



基于地质知识蒸馏学习的油气储集层识别方法

李徵¹, 刘淇^{1*}, 王喆锋², 郑毅², 林霞³, 怀宝兴², 米兰³, 陈恩红¹

1. 中国科学技术大学大数据分析与应用安徽省重点实验室, 合肥 230027

2. 华为技术有限公司, 杭州 310051

3. 中国石油勘探开发研究院, 北京 100083

* 通信作者. E-mail: qiliuql@ustc.edu.cn

收稿日期: 2020-06-16; 修回日期: 2020-08-22; 接受日期: 2020-10-14; 网络出版日期: 2020-12-25

国家自然科学基金 (批准号: 61922073, 61672483, U1605251) 资助项目

摘要 油气储集层识别是石油能源企业在勘测和开发业务中核心的任务之一. 长期以来, 油气行业一直依靠专家人工分析海量测井数据以对地下油气储集层进行定性分析, 虽然专家解释结论有着很高的精准度, 但是时间与经济成本都十分高昂. 近些年来, 随着以深度学习为代表的人工智能技术的迅速发展, 智能油气储集层识别技术成为学术界和工业界共同关注的问题. 然而, 真实工业环境存在严重的传感数据不一致问题, 给传统的监督学习模型带来巨大的挑战. 本文针对传感器不一致情境中油气储集层识别任务展开研究, 提出多尺度地质知识蒸馏网络的方法. 首先, 该方法提出一种多尺度特征自注意力融合机制来学习地质信息的多尺度动态表征. 其次, 该方法设计一种地质知识蒸馏学习模型, 从非一致传感数据中学习额外的地质知识, 进一步提升模型准确度. 最后, 在真实数据集上进行大量实验, 结果充分证明本文提出的模型在油气储集层识别任务上的有效性和鲁棒性.

关键词 油气储集层识别, 地质知识, 蒸馏学习, 传感数据, 深度神经网络

1 引言

油气储集层的识别与预测是石油能源企业在勘测和开发业务中最核心的环节之一, 也是国内外学术界一直关注的重点研究课题^[1]. 传统的油气储集层识别方法一般首先通过测井传感器收集地下地质情况的物理、化学指标并根据对指标随着深度变化的传感数据曲线的分析构建储集层模型, 用以分析储集层的岩性、物性、电性, 以及含油气性; 然后根据岩性以及地质资料判断高潜储集层; 最后通过诸如直观判别法^[2]、交会法^[3]、多孔隙度法^[4]等分析方法对高潜储集层进行油气判定.

传统的人工分析方法需要建立岩石体积模型^[5]、测井响应方程^[6]、统计经验公式^[7,8]等较为繁杂的统计与经验模型, 同时需要大量的专家经验干预过程, 例如: 参数选取、地质类比等. 虽然专家干预

引用格式: 李徵, 刘淇, 王喆锋, 等. 基于地质知识蒸馏学习的油气储集层识别方法. 中国科学: 信息科学, 2021, 51: 40-55, doi: 10.1360/SSI-2020-0178

Li Z, Liu Q, Wang Z F, et al. Method of oil and gas reservoir detection based on geological knowledge distillation learning (in Chinese). Sci Sin Inform, 2021, 51: 40-55, doi: 10.1360/SSI-2020-0178

过程能够使得油气层识别方法获取较高的准确度同时具有一定的可解释能力,但是这种方法需要专家从海量的测井数据中逐一对比分析,解释效率十分低下.另一方面,领域专家培养成本十分高昂,这也进一步给油气能源企业带来十分高昂的勘测成本,因此,油气行业迫切需要自动化的处理方法能更加高效地、智能地完成基础解释工作,进一步将专家时间从繁杂的基础处理流程中释放出来,以节约油气勘测过程中的时间和人工成本.

事实上,随着计算机技术的不断发展,尤其是人工智能技术的突破,近些年来越来越多的研究者开始关注采用计算机技术辅助测井解释以及智能油气储集层识别任务^[9,10].真实工业场景下的智能油气储集层识别任务不同于传统的机器学习任务,其中最大的困难和挑战就是真实场景中传感数据的不一致性.产生这种不一致性的可能原因主要有两个:一是随着地球物理学学科不断发展,测井技术以及石油工程技术不断进步,使得用于测井的传感器发生了大幅变化,例如从最早的电信号传感技术到 20 世纪 70 年代的声波传感技术^[11]再到 90 年代核磁共振成像技术^[12]等,这些技术变革导致测井过程收集到的传感数据存在严重的不一致性;另一方面则是油气能源企业出于成本考虑,对不同地质条件下的测井选用不同的传感器.这种传感数据不一致性给传统基于监督学习的算法带来了巨大挑战.同时,这种传感器的不一致性并不能直接采用删除或者填充的方式解决,这是因为对于年代久远的传感器,例如声波传感,并不是由于其不可靠、不可用,而是因为声波所反映的地下岩层物理性质比较片面,加入新的如核磁共振传感数据后,记录的测井数据能够获取对地下岩层物理化学性质更多的描述,然而单独利用核磁共振技术也会造成类似的片面性;同时,由于传感器的缺失是整口井的数据缺失,因此传统的插值填充等方法并不适用.除了传感数据不一致性挑战,油气储集层识别任务中,由于地质信息的复杂易变特性,尤其是因为地壳运动导致难以预估的层断裂等,使得地下传感数据分析具有多尺度特征关联,仅仅从单一尺度上很难获取这种复杂的动态变化过程,这也对传统的机器学习模型带来了挑战.

本文针对上述真实工业场景下油气储集层智能识别问题,设计了一种基于多尺度地质知识蒸馏的油气储集层识别方法.该方法首先提出一种地质特征多尺度动态学习模型,并设计一种基于自注意力机制的多尺度特征融合模型,用来挖掘不同尺度特征之间的依赖关联并构建多尺度融合特征.接下来,针对真实工业场景存在传感数据不一致性的挑战,设计了一种基于地质特征知识蒸馏学习方法,从差异化的特征中学习额外的地质知识,使得模型获得更好的识别准确度.最后,通过在真实的工业场景数据上进行实验验证,结果充分说明本文所设计方法在传感特征差异情境下的油气储集层识别任务中的有效性和鲁棒性.

2 相关工作

在油气储集层识别的任务中,主要依赖测井的传感数据,这些数据是按照测井深度的递增而逐步采样记录的.因此,测井所收集的传感数据是随着地质勘测深度变化的序列数据,由此,油气储集层识别任务可以看作是沿地层纵深维度的序列分类问题.本节首先总结目前主流的序列分类算法,包括传统方法和深度学习方法,然后讨论机器学习算法在地质勘测以及油气储集层识别问题中的相关工作以及最新研究进展.

2.1 序列分类方法

序列分类问题一直以来就是学术界研究的重点问题,也是数据挖掘领域的被广泛应用的任务之一.传统的时间序列分类模型主要可以分为两大类,一类是基于相似性建模,另一类是特征语义建模.

其中, 基于相似性建模主要是依赖时间序列数据间的相似度量方法, 对序列类别按照相似关系进行预测. 此类方法主要包括 3 种不同的相似度量形式^[13], 例如形状相似度量^[14, 15]、时域相似度量^[16~18], 以及变化相似度量^[19~21]. 基于特征语义建模是指在保持时间序列语义不变前提下, 将高维时间序列转化成低维符号表示^[22]. 其中, 最为经典的就是词袋模型 (bag of word, BOW)^[23] 以及在此基础上衍生出来的基于特征语义模型 (bag of pattern)^[24] 以及基于符号傅里叶近似 (symbolic Fourier approximation, SFA)^[25] 的时间序列语义模型 (bag of SFA-symbols, BOSS)^[26]. 上述方法都是基于传统机器学习的序列分类模型, 虽然这些方法都被广泛验证对于时间序列的分类有着很好的可解释能力和较好的准确度, 但是, 大部分的方法过多依赖于人工特征的手动选取和构建, 在特征工程上消耗大量的时间和空间成本^[27].

