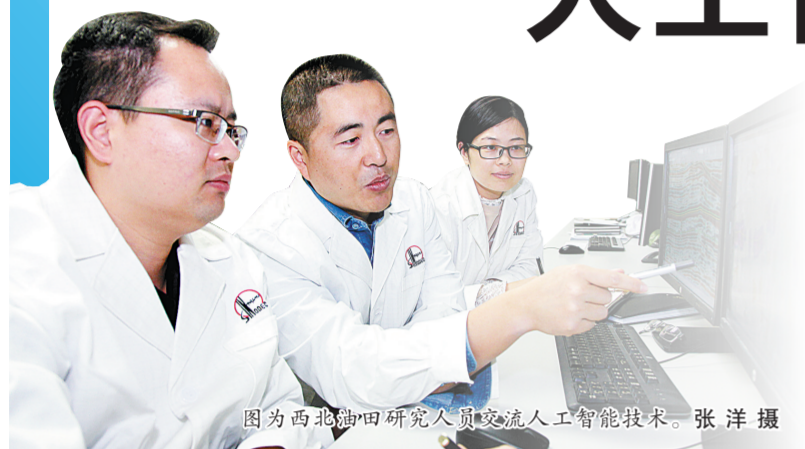


确保“十四五”开好局 以优异成绩庆祝建党100周年

利用人工智能识别复杂的断裂和断溶体,就像教小朋友识别动物,要经过“制作一套适用的识图卡片、训练小朋友识别动物、小朋友在大自然中能认识动物”三个步骤。西北油田开发的首个人工智能项目应用于顺北1号断裂带南段工区,对较大断裂和断溶体的识别已初见成效,计算效率远高于传统方法

人工智能“AI”上油气勘探



图为西北油田研究人员交流人工智能技术。张洋 摄

□本报记者 王福全 通讯员 王超

“这事儿太烧脑了,比我儿子难多了。”3月10日21时,刘军嘟囔着从工位上站起来,伸了一下腰,关上电脑,准备回家。此时的乌鲁木齐,天已黑了,空气中有春天的气息。刘军是西北油田“顺北1号断裂带南段断溶体地震智能识别及规模断溶体空间描述技术研究(2019)”项目组的负责人。一年多来,每天加班两个多小时是他和项目组同事的常态。1月初,这个项目通过了西北油田专家组验收,获得一致好评。中国石化提出打造技术先导型公司,要求加强关键核心技术攻关,聚焦“卡脖子”和短板技术,加快推进数字化转型,深化人工智能、物联网、5G等新一代创新技术应用。西北油田于2019年底设立首个人工智能项目,联合中国地质大学、石油物探技术研究院开展一体化攻关。“我们希望借助前沿技术创新突破技术瓶颈,推动油气勘探关键技术大跃升,同时以此为契机打造产学研深度融合新格局。”西北油田勘探开发研究院院长李宗杰说。目前最成熟的人工智能应用,是计算机深度学习下的自然语言处理和图像识别。西北油田首个人工智能项目要获得突破,关键要破解制作标签样本、训练样本和生产实践应用三道难题。这就像教小朋友识别动物,要经过“制作一套适用的识图卡片、训练小朋友识别动物、小朋友在大自然中能认识动物”三个步骤。研究思路很清晰,可实施起来并不容易。

01 样本标签 画出所有“动物”的各种形态

对于地震资料解释人员来说,识别剖面上的断裂、溶洞或储集体,首先要开展储层特征研究,确立各项参数,然后利用资料建立不同特征的地震地质模型,再通过地震物理模拟实验建立断裂及断溶体特征数据集。

这就像在一幅大森林的图片里寻找大象、狐狸、兔子等动物。研究人员有丰富的经验,或者能够科学套用相应“动物”模型,识别的准确度才高、速度才快。

目前,完成一块200平方千米三维地震资料的断溶体识别与描述,研究人员要花1-2个月。而运用人工智能技术,只需1周就能完成断溶体的识别和预测,在此基础上,研究人员再用1周就能完成断溶体描述。

人工智能技术能够快速准确识别各种“动物”的关键,是通过学习“识图卡片”,准确记住每种“动物”的样子。

研究人员将这种“画有动物模样的学习卡片”称为样本标签。顺北油气田地质结构复杂,地下断裂和断溶体多样。在地震剖面上准确“画”出不同断裂、断溶体在不同状态下的样子,是制作样本标签首先要解决的问题。

地震物理模拟实验是研究人员的首选。这就像制作一套等比例的楼盘模型,然后对其进行分析研究。

研究人员利用生产中获得的测井资料、岩芯数据,对整条断裂带的地层速度、密度等参数进行大数据分析,研究不同岩石的物理弹性参数变化规律,并利用钻井、测井、录井等资料开展地震地质建模和地震物理模拟,用模拟结果不断修改完善断裂模型数据体和溶洞模型数据体,再利用褶积模型和偏移成像方法对人工合成的断裂模型数据体成像,最后将成像结果转化为训练样本。

