实体经济深度融合的关键底座。工信部最新数据显示，我国已累计建设数字化车间和智能工厂近万家，工业互联网已融入45个国民经济大类，连接设备近9000万台(套)，对经济社会发展的带动作用日益加大。

近日召开的全国工业和信息化工作会提出，加快改造提升传统产业。实施制造业技术改造升级工程，加快钢铁、有色、轻工等重点行业改造升级。推动制造业“智改数转网联”……作为行业转型升级的数字“底座”，工业互联网企业不断推动产品升级，激发数字“生产力”，为新型工业化写下生动注脚。

数据已成企业核心竞争力

在广汽本田增城工厂里，无人物流小车“长了眼”，生产设备“会说话”……在这些背后，“数据大脑”是重要支撑，透过广泛连接并实时收集的各项数据，工作人员坐在管控中心内，通过平台终端就可以对生产状况“了如指掌”。

“数据已成为企业的核心竞争力。”中国工程院院士谭建荣表示，制造过程中的产品设计、加工、装配、销售、使用、维护等环节会产生大量数据，谁能更深入地挖掘和利用好数据资源，谁就能在数字化起跑线上占得先机。

目前，许多工业互联网企业均提出了各自的解决方案，帮助企业“唤醒”生产现场的各项数据，实现制造环节的数据透明可视，运营管理环节的精益化、智能化。

在树根互联股份有限公司副总裁吉海星看来，要让数据成为生产力，首先应按照第一性原理，梳理生产制造环节中的核心指标。再通过流程信息化系统和现场采集的数据，以及其他非结构化数据等多类型数据，以工业知识模型分析为支撑，开展核心业绩指标的智能化追踪和分析。此后，让企业从发现问题、解决问题中实现实时闭环改善。

树根互联的这一方法论，让一众企业尝到了甜头。比如，三一重工依托树根互联工业互联网操作系统“云端+终端”等，构建起了覆盖全球的智能服务平台和网络；通过数据分析和智能算法，实现设备健康管理、操作模式优化和工程施工信息服务等增值服务。

寻求标准化与个性化的平衡

近日发布的《中国工业互联网产业经济发展白皮书(2023年)》显示，2023年我国工业互联网核心产业将达到1.35万亿元，融合应用不断向纵深拓展。业内专家表示，当前我国工业互联网在各行业融合应用范围与深度不一，需要通过实践进一步解决标准化方案与个性化需求之间的矛盾。

源于传统制造企业基因，行业中已涌现出家电制造业的海尔卡奥斯、钢铁行业的宝信软件等一批重点深耕某些行业的工业互联网平台。还有一类平台，深谙制造业底层逻辑，对行业数字化转型需求充分理解的同时，在技术上保持创新。

树根互联的“根云”平台就是其中之一。据了解，现阶段，树根互联重点面向汽车及零部件、装备制造、钢铁建材等行业，整合企业的共性需求，打造出通用的工业互联网操作系统，构建起丰富的应用生态；并在各行业重点选择有实力的合作伙伴，以推出组合产品等方式，为客户提供咨询到实施一体化的端到端数字化服务等；同时，根据不同行业、不同场景的个性问题，提出有针对性的解决方案。

现阶段树根互联基于前些年根云平台在不同产业和企业数智化转型领域的深耕，业务重点持续迭代，正是平台能力积累到一定程度公司战略调整的新成果。据吉海星透露，树根互联今年积极实现产品升级，推出五大标准产品和三大解决方案——根云—工业连接平台、根云—工业智慧管控、根云—移动物联网终端、根云—数据智能、根云IIOT平台和从咨询规划到落地申报的端到端灯塔工厂整体解决方案、以产品智能化为核心的智能运营解决方案、针对工业企业尤其是中小企业数字化转型核心痛点的开箱即用标准化产品工具箱，致力于成为行业转型升级的坚实“底座”。

树根互联的解决方案有助于企业更好地进行更深度的、更彻底的数据贯通和能力建设，整合新兴数字技术，提升生产和运营管理效率。接下来，树根互联凭借丰富的数智化赋能实践和对工业的深刻理解，持续引领企业激活数据价值落地更多创新应用。

中小企业亟需“即插即用”的解决方案

量大面广的中小企业是制造业的生力军，然而这一群体由于IT基础相对薄弱，且资金压力较大，数字化转型面临重重挑战。

在业内专家看来，中小企业数字化转型，需要走“企业服务企业”的路子，降低数字化转型门槛，给中小企业“端到端”的应用，让解决方案“即插即用”。

中小企业如何在无须“补课式”信息化建设的情况下，以更低成本调用工业互联网平台共性能力?

树根互联给出了“拎包入住”的数字化方案，让数字化水平还处于初期阶段的企业接入轻量化的MES、设备管理等SaaS化服务。面对国内某领先汽车制造商高价值的复杂生产设备“压力机”运维难题，该公司提供订阅式数字化服务，从单台数字化开始做起，从被动维修转变成主动预防，设备维保效率提升80%，故障发生率减少40%。

“即插即用”的方案还带动了产业集群的中小企业形成合力，解决生产环节难题。目前，依托树根互联打造的湖南省嘉禾县“工业互联网+区块链”共享铸造高质量发展创新平台，通过构建能源管理系统和产业中心可视化大屏，让中小企业数字化转型“一步到位”，拥有能耗优化、生产可视化的能力，解决了中小企业过去“不敢转、不会转”的困扰。