为解决这些问题, 近些年来, 越来越多的时间序列分类研究开始逐渐与深度学习模型相结合. 最早研究人员就开始探讨多通道卷积神经网络在多变量时间序列中的应用. 由于深度学习模型中卷积神经网络 (convolutional neural networks, CNN)^[28, 29] 以及递归神经网络 (recurrent neural network, RNN)^[27, 30] 对序列数据特征都有着很好的非线性学习能力并能够较好地建模序列的动态演变过程, 越来越多的工作开始围绕构建混合深度学习模型来深入挖掘序列数据的潜在非线性关联^[28, 29]. 其中基于长短期记忆网络以及全卷积网络的混合模型 LSTMFCN^[27, 30] 就在多个时间序列分类数据集上被验证已超越传统的时间序列模型, 并进一步达到最佳的分类效果. 值得注意的是, 区别与传统的机器学习模型, 基于深度学习的时间序列分类模型不需要过分依赖特征工程, 能够极大程度减少特征工程的时间和空间消耗.

虽然现有的时间序列分类方法已经能够很好地针对常见时间序列进行分类, 但是目前这些深度方法都没有深入探索不同尺度时序特征的相互依赖关系对序列分类问题的影响, 同时也没有考虑到特征差异情境下的序列分类问题. 实际上, 在真实工业场景中, 这两个都是不可忽略的重要问题, 这也是本文设计模型的出发点.

2.2 地质与油气智能勘测研究

在地质与油气智能勘测领域, 一直就有计算机辅助手段帮助领域专家处理、分析地球物理手段收集到的地质数据. 早期的研究主要集中在如何构建专家知识系统用来进行简单的数据分析与建模^[31]. 但是这种方法还是主要依靠人工专家进行测井解释, 分析海量的传感数据并手动识别油气储集层. 这就导致类似方法对专家工作辅助的有限性, 未能完全将专家时间从繁杂的基础性工作中释放, 对专家的人工依赖过高, 所需要的时间成本过于高昂. 因此, 自动化的油气储集层识别技术一直以来就是学术界以及工业界探索的重要课题^[9, 10]. 近些年来, 随着计算机技术的不断发展, 尤其是以机器学习、深度学习为代表的人工智能技术, 使得越来越多的研究者开始尝试使用智能技术来对海量的地质测井数据进行分析解释. 早期针对岩层预测就有研究者尝试利用人工智能算法, 例如深度网络模型^[32]. 还有一部分工作使用将深度学习技术应用到地震反演等研究中^[33]. 最近, 还有研究人员采用循环神经网络探索智能生成地层下岩性描述^[34].

3 预备知识与问题定义

本节首先介绍油气储集层识别的预备知识, 并给出样例说明, 最后介绍地质特征差异情境下油气储集层识别问题以及本文所提出模型的基础架构.

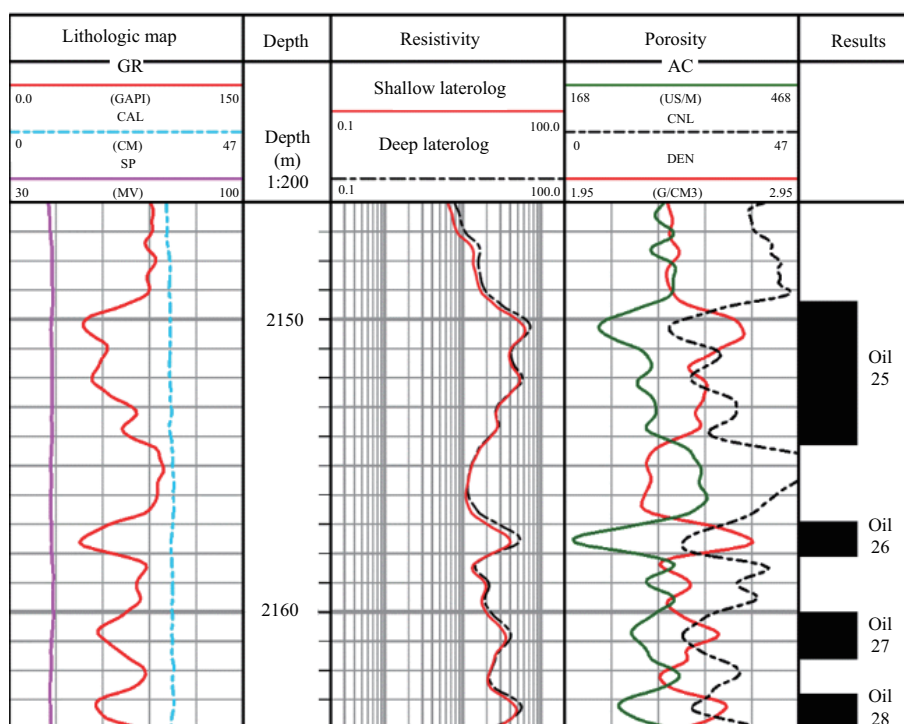


图 1 (网络版彩图) 油气储集层识别示例

Figure 1 (Color online) The sample of oil and gas reservoirs detection

3.1 预备知识

在石油勘测工业中, 油气储集层识别是一项重要任务. 考虑到真实生产井成本高昂, 通常会首先采用钻井成本较低的测井对地下的地质情况进行收集、分析以判断具有开采价值的地区和地下层位. 具体来说, 就是在测井中放入不同传感器用于记录地下岩层的物理、化学特性指标随着测井深度的变化. 测井数据记录以后, 首先根据数据的变化情况通过简易规则判断出潜在的储集层, 并对潜在储集层深度进行划分. 随后, 再通过精细化分析每一个潜在储集层的物理、化学特性, 判断该储集层的含油气水等流体的类别, 从而对油气储集层进行识别. 图 1 显示了一口测井数据以及相应的油气储集层识别结果, 图中最左边一栏是显示岩性知识的传感器数据, 左边第 2 栏显示的是深度指标, 第 3 栏是电阻率曲线, 第 4 栏指示孔隙度的传感器数据, 最后一栏是基于这些传感数据分析的储集层结果以及层位编号, 在油气领域, 这个基于测井数据油气储集层识别的过程又被称作测井解释, 因此最后一行的油气储集层识别结果又被称作解释结论. 本文研究的问题就是如何在测井传感特征非一致情境下对测井数据智能分析并得出储集层的油气解释结论.

3.2 问题定义

定义1 (测井曲线) 测井曲线是指在测井活动中, 由各种传感器采集的随着测井深度变化的指标曲线数据. 测井曲线是判断地下层位的岩性、物理性质、储集性, 以及含油性的基本依据, 也是油气储集层识别的地质特征.

