

基于卷积神经网络的核磁共振测井渗透率预测

马欣然, 唐欣怡*

广西大学数学与信息科学学院, 南宁 530004, 广西, 中国

摘要: 渗透率是地下流体流动分析、油气藏管理、烃类采收与二氧化碳封存中的关键参数。传统的渗透率测定方法依赖成本高、耗时长、室内实验或与井相关的资料。机器学习——特别是卷积神经网络 (CNN) ——因其能够挖掘输入变量与输出变量之间的内在联系, 被认为是一种低成本、快速的渗透率预测手段。本文用卷积神经网络建立了渗透率的经验关联式: 先对 40 条核磁共振 (NMR) T2 谱与 89 个 NMR T2 对数平均分布 (T2lm) 作预处理与筛选, 再用主成分分析 (PCA) 识别出关键谱段; 随后采用自行设计的卷积神经网络结构, 充分利用核磁共振数据中蕴含的空间模式与复杂关系来构建关联式; 并用 k 折交叉验证方案对模型作了严格的训练与验证, 以保证其稳健性与泛化能力。评价模型精度所用的指标包括决定系数 (R^2)、均方根误差 (RMSE)、绝对误差 (MAE)、标准差 (SD)、绝对偏差 (AD)、平均绝对偏差 (AAD)、平均绝对百分比相对误差 (AAPRE) 与最大误差 (Emax)。在所考察的各折当中, 第 1 折表现最佳, R^2 达到 0.9544。该基于卷积神经网络的关联式在 R^2 、Emax、AD、AAD、AAPRE 等指标上均优于常规模型与其他人工智能模型。总体而言, 本研究表明卷积神经网络能够有效预测渗透率, 可替代成本高昂且适用范围有限的传统方法, 其中第 1 折的结果最为理想。

关键词: 卷积神经网络; 核磁共振; 渗透率; 主成分分析

Prediction of Permeability Via Nuclear Magnetic Resonance Logging Using Convolutional Neural Networks

XinRan Ma, XinYi Tang*

College of Mathematics and Information Science, Guangxi University, Nanning 530004, Guangxi, China

Abstract: Permeability is a critical parameter in subsurface fluid flow analysis, reservoir management, hydrocarbon recovery, and carbon dioxide sequestration. Traditional permeability measurement methods involve costly and time-consuming laboratory tests or well-related data. Machine learning (ML), specifically convolutional neural networks (CNN), is proposed as a cost-effective and rapid permeability prediction solution, harnessing interrelationships of input-output variables. In this study, empirical permeability correlation was developed using CNN. Forty nuclear magnetic resonance (NMR) T2 spectrums and 89 logarithmic mean NMR T2 distributions (T2lm) were preprocessed, screened and key spectra were identified using the principal component analysis (PCA). To develop the correlations, a custom-designed CNN architecture was employed to leverage the spatial patterns and intricate

relationships embedded in the NMR data. The model was trained and validated rigorously using k-fold cross validation scheme to ensure robustness and generalization. Performance metrics like R-squared (R²), root mean squared error (RMSE), mean absolute error (MAE), standard deviation (SD), absolute deviation (AD), average absolute deviation (AAD), average absolute percentage relative error (AAPRE), and maximum error (Emax) were deployed to evaluate the model's accuracy and ability to predict permeability values accurately. Among the folds considered, the fold 1 emerged as the best-performing model with the highest R² value of 0.9544. This CNN-based correlation outperformed conventional and other AI-based models in terms of R², Emax, AD, AAD, AAPRE, among other metrics. Overall, the study demonstrates the effectiveness of CNN in predicting permeability, offering a superior alternative to costly and limited traditional methods, with fold 1 showing the most promising results.

Keywords: Convolutional neural networks; Nuclear magnetic resonance; Permeability; Principal components analysis

一、引言

用数值方法预测多孔介质的渗透率是一条很有价值的辅助途径, 相较实验测定, 它成本更低、效率更高[1]。分析储层的输运与储集特征, 有助于评价其储量与地层条件。特别是渗透率的分布, 在制定油气田开发方案与控制生产时至关重要。除油气行业之外, 准确估算渗透率对二氧化碳封存、水资源管理、地热开发等地下流体流动类应用同样关键[2-3]。石油工程借助渗透率来评价岩石地层让油、气、水通过的能力, 这一属性直接关系到烃类储层的产能与经济可行性: 渗透率高, 说明岩石孔隙发育、连通良好, 流体易于流动; 渗透率低, 则表明地层致密, 流体难以通过。