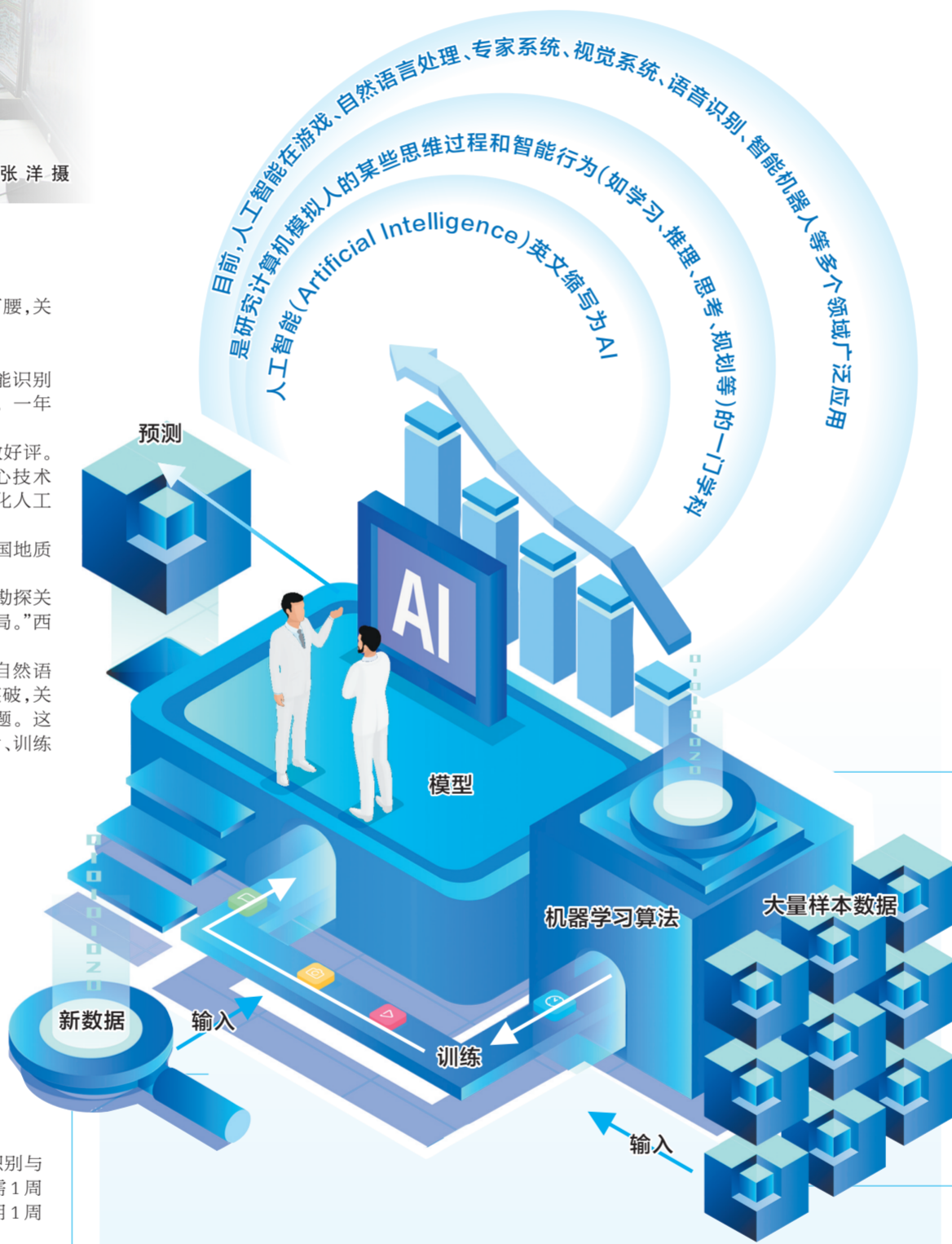
经过上述一系列专业性很强的复杂操作,研究人员就能“画”出不同地质结构在地震剖面上的样子。

在项目攻关中,研究人员累计“画”出2.7万组实际数据样本集,覆盖了各种地质结构类型。

地质结构类型通常不是一种形态存在,就像家里养的宠物狗,有时跑、有时跳、有时蹲着,有时只看到一条摇动的尾巴,如何能快速准确判断哪只狗才是自己想找的宠物?