定义2 (特权特征) 在真实的工业生产中, 由于采用传感器的不同, 对于智能油气储集层识别模型的训练与测试阶段会产生差异化特征. 对此, 我们定义在训练数据中所使用的差异化特征为特权特

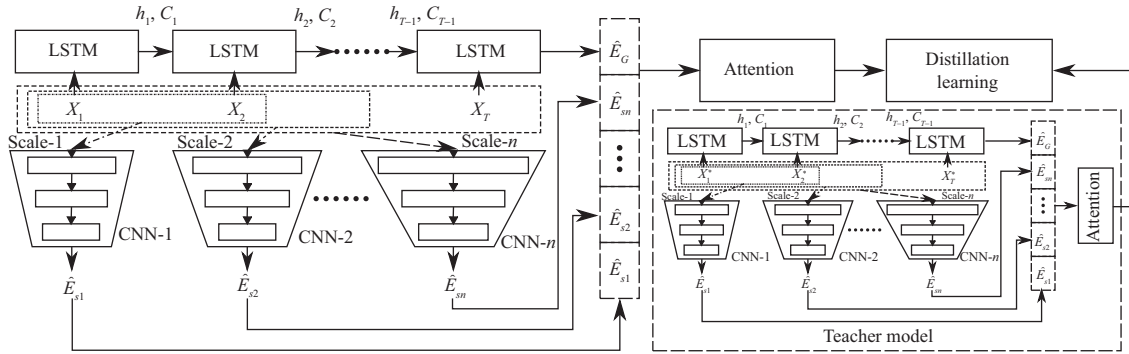


图 2 基于地质知识蒸馏的多尺度网络模型示例

Figure 2 Illustration of GKDMN model

征. 此类特权特征在测试数据中由于相应传感器的缺失而不可获取与利用.

本文所研究的是在真实工业场景中, 由于测井所装备的传感设备不一致导致的数据特征差异化情境下油气层智能识别问题. 具体而言, 就是对于训练数据样本 $X^* = X \cup X'$, 其中 X 为训练与测试样本共同的传感特征, X' 表示仅在训练数据中存在的特权特征. 本文研究问题就是如何利用训练数据与测试数据中的特权特征, 更好地提升在真实工业场景中油气储集层识别模型的性能.

值得注意的是, 本文所关注的是真实工业场景传感数据非一致问题, 不同于传统的数据部分缺失, 采用基于特征工程的插值或者均值填充在本问题中并不适用, 因为这种不一致性是一口井中特权特征数据全部缺失, 这样无法通过序列情境数据进行联想和填充. 因此, 相比较传统数据缺失问题, 这种特征不一致情境下数据挖掘问题更具有挑战性.

3.3 模型架构

针对测井曲线深度序列变化, 尤其是考虑到地质属性多尺度变化特性, 本文提出一种基于地质知识蒸馏的多尺度网络 (geological knowledge distillation multi-scale networks, GKDMN). 如图 2 所示, 模型主要包含两个部分, 具体地, 模型首先提出一种多尺度地质特征提取模型建模不同深度尺度 (scale) 下的测井曲线表征, 通过一个自注意力机制, 自动学习不同尺度之间的依赖关联, 并对尺度信息重要程度进行赋权融合. 接下来, 模型设计基于地质知识蒸馏学习 (distillation learning) 模块, 利用包含有特权特征的测井曲线样本首先训练一个老师模型, 并利用蒸馏学习方法在测井曲线差异情境中学习额外的地质特征, 进一步提升模型对于真实工业场景中油气储集层识别任务的准确性和鲁棒性.

4 多尺度地质动态知识融合模型

本节针对多尺度地质动态知识融合模型进行详细介绍.

4.1 全局动态建模

对于序列数据处理中, 全局动态信息理解是一个很重要的过程, 序列的全局信息不仅仅需要蕴含序列数据的基本属性, 更要能够反馈序列数据的全局变化特性. 对于序列数据尤其是长数据的处理, 其长程依赖关系学习一直是序列建模的关键问题 [35]. 长短期记忆网络 (long short-term memory network, LSTM) [35] 为解决经典循环神经网络对于长程依赖学习过程中的梯度消失问题, 提出非线性“门” (gate) 结构针对性地对前序信息进行学习, 并对不重要的长期信息进行选择性遗忘. 由于 LSTM

网络对序列长程依赖问题以及对长短期动态变化建模的有效性^[36], 本模型采用 LSTM 网络用来学习传感器数据在深度序列中的全局特征, 模型输入的测井指标多序列数据 $x_t = [x_t^1, x_t^2, \dots, x_t^n]$, 其中对于每一深度步都由 n 个测井传感器收集的指标数据组成, x_t^k 表示在当前 t 深度步中第 k 个传感器收集的测井指标 (例如自然伽马、井径等). 接下来, 根据下式 (1) 获取序列中包含有前序信息的隐状态:

$$\begin{aligned} i_t &= \delta(W_i x_t + W_{hi} h_{t-1} + W_{ci} c_{t-1} + \hat{b}_i), \\ f_t &= \delta(W_f x_t + W_{hf} h_{t-1} + W_{cf} c_{t-1} + \hat{b}_f), \\ c_t &= f_t c_{t-1} + i_t (W_c x_t + W_{hc} h_{t-1} + \hat{b}_c), \\ o_t &= \delta(W_o x_t + W_{ho} h_{t-1} + W_{co} c_t + \hat{b}_o), \\ h_t &= o_t \tanh(c_t), \end{aligned} \quad (1)$$

其中, i_t , f_t 和 o_t 分别表示输入门、遗忘门、输出门, 这种门结构可以有针对性的对前序输入进行学习, 并将不重要的信息进行遗忘, 从而有效地解决长序列依赖学习中可能导致的梯度消失问题. 同时, $\delta(\cdot)$ 是 sigmoid 函数, c_t 和 h_t 分别表示记忆单元和当前时序步 t 的隐状态.

随着测井曲线随着深度序列输入到 LSTM 中, 模型就能够学习地质属性随着地下深度的演变过程, 由此, 本模型用最后的时序步 T 的隐状态来表征当前深度序列下的全局动态特征 $E_G = h_T$.

4.2 多尺度局部动态建模

考虑到地下由于地质条件的复杂性, 尤其是受板块运动作用的影响, 地下的层位在局部会形成交错重叠, 其物理、化学特性指标的变化情况也会发生相应地改变, 这就导致局部深度区间中所收集的传感器数据数值上呈现剧烈变化. 对于地下油气储集层, 其所处的地层区间通常薄厚不一, 且不同深度区间下, 测井指标对单一深度尺度下地质属性具有明显的易变性, 因此局部特征建模需要综合考虑到地层深度的多尺度特性. 本文基于全卷积网络 (fully convolutional networks, FCN)^[37] 设计一种多尺度的局部特征学习模型用来对测井指标随深度局部变化的建模, 具体地, 针对每一个输入序列 $X = [x_1, x_2, \dots, x_T]$, 以 s 为深度尺度采用一维滤波器的全卷积神经网络, 具体地对 FCN 中的层 l 每一个序列步 t 有

$$E_{s,t}^{(l)} = f_s \left(\sum_{t'=1}^s \langle W_{t'}^{(l)}, E_{t+s-t'}^{(l-1)} \rangle + \hat{b}^{(l)} \right), \quad (2)$$

其中, $E_{s,t}^{(l)}$ 表示对 l 层以 s 为卷积尺度 t 步下的卷积输出, W 和 \hat{b} 分别表示权重以及偏差, $f_s(\cdot)$ 是以修正线性单元 (rectified linear unit, ReLU) 为激活函数的激活层. 通过式 (2) 模型即可获取对测井指标随井深变化的局部感受野, 将 FCN 最后一层的卷积结果 $E_s^{(L)} = [E_{(s,1)}^{(L)}, E_{(s,2)}^{(L)}, \dots, E_{(s,T)}^{(L)}]$ 输入式 (3) 的全局平均池化函数就可以生成当前深度尺度下地质属性的特征向量 \hat{E}_s :

$$\hat{E}_s = \text{GlobalAveragePooling}(E_s^{(L)}). \quad (3)$$

通过设置卷积过程中不同的深度尺度 s , 模型即可获取多尺度深度区间下地质属性的表征 $\hat{E}_{s_1}, \dots, \hat{E}_{s_n}$. 由此, 模型可获得在深度空间上测井指标多尺度动态变化的感知能力.