室内岩心测量与试井压力恢复测试是估算储层渗透率的常用手段[4]。这两类方法都费时费力、资源投入大, 而且只能得到有限几口已钻井的渗透率数据。作为替代, 地球物理测井也可以用来确定多孔介质的渗透率。

然而, 仅凭常规测井曲线很难有效评价渗透率。这类曲线测量的是孔隙度, 要预测渗透率还得依赖经验或半经验关联式。近几十年来, 核磁共振(NMR)测井因其估算渗透率更为精确而日渐普及[4]。与受矿物成分影响很大、同时响应固相与流体的常规测井不同, 核磁共振测井的响应对应的是多孔介质中流体的体积、组成、黏度与分布状态, 因此它比其他测井方法更能准确地刻画岩石的导流能力[5]。不过, 核磁共振测井仪器采集成本高、分析耗时长, 而且不适用于套管井, 这些限制阻碍了它的广泛应用, 也使得核磁共振资料相当稀缺。为了用核磁共振测井估算渗透率等储层属性, 人们尝试由更易获取的常规测井曲线合成核磁共振测井曲线(或核磁共振派生成果)。

本文其余部分安排如下: 第二部分讨论深度学习中的卷积神经网络技术; 第三部分说明本研究所采用的方法; 第四部分给出结果与讨论; 第五部分总结全文并提出后续工作的建议。

二、文献综述

本部分从概念上梳理人工神经网络(ANN), 其中既包含机器学习(ML), 也包含深度学习(DL)——本文用以建立渗透率预测经验关联式的卷积神经网络技术即属于后者。机器学习在石油工程与地球科学领域日益流行。由于在实际工况下行之有效, 用机器学习算法估算无岩心井段渗透率的做法在油气行业已被广泛采用[6]。石油工业已成功地把机器学习用于储层表征、地震模式识别、属性预测、岩相识别、单井产量趋势估计以及多样化的储层组合评价[7]。近年来, 不少研究者尝试把岩心渗透率与测井曲线关联起来。理解渗透率是储层表征的关键: 地层渗透率会影响完井方式、增产措施以及储层管理策略[1]。取心与试井既费时又昂贵, 且难以刻画储层的非均质性, 这限制了可行钻井作业的数量; 某些层段缺失岩心或岩心资料不全, 也会带来困难。

试井中的压力瞬变分析给出的是所研究储层体积的平均渗透率, 试井时间必须足够长, 才能取得可靠数据并反映相当一部分储层体积[8]。第二条途径是为每口井建立方程并确定其中的系数, 但适用于某一地层的方程未必适用于其他区块。多数测井技术并不直接测量渗透率, 需要经过解释才能得到。研究者曾用非线性回归把深度、孔隙度等储层属性与渗透率间接关联起来, 然而这种方法更适合砂岩储层, 在复杂碳酸盐岩地层中往往失效——成岩作用加剧了这类地层的非均质性, 使该方法难以奏效。因此, 需要一种能够跨越储层非均质性、可靠预测渗透率的方法[9-12]。石油工程师正致力于改进渗透率估算, 以降低油气勘探开发中的不确定性; 他们从自然界汲取灵感, 研究那些历经长期演化而成的解决方案。自然界的机制能够适应多种情形, 是构建新颖而可靠方案的理想范本。

受自然启发的方法包括人工神经网络、进化计算 (EC)、群体智能 (SI)、支持向量机 (SVM)、高斯过程回归方法 (GPRM) 与模糊系统 (FS) 等。把这些方法整合起来, 工程师便能构建出稳健且一体化的系统[7]。近来有研究引入高斯过程回归与支持向量机来建立渗透率经验关联式: 该工作对 40 条核磁共振 T2 谱与 89 个 T2 对数平均分布 (T2lm) 作了预处理与筛选, 并用主成分分析识别关键谱段; 随机选取 80% 的数据用于建模, 其中一半用于训练、一半用于测试, 并考察了十种不同的核函数。结果显示, 采用平方指数核的高斯过程回归 ($R^2=1.0$, $RMSE=2.1512 \times 10^{-4}$) 与采用三次核的支持向量机 ($R^2=0.97$, $RMSE=2.7393 \times 10^{-2}$) 效果最佳; 将这些结果与人工神经网络、SDR 模型和 Timur-Coates 模型作对比, 凸显了高斯过程回归模型的优越性[5]。另有研究把扫描电镜预测的渗透率与实测值作了比较, 发现影响孔喉形状的表面粗糙度会导致渗透率被低估。