训练数据成为人工智能识别个体的核心。研究人员用地震物理模拟合成的断裂和生产解释过的断层作为训练样本,先生成水平反射层,再做褶积、加入噪声等随机变形。就像对“狗”进行三动态全景拍照,然后用各种参数和算法,得到“狗”不同角度、不同部位的照片,形成海量的多样性样本,覆盖所有形态和部位特征。识别时,人工智能只需要把记忆中的样子与实体或部分实体进行对比,就能够快速判断“狗”的身份。

目前,在西北油田样本标签数据库中,研究人员已累计合成张扭、压隆型等断裂训练样本25万个,其他样本标签的种类和数量也在不断丰富中。



人工智能与油气那些事儿



人工智能在石油工业中的应用



02 训练识图 让“她”记住每一种“动物”的样子

有了足够多的样本标签,如何让人工智能做个“乖宝宝”,深度学习这些内容呢?

“这事儿太烧脑了,比我儿子难多了。”刘军说。

所谓深度学习,就是通过组合低层特征,形成更加抽象的高一层表示属性的类别或特征。这类似“盲人摸象”,人工智能根据海量的样本标签,分别识别“大象”的身体、牙、腿、尾巴后,系统自动组合这些特征,判断出识别的主体是“大象”,而不是墙、棍子、柱子或绳子。

研究人员想到了一个“类别”学习的方法。他们发现,断裂、断溶体和断层、溶洞在地震剖面的分布是有规律可循的,就像森林里的野兽基本都有一个脑袋、四条腿、一条尾巴。他们先把不同地质构造类型中的相同值或相似的特征值提炼出来,形成共性知识供人工智能学习,训练他们区别“动物”“植物”“岩石”等实体。

研究人员将第一个层面的训练结果传递应用到下一个层面,训练人工智能识别各个类别中每一个体的特征值,进行个体细节区分训练。这一学习方法既能节约计算机并行计算时间,又能提高人工智能的识别精度。

根据这一思路,研究人员开始构建深度学习网络。经过6个月的努力,他们构建了断裂深度残差网络、断裂U-Net识别、断溶体轮廓识别、溶洞体识别等方法及模型。深度学习网络架构搭建完成后,他们又多次优化相关参数,根据不同内容建立了快节奏学习、慢节奏学习、预习后学习等不同模式,有针对性地增强人工智能的学习效果。

随着样本标签增加,深度学习的识别能力与精准度也同步提升,误差会逐步减小。

那么,样本标签越多越好吗?并不尽然。计算机资源是有限的,深度学习样本标签需要耗费大量的计算机资源,因此要找到样本数量与学习深度之间的最佳平衡点,否则将影响工作质效,甚至事倍功半。

研究人员利用迁移学习的方法有效解决了这一问题。例如,人工智能在识别水果时,经常会把苹果的样本标签识别成梨。他们就对苹果、梨的特征相似性出发,利用迁移学习的特点,将旧领域学习过的模型应用到新领域,在认识苹果的同时,也认识了梨,而不用单独教“她”学习梨的样子。这一方法在减少目标网络数据和标签数量的同时,提升了识别的精度和可靠性。

2020年9月,在西北油田首个人工智能项目中,人工智能系统终于学有所成。

03 学以致用 “大动物”识别较好,“小动物”识别有待提升

“我们在顺北1号断裂带南段工区划出200平方千米区域作为项目研究目标靶区。一方面,顺北油藏类型特殊,地质结构复杂,储层预测难度大,地震资料刻画常常存在多解性,特别是地震波场的认识和储层精细刻画已到了技术应用的瓶颈期,常规方法难以有效突破。另一方面,该区域的油气开发已规模展开,各类资料较为完善,生产数据丰富,有利于进行成果对比。”西北油田勘探开发研究院总工程师李海英说。

目前,西北油田应用人工智能技术进行地震资料识别,已在三个方面取得进展。

第一,人工智能提高了识别效率。利用人工智能技术识别顺北1号断裂带的断裂、断溶体,计算效率远高于传统地震属性算法与追踪方法,计算时间也少。

第二,人工智能识别精度初见成效。因不受地震数据同轴能量强弱的影响,人工智能技术能够很好地发现隐藏的实体间的联系,发现肉眼和以往经验不能识别的内容。从实际应用情况看,针对较大断裂和断溶体的识别已初见成效,较大断溶体边界刻画清晰,与实际地震资料吻合度较高,但针对小断裂和小溶洞体需要进一步提高识别精度。也就是说,经过大量训练,“大动物”识别较好,“小动物”识别还有待提升。

第三,人工智能训练的模型可保存,可在相似新断裂及断溶体的识别中重复利用,修改相应参数即可深度学习。

人工智能识别结果与实际生产数据吻合程度较高。例如,在顺北1号断裂带南段,他们将人工智能识别结果和多口井进行资料比对,发现放空漏失、井漏、油气显示等特征吻合较好。

研究人员说,虽然人工智能在油气领域应用已取得一些成果,但还有很多方法、算法尚未深入研究应用。比如,反向传播、卷积神经网络方法等只是初步应用,还有一些新算法值得探索。人工智能的新方法有利于打破旧模式,新算法也有利于消除旧算法的劣势,应用更多新方法、新算法,必可实现油气勘探开发质量效益跨越式提升。