4.3 多尺度自注意力知识融合机制

由于地质情况复杂多变, 不同的地质环境地下的各种流体 (油、气、水等) 对测井指标在不同深度区间下动态变化的反应也有所不同. 因此, 考虑到测井曲线特征在不同的深度区间中的相互依赖关联

(例如: 自然电位与声波特征对地下岩层特性描述的共同作用等), 模型应能够自动挖掘不同深度尺度下地质特征对当前地下深度的油气水层类别预测的影响程度; 同时, 对于不同尺度下的测井指标特征变化, 模型需要综合当前全局与局部动态变化特征, 并能学习到不同的尺度下测井指标变换对当前地下深域的地质属性以及流体性质所起的关键作用. 因此, 本文设计一种多尺度自注意力融合机制来建模不同尺度之间从动态测井指标特征获取地质知识 (例如: 地下岩层的物理、化学特性) 的相互依赖关联, 进一步挖掘不同尺度特征对当前地质属性刻画的重要程度.

在前两节中, 模型可以自动从测井深度指标中获取该井地下储集层的全局和多尺度局部特征表示, 进一步, 可以用 $\hat{E}(X)$ 表示对输入测井序列 X 的全局与多尺度局部融合特征输入矩阵:

$$\hat{E}(X_i) = [h_T, \hat{E}_{s1}, \dots, \hat{E}_{sn}]. \quad (4)$$

考虑到不同尺度下, 地质属性特征之间存在的相互依赖关联, 基于多头注意力机制 (multi-heads attention mechanism) [38] 设计一种多尺度特征融合方法. 首先, 对于多尺度特征矩阵 $\hat{E}(X)$, 将其映射到多个独立空间中: $\hat{E}_i^p = W^p \hat{E}_i$, 其中, \hat{E}_i 表示特征矩阵 $\hat{E}(X)$ 中第 i 个列向量, $W^p \in \mathbb{R}^{d \times d'}$ 是将 d 维列向量映射到第 p 个 d' 维空间的映射矩阵, 通常 $d' = d/P$, P 表示映射空间总数. 接下来, 计算不同尺度的地质特征在不同独立空间中相互依赖关系:

$$\alpha_{i,j}^p = \frac{W_p^q \hat{E}_i^p \cdot (W_p^k \hat{E}_j^p)^T}{\sqrt{d'}}, \quad (5)$$

其中, $\alpha_{i,j}^p$ 表示第 i 和 j 个尺度下地质特征之间的关联注意力值, 进一步通过如式 (6) 所示的 Softmax 函数可以对上述注意力值进行归一化:

$$\hat{\alpha}_{i,j}^p = \frac{\exp(\alpha_{i,j}^p)}{\sum_{j=0}^{sn} \exp(\alpha_{i,j}^p)}. \quad (6)$$

由此, 模型获取特征组合关系注意力之后, 即可以通过式 (7) 进一步获得注意力加权的多尺度注意力融合特征, 具体地, 对于单个尺度特征 \hat{E}_i^p 在空间 p 表征 v_i^p 如下:

$$v_i^p = \sum_{j=0}^{sn} \hat{\alpha}_{i,j}^p \hat{E}_j^p. \quad (7)$$

通过注意力机制即可获取在不同空间中的多尺度融合特征矩阵 $H_p = [v_0^p, v_1^p, \dots, v_{sn}^p]$, 接下来, 通过对不同空间融合特征矩阵进行拼接, 并如式 (8) 所示, 通过一个多层感知机 (multi-layer perceptron, MLP) 网络将多空间多尺度特征进行融合:

$$G(X) = f_m(\text{Concate}(H_1, H_2, \dots, H_P)W^H), \quad (8)$$

其中 $G(X)$ 表示对输入测井序列 X 的多尺度特征自注意力融合表征, $f_m(\cdot)$ 函数表示以 ReLU 为激活函数的 MLP 网络输出, $\text{Concate}(\cdot)$ 函数表示多空间融合向量拼接函数, W^H 表示权重矩阵. 最后如式 (9) 所示, 通过一个 Softmax 函数即可计算最后的油气储集层识别类别分数 y' :

$$y' = \text{Softmax}(G(X)). \quad (9)$$

5 地质知识蒸馏学习

在上述介绍的地质特征多尺度学习模型基础上, 本节详细介绍传感数据不一致情境下地质知识蒸馏学习的具体内容. 具体而言, 首先介绍基于类别加权的优化函数设计, 随后介绍基于地质知识蒸馏优化函数改进方法.

5.1 基于类别加权的优化函数设计

对于传统的分类任务,一般最小化式 (10) 交叉熵损失函数以优化模型,获取更高的模型分类准确度.

$$L_{\text{ori}} = -\frac{1}{N} \sum_{i=1}^N \sum_{m=1}^M y_m \log(y'_m), \quad (10)$$

其中, N 和 M 分别表示训练集总样本数以及油气储集层类别数, y_m 和 y'_m 分别表示类别真实值和模型对该类别的预测值.

虽然,交叉熵函数被广泛用于分类预测任务中,但是,对于多标签分类,交叉熵损失函数未考虑到多标签任务中的类别不均衡问题,由此就会带来结果准确度上的损失^[39].同时,对于油气储集层识别任务,一方面受地质环境影响,其类别分布不均衡性十分突出,另一方面,对于不同类别因为其经济效益的不同,对其准确度的要求也不同,例如工业界对于油层、气层等具有开采价值的储集层准确度要求就会高于对于水层、干层等开采价值较低的储集层.基于此,本小节进一步提出一种基于类别加权的非均衡学习优化函数,具体如下所示:

$$\begin{aligned} S_m &= y'_m \cdot y_m + (1 - y'_m) \cdot (1 - y_m), \\ L_w &= -\frac{1}{N} \sum_{i=1}^N \sum_{m=1}^M (\alpha_m (1 - S_m)^\gamma \cdot \log(S_m)), \end{aligned} \quad (11)$$

其中, S_m 是对于类别 m 的预测模型评分, α_m 表示 m 类别的权重,默认设置为类别在训练本中所占比例, γ 是可调参数,用来平衡模型学习过程中对于样本多的类别过于激进或者惰性, γ 默认值是 2.

5.2 基于地质知识蒸馏优化函数改进

如前文所述,对真实工业场景的测井数据学习中,对传感器不一致情境下的特征进行学习十分困难,传统特征工程方法,例如直接删除、插值填充的方式是难以应用的.鉴于此,本文设计一种基于地质知识蒸馏学习的方法用来从传感器不一致情境的特权特征中学习额外的地质属性.具体地,对于测井的传感数据 $X^* = X \cup X'$, 其中, X 表示是所有测井都具有的通用传感器数据, X' 是训练测井具有特权特征^[40].值得注意的是,在讨论特征差异化问题中, X' 是测试井中所不具备的特征,因此,本文首先使用包含有特权特征的传感数据训练一个油气储集层识别任务的老师模型,对于老师模型的预测评分 y'_{Teacher} 可修改式 (9) 得到:

$$y'_{\text{Teacher}} = \text{Softmax}(G(X^*)). \quad (12)$$

进一步,通过最小化式 (11) 即可优化包含有特权特征的老师模型,从而获取包含有额外地质属性知识的油气储集层识别模型.然后,通过蒸馏学习方法^[41]如式 (13) 所示,即可获取蒸馏后蕴含特权知识的模型输出:

$$y_m^D = \text{Softmax}(G(X^*)/\text{Temp}), \quad (13)$$

其中, y_m^D 表示老师模型经过蒸馏后的预测输出, Temp 是温度参数,用来控制老师模型的输出结果的平滑程度.接下来,通过对油气储集层识别模型学习过程中引入老师模型蒸馏的知识输出即可获取额外的特征差异情境下的地质知识,从而进一步提升油气储集层预测模型的准确度.如式 (14) 所示,具体学习过程即可通过最小化包含老师模型特权知识的优化函数.

$$L_D = -\frac{1}{N} \sum_{i=1}^N \left(\sum_{m=1}^M (\alpha_m (1 - S_m)^\gamma \cdot \log(S_m)) + \lambda \sum_{m=1}^M (y_m^D \log(y'_m)) \right), \quad (14)$$

表 1 数据集统计
Table 1 Statistics of the datasets

Statistics	#9FAB2	#BF8A9
Number of total wells	299	180
Number of total samples	749209	541717
Number of reservoir classes	7	6
Number of wells in train set	239	144
Number of sensor features in train set	21	12
Number of wells in test set	60	36
Number of sensor features in test set	5	5

其中, λ 表示可调参数, 用来平衡模型从原始特征数据中学习以及从老师模型中学习的偏向权重. 最后, 模型使用初始学习率为 0.001 的 Adam 算法^[42] 优化模型参数直至模型收敛.