An 等用规则网络模型研究了孔喉比、配位数与孔喉取向对绝对渗透率的影响。Shah 等则预测了孔隙网络的特征参数, 如孔隙数与喉道数、平均孔喉半径、配位数等, 并用格子玻尔兹曼方法 (LBM) 与孔隙网络模型 (PNM) 模拟了单相与两相流动条件。卷积神经网络属于深度学习, 是神经网络的一种扩展, 主要用于图像识别与图像处理任务。它在检测图像中的模式与特征方面极为有效, 已成功应用于目标检测、人脸识别、自动驾驶等诸多场景。图 1 给出了卷积神经网络的架构。

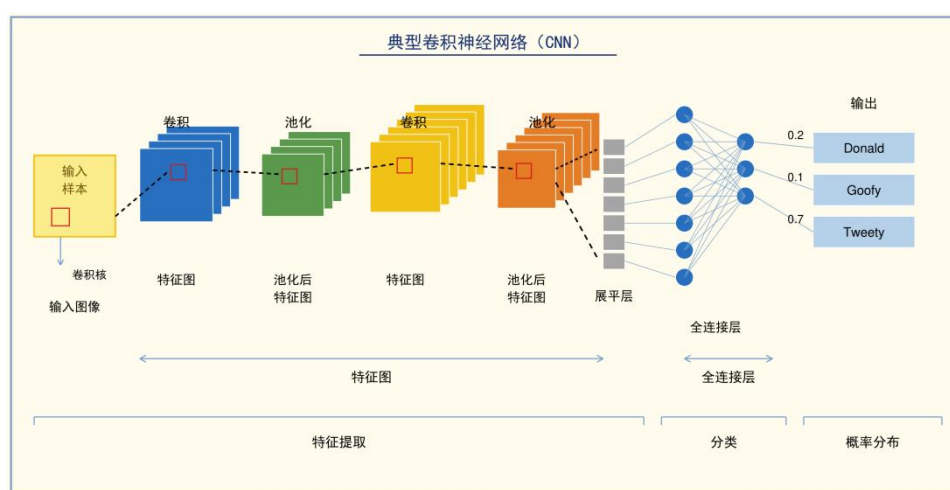


Figure 1 CNN Architecture

图 1 卷积神经网络的架构

卷积神经网络的主要特点, 是能够从所获取的数据中自动学习并发现特征。这一过程由卷积层完成: 卷积层对输入数据施加滤波器, 生成特征图; 随后特征图再经过池化层、全连接层等若干层的处理, 得出最终结果。除图像处理之外, 卷积神经网络在自然语言处理 (NLP)、语音识别乃至金融等领域也展现出巨大潜力, 同样

可以用于地质与地球物理领域, 例如由核磁共振测井预测岩石渗透率。总体来看, 卷积神经网络是从复杂高维数据中学习的有力工具, 其通用性使它成为机器学习与人工智能中重要的研究与开发方向。

卷积神经网络已成为分析包括核磁共振数据在内的复杂数据的一种强有力的机器学习技术, 在图像识别、语音识别、自然语言处理、药物发现等多个方面均有成功应用, 也已被用于由测井曲线预测岩石物理属性(含渗透率)。测定储层岩石的渗透率, 对评价油气藏的可行性与产能至关重要。核磁共振测井是测量岩石物理属性(含渗透率)的常用方法, 但由于岩石的信号响应复杂, 其解释颇具难度, 因此卷积神经网络被提出作为由核磁共振测井准确预测渗透率的一条可行路径。

核磁共振测井测量的是岩石孔隙流体与固体骨架中氢质子的弛豫时间, 可据以估算岩石的孔径分布与渗透率。然而, 解释核磁共振测井需要对岩石物理性质有深入的理解, 传统分析方法既耗时又容易出错。正因如此, 人们对把卷积神经网络等机器学习技术用于核磁共振测井自动解释的兴趣不断增长。

卷积神经网络是一类在图像识别与图像处理任务中取得显著成功的深度学习算法。已有多项研究提出用卷积神经网络由核磁共振测井预测渗透率。例如 Liu 等用卷积神经网络由核磁共振数据预测渗透率, 精度达 84.4%, 优于传统回归方法; 该研究还发现, 卷积神经网络能够捕捉岩石中孔隙的空间分布, 而这正是渗透率估算的关键因素。