本文提出的 GKDMN 模型首次探索在传感数据不一致情境下的油气储集层识别任务. 考虑到油气储集层识别问题收到地下多尺度的地质属性影响, 本文提出一种基于自注意力的地质特征多尺度融合模型. 进一步地, 针对测井曲线差异化问题设计一种地质特征差异下蒸馏学习框架. 这两个点改进为本模型最重要的创新点.

6 实验评估

本节通过在真实工业场景数据集上构建实验验证所设计模型在真实工业场景油气储集层识别问题上的有效性与鲁棒性.

6.1 实验数据

本文实验数据全部来源于中国石油天然气股份有限公司的真实工业测井数据, 数据包含两个主要油田区块, #9FAB2 和 #BF8A9 区块. 其中 #9FAB2 区块总共包含 299 口测井数据, 有效储集层样本条目为 749209 条, 储集层类别有 7 类 (包含水层、油层、气层、干层、油水同层、含油水层, 以及差油层), 总共包含 21 条测井传感数据, 其中 5 条数据来自于通用传感器 (包括井径、自然伽马、自然电位、声波测井, 以及电阻率), 其他诸如中子密度等为差异化的测井曲线, 并在测试井中所不具备的测井曲线. #BF8A9 区块包含有 180 口井, 541717 条有效储集层样本, 类别有 6 类 (包含水层、油层、干层、油水同层、含油水层, 以及差油层), 有和 #9FAB2 油田区块一致的 5 条通用测井曲线以及 7 类特权特征. 具体的数据统计见表 1.

6.2 实验设置

6.2.1 实验平台

实验中, 本文所提出的 GKDMN 模型以及全部对比方法均基于 Python 3.5 以及 TensorFlow 1.10.0 实现, 模型 LSTM 以及 CNN 网络中神经元个数为 128, 模型设置 Dropout 为 0.7, Batch size 设置为 30000. 本文设置 4 个不同尺度的局部动态建模网络, 其深度尺度分别设置为 0.625, 1.25, 2, 2.5 m. 多尺度特征融合映射空间总数 P 设置为 8. 本文所有程序均在配置两块 2.10 GHz Intel Xeon E5-2620 v4 处理器以及两块 Tesla P40 的 Linux 服务器下运行.

6.2.2 对比方法

为对比验证本文所提出的 GKDMN 模型效果,共选取如下 5 种方法作为对比.

(1) 梯度提升决策树 (GBDT)^[43]. 经典的集成学习算法,输入为依照序列顺序排列的测井曲线.

(2) 长短期记忆网络 (LSTM)^[35]. 多序列输入的 LSTM 网络,输入是多传感器数据序列,输出是油气储集层类别概率.

(3) 全卷积神经网络 (FCN)^[37]. 采用一维全卷积网络的模型,该模型被证明在序列分类问题上有着很好的准确性.

(4) 长短期记忆全卷积网络 (LSTMFCN)^[27,30]. 目前最好的序列分类深度学习模型,在多个传统序列分类任务中达到最优效果.

(5) 带注意力的长短期记忆全卷积网络 (ALSTMFCN)^[27,30]. 基于 LSTMFCN 改进模型,在原模型中增加一个带有注意力机制的 LSTM 模块,这是另一个最优深度学习序列分类模型.

值得注意的是,因为这些模型的输入限制,因此都没有考虑到数据不一致情境中的特权特征,因此采用的都是基础的共同特征.同时,为进一步验证本文提出模型的每一个组件的有效性,在本文提出 GKDMN 基础上共设计 4 种变种模型.

(1) 不带注意力融合机制的基础模型 (记为 GMN-a).在不考虑特权特征情况下,没有自注意力融合机制的模型,采用直接拼接的方法合并多尺度特征.

(2) 多尺度注意力融合机制的基础模型 (记为 GMN).在不考虑特权特征的情况下,使用多尺度特征自注意力融合模型.

(3) 不带注意力融合机制的蒸馏模型 (记为 GKDMN-a).考虑特权特征,采用蒸馏学习框架下的没有自注意力融合机制的模型.

(4) 多尺度注意力融合机制的蒸馏模型 (记为 GKDMN).包含有全部模块的模型.

6.2.3 评测指标

为衡量本文所提出 GKDMN 模型以及所有对比方法在油气储集层识别任务中的有效性,本文使用精确度 (precision)、召回率 (recall),以及 $F1$ 指标来评价模型在不同区块数据集上的油气储集层识别效果,具体地,评价指标定义如下所示:

$$\begin{aligned} \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ F1 &= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \end{aligned} \quad (15)$$

其中,TP 将该油气储集层识别为正确类别的样本数,FP 为将非该类别油气储集层识别为该类别的样本数,FN 是将该类别油气储集层识别为其他类别的样本数.值得注意的是,由于是多标签分类问题,考虑到油气储集层的类被不平衡问题,在统计评测指标的时候采用的是加权平均的方法,计算各指标在各个类别下的加权平均值.

6.3 实验结果与分析

本文分别在两个不同区块数据集中进行实验,并针对模型中的学习参数进行讨论,以观察学习参数对油气储集层识别模型效果的影响,进一步验证模型的鲁棒性.

表 2 不同数据集中油气储集层识别模型性能比较^{a)}Table 2 The performance comparisons of oil and gas detection model in different datasets^{a)}

Method	#9FAB2			#BF8A9		
	Precision	Recall	F1	Precision	Recall	F1
GBDT	0.5750	0.6579	0.5872	0.7099	0.7555	0.7289
LSTM	0.5565	0.6625	0.5779	0.7426	0.7655	0.7484
FCN	0.5758	0.6700	0.5812	0.7461	0.7509	0.7483
LSTMFCN	0.6104	0.6841	0.6069	0.7277	0.7951	0.7493
ALSTMFCN	0.6155	0.5896	0.6006	0.7305	0.7924	0.7546
GMN-a	0.6126	0.6863	0.6115	0.7380	0.7965	0.7609
GMN	0.6349	0.6995	0.6394	0.7411	0.8004	0.7640
GKDMN-a	0.6330	0.6892	0.6475	0.7583	0.7870	0.7686
GKDMN	0.6486	0.7045	0.6540	0.7488	0.8086	0.7734

a) The best results are in bold.

6.3.1 对比实验分析

表 2 给出 9 种方法在不同数据集上的精确度、召回率, 以及 $F1$ 值. 从表中可以清晰地观察到本文提出的 GKDMN 模型以及其变种模型效果显著优于已有的模型方法. 整体而言, GKDMN 在两个数据集上相对于现有最佳深度学习方法 $F1$ 值分别提升 7.8% 以及 2.5%, 并且相对于传统机器学习方法 GBDT 在两个数据集的精确度指标上取得 12.8% 以及 5.5% 的提升, 在召回率指标上取得 7.1% 以及 7.0% 的提升, 在 $F1$ 指标上分别取得 11.4% 以及 6.1% 的效果提升. 同时, 所有模型在 #BF8A9 区块数据集上表现要好于 #9FAB2 区块数据集, 这是因为 #BF8A9 区块中油气储集层类别较少, 导致油气储集层识别任务相比较 #9FAB2 区块更为容易. 同时, 可以观察到本文提出模型在更复杂区块数据上的效果提升更加明显, 进一步说明 GKDMN 模型能够应对更加复杂的油气储集层识别工业场景. 进一步对比其他方法, 可以观察到深度学习模型在两个数据集上精确度、召回率, 以及 $F1$ 指标上均取得优于 GBDT 模型的效果, 这一方面是因为 GBDT 模型无法很好地建模序列数据的动态演变过程, 另一方面是在数据量充足的情况下, 深度学习模型能够更好地建模特征之间的非线性关联, 深度挖掘测井传感数据之间的特征依赖关系.

通过对比本文提出的 4 种变种模型, 可以看到在两个数据集中包含全部模块的 GKDMN 模型在绝大多数情况下取得最优的效果. 其中, 对于多尺度特征自注意力融合机制的讨论, 包含有该融合机制的基础模型对比未采用该机制的模型在两个数据集上 $F1$ 指标分别提升 4.6% 以及 0.4%, 包含有该机制的蒸馏学习模型相比较未包含的模型 $F1$ 指标分别提升 1.0% 以及 0.6%. 这说明本文提出的多尺度特征自注意力融合机制的有效性. 同时, 相比于基础模型, 采用蒸馏学习的模型在两个数据集中 $F1$ 指标分别提升 2.3% 以及 1.2%. 这说明采用知识蒸馏的模型能够很好地利用特权特征中额外蕴含的地质知识. 进一步验证在测井曲线特征差异情境中地质知识蒸馏模型的有效性.

6.3.2 参数讨论

为进一步验证模型参数的敏感性, 图 3 和 4 展示了 GKDMN 模型在两个数据集中不同的可调参数 λ 以及蒸馏学习温度参数 Temp 设置时, 油气储集层识别性能, 其中选取 4 组温度参数分别是 10, 30, 50 和 100 以及 10 组蒸馏学习权重参数. 从图 3 和 4 中可以清晰观察到 GKDMN 在不同参数设置

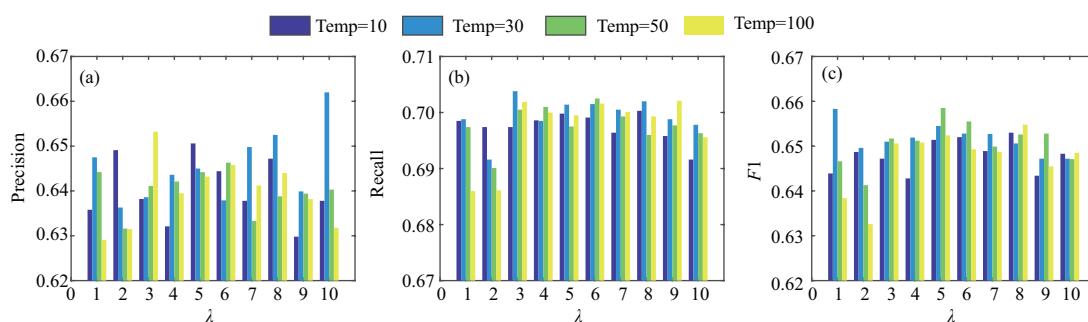


图 3 (网络版彩图) #9FAB2 区块数据中不同的参数下模型油气储集层识别实验结果 (横轴表示可调节参数 λ 的值, 纵轴表示各指标的值)

Figure 3 (Color online) The model experimental results of the oil and gas detection in the different parameters on #9FAB2 block dataset (the horizontal axis is the values of parameter λ , and the vertical axis indicates the matrix scores). (a) Precision; (b) recall; (c) F1-score

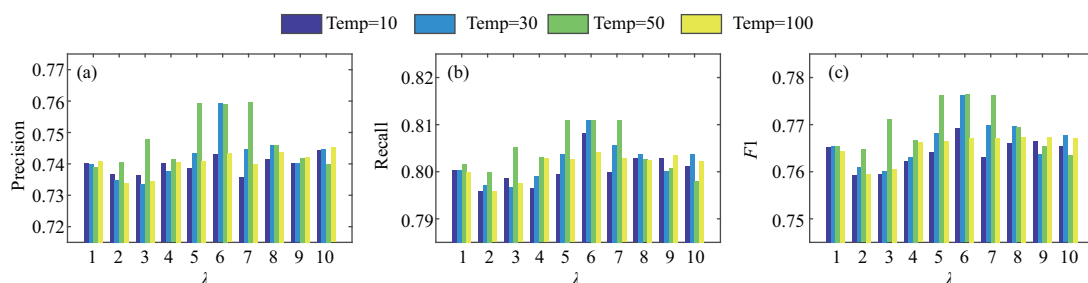


图 4 (网络版彩图) #BF8A9 区块数据中不同的参数下模型油气储集层识别实验结果 (横轴表示可调节参数 λ 的值, 纵轴表示各指标的值)

Figure 4 (Color online) The model experimental results of the oil and gas detection in the different parameters on #BF8A9 block dataset (the horizontal axis is the values of parameter λ , and the vertical axis indicates the matrix scores). (a) Precision; (b) recall; (c) F1-score

的情况下就有较好的鲁棒性, 同时, 进一步验证在传感特征差异下知识蒸馏模型的有效性. 对于两个区块数据集, 模型绝大部分情况在蒸馏参数为 5 到 7 之间的时候表现最佳. 图 3 显示的是 #9FAB2 区块数据中不同的参数下模型油气储集层识别实验结果, 可以观察到模型在蒸馏参数为 5, 蒸馏温度为 50 的时候获得较为均衡的性能表现, 在 $F1$ 指标上取得最好的结果. 图 4 中实验结果表明在 #BF8A9 区块数据中最优的蒸馏温度是 50. 这些观察结果进一步表明面对不同的数据集以及不同的地质特征差异情境, 需要选取不同的蒸馏模型参数以从特权测井曲线中获取对最佳的地质知识蒸馏效果. 同时, 进一步对比在不同的参数设定下模型油气储集层识别结果的变化, 可以发现对于不同的蒸馏温度的设定, 需要进一步优化蒸馏学习的损失函数参数, 也就是说对应不同的蒸馏参数需要调节不一样的损失函数权重, 才能够获得更好的蒸馏效果.

6.3.3 地质知识蒸馏分析

图 5 展示的是不同训练轮次下模型油气储集层识别实验结果. 图中可以清晰观察到相较于不加地质知识蒸馏的模型 GMN, 带有地质知识蒸馏模块的 GKDMN 虽然收敛的速度更慢但收敛效果更优. 在训练轮数少的情况下, 由于 GMN 直接从原始油气层标签中学习, 因此会收敛的更快, 因此在结果上也会优于还没有完全训练好的蒸馏模型. 此外, 由于 GKDMN 模型会从带有特权特征知识的老师模型出学习额外的特征知识, 因此, 在训练过程中需要平衡老师模型知识以及原始标签的监督信号, 所

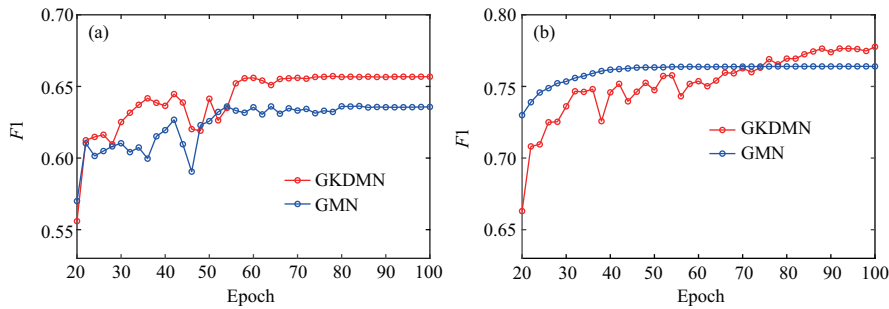


图 5 (网络版彩图) 不同训练轮次下模型油气储集层识别实验结果. (a) #9FAB2 区块数据中 $F1$ 指标; (b) #BF8A9 区块数据中 $F1$ 指标 (横轴表示训练轮次, 纵轴表示 $F1$ 指标的值)

Figure 5 (Color online) The model experimental results of the oil and gas detection over the different epochs. (a) $F1$ scores of #9FAB2 block; (b) $F1$ scores of #BF8A9 block (the horizontal axis is the number of epochs and the vertical axis indicates the $F1$ scores)

以收敛更慢. 而随着训练轮次的增加, 带有知识蒸馏模块的 GKDMN 模型性能逐渐超越 GMN 模型, 并在完全收敛的时候取得明显的效果提升. 这也进一步说明对于特权特征进行地质知识蒸馏学习的必要性以及有效性.

7 总结

本文针对真实工业场景中传感特征差异情境下的油气储集层智能识别问题进行研究, 首先介绍了油气储集层识别任务的研究背景与挑战. 然后回顾了基于传统机器学习和深度学习模型的序列分类方法, 并对当前人工智能技术在地质与油气勘测领域的应用进行了简要介绍. 针对真实工业场景地质情况复杂多变, 传感数据不一致性等特点, 本文提出了一种基于地质知识蒸馏多尺度神经网络的油气储集层识别方法. 首先, 该方法设计了一种多尺度测井特征提取模型, 可以实现对全局和多个尺度序列动态特征的感知与学习. 其次, 设计了一种基于自注意力机制的多尺度特征融合机制, 进一步挖掘多尺度特征之间的相互关联. 然后, 通过对地质特权特征的知识蒸馏学习, 加强了模型在传感特征差异情境的油气储集层识别的准确性. 最后, 在大规模的真实工业数据上进行了大量实验, 结果表明本文所提出的基于地质知识蒸馏学习的多尺度网络方法在多个评价指标下取得了最佳的预测性能, 充分证明了该方法在传感特征差异情境的油气储集层识别任务中的有效性与鲁棒性.

本文深入探索油气储集层识别问题在传感数据不一致情境下的挑战, 并建模多尺度多变量测井曲线对油气储集层识别的影响, 同时利用知识蒸馏学习方法进一步提升模型对测井曲线的特权信息的学习, 提升模型的准确性, 为智能油气储集层识别提供一种新思路. 未来的研究可以从以下几个方面展开: (1) 考虑到在油气层识别任务中, 特权特征与共同特征都是反应地下岩层的物理、化学特性, 因此, 两类特征是存在着相互的关联以及共同反应同一类地质知识的, 未来, 我们计划进一步探索在这种特征差异下的模型特征关联学习问题; (2) 在油气勘测领域, 专家经验知识对于油气层识别任务起到重要作用, 下一步我们也将继续结合领域专家知识信息, 提升模型在真实工业应用场景中的可解释能力.

参考文献

- 1 Lin N T, Zhang D, Zhang K, et al. Predicting distribution of hydrocarbon reservoirs with seismic data based on learning of the small-sample convolution neural network. *Chinese J Geophys*, 2018, 61: 4110–4125 [林年添, 张栋, 张凯, 等. 地震油气储层的小样本卷积神经网络学习与预测. *地球物理学报*, 2018, 61: 4110–4125]

- 2 Zhang Y B, Liu M Y, Xiang J L. Discriminating method of the lateral sealing oil and gas by the transporting fault in hydrocarbon accumulation period and its application. *Petrol Geol Oilfield Dev Daqing*, 2019, 38: 33–40 [张永波, 刘明杨, 相加林. 油气成藏期输导断裂侧向封闭油气的判别方法及其应用. *大庆石油地质与开发*, 2019, 38: 33–40]
- 3 Zhang Y Q. Application analysis of log data intersection graph method in volcanic rock lithology identification. *West-China Explor Eng*, 2019, 31: 53–54 [张晏奇. 测井资料交会图法在火山岩岩性识别中的应用探讨. *西部探矿工程*, 2019, 31: 53–54]
- 4 Li X Y, Qiao H W, Zhang J K, et al. Linear relationship between mineral content and porosity of Chang 6 reservoir in Jiyuan area, Ordos Basin. *Lithologic Reservoirs*, 2019, 31: 66–74 [李小燕, 乔华伟, 张建魁, 等. 鄂尔多斯盆地姬塬地区长6储层矿物含量与孔隙度的线性关系. *岩性油气藏*, 2019, 31: 66–74]
- 5 Neiman I B. Volume models of the action of cylindrical-charge explosion in rock. *Sov Min Sci*, 1987, 22: 44–52
- 6 Zhu Y, Xie J Z, Yang W H, et al. Method for improving history matching precision of reservoir numerical simulation. *Pet Explor Dev*, 2008, 35: 225–229 [朱焱, 谢进庄, 杨为华, 等. 提高油藏数值模拟历史拟合精度的方法. *石油勘探与开发*, 2008, 35: 225–229]
- 7 Liu Z D. Studying on inversion method and application of logging while drilling response. Dissertation for Ph.D. Degree. Chengdu: Southwest Petroleum University. 2006 [刘之的. 随钻测井响应反演方法及应用研究. 博士学位论文. 成都: 西南石油大学, 2006]
- 8 Mao Z Q, Li J F. Method and models for productivity prediction of hydrocarbon reservoirs. *Acta Petrolei Sin*, 2000, 21: 58–61 [毛志强, 李进福. 油气层产能预测方法及模型. *石油学报*, 2000, 21: 58–61]
- 9 Wu P Y, Jain V, Kulkarni M S, et al. Machine learning-based method for automated well-log processing and interpretation. In: *SEG Technical Program Expanded Abstracts 2018*. Tulsa: Society of Exploration Geophysicists, 2018. 2041–2045
- 10 Gupta A, Soumya U. Well log interpretation using deep learning neural networks. In: *Proceedings of International Petroleum Technology Conference, 2020*
- 11 Haldorsen J B U, Johnson D L, Plona T, et al. Borehole acoustic waves. *Oilfield Rev*, 2006, 18: 34–43
- 12 Walsh D, Turner P, Grunewald E, et al. A small-diameter NMR logging tool for groundwater investigations. *Groundwater*, 2013, 51: 914–926
- 13 Yuan J D, Wang Z H. Review of time series representation and classification techniques. *Comput Sci*, 2015, 42: 1–7 [原继东, 王志海. 时间序列的表示与分类算法综述. *计算机科学*, 2015, 42: 1–7]
- 14 Ye L X, Keogh E J. Time series shapelets: a new primitive for data mining. In: *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2009*. 947–956
- 15 Ye L X, Keogh E. Time series shapelets: a novel technique that allows accurate, interpretable and fast classification. *Data Min Knowl Disc*, 2011, 22: 149–182
- 16 Chan K P, Fu A W C. Efficient time series matching by wavelets. In: *Proceedings the 15th International Conference on Data Engineering, 1999*. 126–133
- 17 Ding H, Trajcevski G, Scheuermann P, et al. Querying and mining of time series data: experimental comparison of representations and distance measures. *Proc VLDB Endow*, 2008, 1: 1542–1552
- 18 Wang X Y, Mueen A, Ding H, et al. Experimental comparison of representation methods and distance measures for time series data. *Data Min Knowl Disc*, 2013, 26: 275–309
- 19 Zhong S, Ghosh J. HMMs and coupled HMMs for multi-channel EEG classification. In: *Proceedings of International Joint Conference on Neural Networks, 2002*. 1154–1159
- 20 Povinelli R J, Johnson M T, Lindgren A C, et al. Time series classification using Gaussian mixture models of reconstructed phase spaces. *IEEE Trans Knowl Data Eng*, 2004, 16: 779–783
- 21 Wang S C, Zheng F, Gao R. Dynamic full Bayesian ensemble classifiers for small time series. *Sci Sin Inform*, 2017, 47: 1445–1463 [王双成, 郑飞, 高瑞. 小时间序列动态完全 Bayesian 集成分类器研究. *中国科学: 信息科学*, 2017, 47: 1445–1463]
- 22 Wu T H, Weng X Q, Shan Z N. Survey of time series classification based on symbolic representation. *J Hebei Acad Sci*, 2019, 36: 11–20 [武天鸿, 翁小清, 单中南. 基于符号表示的时间序列分类综述. *河北省科学院学报*, 2019, 36: 11–20]
- 23 Wang J, Liu P, She M F H, et al. Bag-of-words representation for biomedical time series classification. *Biomed Signal Process Control*, 2013, 8: 634–644

- 24 Lin J, Khade R, Li Y. Rotation-invariant similarity in time series using bag-of-patterns representation. *J Intell Inf Syst*, 2012, 39: 287–315
- 25 Schäfer P, Höggqvist M. SFA: a symbolic fourier approximation and index for similarity search in high dimensional datasets. In: *Proceedings of the 15th International Conference on Extending Database Technology*, 2012. 516–527
- 26 Schäfer P. The BOSS is concerned with time series classification in the presence of noise. *Data Min Knowl Disc*, 2015, 29: 1505–1530
- 27 Karim F, Majumdar S, Darabi H, et al. Multivariate LSTM-FCNs for time series classification. *Neural Netw*, 2019, 116: 237–245
- 28 Zheng Y, Liu Q, Chen E H, et al. Exploiting multi-channels deep convolutional neural networks for multivariate time series classification. *Front Comput Sci*, 2016, 10: 96–112
- 29 Zheng Y, Liu Q, Chen E H, et al. Time series classification using multi-channels deep convolutional neural networks. In: *Proceedings of International Conference on Web-Age Information Management*, 2014. 298–310
- 30 Karim F, Majumdar S, Darabi H. Insights into LSTM fully convolutional networks for time series classification. *IEEE Access*, 2019, 7: 67718–67725
- 31 Shang F H, Yuan Y, Wang C Z, et al. A logging data processing method of intelligent optimization logging interpretation model based on knowledge-base. *Acta Petrolei Sin*, 2015, 36: 1449–1456 [尚福华, 原野, 王才志, 等. 基于知识库的解释模型智能优选测井数据处理方法. *石油学报*, 2015, 36: 1449–1456]
- 32 Alaudah Y, Michałowicz P, Alfarraj M, et al. A machine-learning benchmark for facies classification. *Interpretation*, 2019, 7: 175–187
- 33 Zhang W, Stewart R. Using FWI and deep learning to characterize velocity anomalies in crosswell seismic data. In: *SEG Technical Program Expanded Abstracts 2019*. Tulsa: Society of Exploration Geophysicists, 2019. 2313–2317
- 34 Tong B, Klinkigt M, Iwayama M, et al. Learning to generate rock descriptions from multivariate well logs with hierarchical attention. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017. 2031–2040
- 35 Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*, 1997, 9: 1735–1780
- 36 Zhao H K, Wu L K, Li Z, et al. Predicting the dynamics in internet finance based on deep neural network structure. *J Comput Res Dev*, 2019, 56: 1621–1631 [赵洪科, 吴李康, 李徵, 等. 基于深度神经网络结构的互联网金融市场动态预测. *计算机研究与发展*, 2019, 56: 1621–1631]
- 37 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2015. 3431–3440
- 38 Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. In: *Proceedings of Advances in Neural Information Processing Systems*, 2017. 5998–6008
- 39 Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection. In: *Proceedings of IEEE International Conference on Computer Vision*, 2017. 2980–2988
- 40 Lopez-Paz D, Bottou L, Schölkopf B, et al. Unifying distillation and privileged information. 2015. ArXiv:1511.03643
- 41 Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. 2015. ArXiv:1503.02531
- 42 Kingma D P, Ba J. Adam: a method for stochastic optimization. 2014. ArXiv:1412.6980
- 43 Ye J, Chow J H, Chen J, et al. Stochastic gradient boosted distributed decision trees. In: *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, 2009. 2061–2064

Method of oil and gas reservoir detection based on geological knowledge distillation learning

Zhi LI¹, Qi LIU^{1*}, Zhefeng WANG², Yi ZHENG², Xia LIN³, Baoxing HUAI²,
Lan MI³ & Enhong CHEN¹

1. Anhui Province Key Laboratory of Big Data Analysis and Application, University of Science and Technology of China, Hefei 230027, China;

2. Huawei Technologies Co., Ltd., Hangzhou 310051, China;

3. Research Institute of Petroleum Exploration and Development, Beijing 100083, China

* Corresponding author. E-mail: qiliuql@ustc.edu.cn

Abstract Oil and gas reservoir detection is one of the major tasks of petroleum energy companies in the exploration and production process. The oil and gas industry has long relied on the expert manual analysis of massive logging data to perform qualitative analyses of oil and gas reservoirs. Although experts' interpretations are highly accurate, the time and economic costs are considerably high. With the rapid development of artificial intelligence technologies such as deep learning in recent years, intelligent oil and gas reservoir detection methods have become a focus in the academia and industry. However, sensor data in real industrial scenarios present serious inconsistencies, which bring great challenges to traditional supervised learning models. This paper presents a focused study on the oil and gas reservoir detection task in the context of sensor inconsistencies and proposes a geological knowledge distillation multiscale network approach. This method proposes a multiscale feature fusion mechanism based on self-attention to learn the multiscale dynamic representation of geological information. Then, the model designs a geological knowledge distillation learning framework to learn additional geological knowledge from inconsistent sensor data. This step further improves the model's accuracy. A large number of experiments on real industrial datasets are subsequently performed. The results fully prove the effectiveness and robustness of the proposed model in oil and gas reservoir detection.

Keywords oil and gas reservoirs detection, geological knowledge, distillation learning, sensor data, deep neural network



Zhi LI was born in 1993. He received his B.E. degree in Software Engineering from Xi'an Jiaotong University, Xi'an, China, in 2015. Currently, he is working toward his Ph.D. degree in Data Science at University of Science and Technology of China, Hefei, China. His main research interests include data mining, industrial intelligence, and recommender systems.



Qi LIU was born in 1986. He received his Ph.D. degree in Computer Science from University of Science and Technology of China, Hefei, China, in 2013. He is currently a professor in the School of Computer Science and Technology at University of Science and Technology of China, Hefei, China. His main research interests include data mining, intelligent education, recommender systems, and social network analysis.



Zhefeng WANG was born in 1991. He received his Ph.D. degree in Computer Science from University of Science and Technology of China, Hefei, China, in 2017. He is currently a researcher of the Language and Speech Innovation Lab in Huawei Cloud, Hangzhou, China. His main research interests include natural language processing, machine learning, and social network analysis.



Enhong CHEN was born in 1968. He received his Ph.D. degree in Computer Science from University of Science and Technology of China, Hefei, China, in 1996. He is currently a professor, executive dean of the School of Data Science, and vice dean of the School of Computer Science in University of Science and Technology of China. His main research interests include data mining, artificial intelligence, recommender systems, and social network analysis.