Zhang 等[1]提出用一个以无标签数据训练的自编码器(AE)模块, 把低分辨率多孔介质映射到高分辨率特征, 以辅助主卷积神经网络作预测。该方法同时利用了自编码器由大量无标签数据学到的高分辨率信息, 以及卷积神经网络由少量有标签数据学到的低分辨率信息。通过计算均方误差(MSE)与决定系数(R^2), 他们把 AE-CNN 由低分辨率图像所得的预测效果与常规卷积神经网络及格子玻尔兹曼方法的结果作了对比: 在 5 折交叉验证下, AE-CNN 测试集的平均 R^2 为 0.896, 不含自编码器的常规卷积神经网络为 0.869; 在表现最好的一折中, AE-CNN 训练集与测试集的均方误差分别为 0.022 与 0.064, 而不含自编码器的卷积神经网络分别为 0.034 与 0.083。这表明自编码器模块能显著提升由低分辨率多孔介质图像预测渗透率的效果。格子玻尔兹曼方法的模拟结果则显示, 低分辨率图像模糊的边界处存在大量数值误差, 使其预测可靠性(平均 R^2 为 0.42; 表现最好的一折中均方误差为 0.37 与 0.36)明显逊于基于卷积神经网络的机器学习方法。

综上, 上述文献以及相关文献中存在的研究空白, 正是本文进一步探讨的动因。用卷积神经网络由核磁共振测井预测渗透率, 已展现出提升油气行业(尤其是储层表征)预测精度的潜力, 但把卷积神经网络用于这一任务仍伴随诸多局限与挑战。

三、方法

(一) 步骤 1: 数据采集

第一步是获取尼日尔三角洲盆地 X 油田的核磁共振测井资料。这批数据以 CSV 格式存储——这是一种被广泛接受、受众多软件支持的表格数据标准格式。载入与解析 CSV 文件所用的软件为 Anaconda, 它是数据科学与机器学习中常用的 Python 发行版, 内置多个基于 Web 的集成开发环境(IDE)。

(二) 步骤 2: 数据预处理

获取核磁共振测井资料之后, 下一阶段是数据预处理, 以提高数据的质量与可用性, 包括去除噪声、处理缺失值, 以及通过特征工程从数据中提炼有意义的信息。预处理安装并使用了 Keras、TensorFlow、Scikit-learn 等 Python 库, 具体步骤如下。

数据清洗: 首先检测并处理数据中的缺失值与离群值。缺失值采用均值插补处理; 离群值会损害模型性能,

予以剔除或校正。

数据归一化: 为使各特征的量纲一致, 采用标准化方法对数据作归一化处理。归一化有助于改善训练过程、加快模型收敛。

(三) 步骤 3: 特征提取

把预处理后的核磁共振测井数据输入卷积神经网络模型以完成特征提取。模型结构的设计目标, 是识别数据特征中的空间模式与相关性, 其中包含以下几类层。

卷积层: 通过对输入数据施加一组滤波器来完成特征提取。每个滤波器在输入数据上滑动, 与重叠区域的数值作点积运算。卷积层的输出是一组特征图, 其中蕴含了数据中的空间模式。

池化层: 借助最大池化这一下采样手段, 缩减特征图的空间维度。该操作有助于减少模型的参数数量, 从而提高计算效率。

全连接层: 最终的分类或回归任务由全连接层完成。这些层接收前面各层展平后的输出, 施加权重后生成最终结果。

这一特征提取过程, 使卷积神经网络模型得以学习并识别预处理后的核磁共振测井数据中的重要模式与关系, 最终有助于更准确地预测渗透率。

(四) 步骤 4: 模型构建

卷积神经网络模型借助 Python 的 Keras 与 TensorFlow 库搭建, 这两个库为指定模型结构、配置超参数与训练模型提供了便利的接口。

模型结构: 卷积神经网络的结构通过 Keras API 建立, 该接口简化了复杂神经网络的搭建。模型包含多个卷积层, 其后依次为池化层与全连接层。

超参数调优: 对学习率、批大小、训练轮数等超参数作了细致调整, 以优化模型性能。

图 2 给出了本研究中用卷积神经网络由核磁共振测井预测渗透率的模型结构说明。

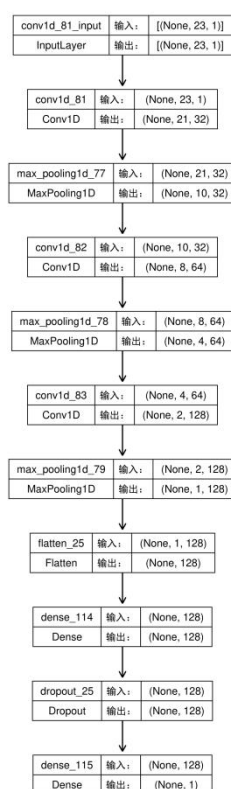


Figure 2 Model Description of the CNN Model

图 2 卷积神经网络的模型结构说明

(五) 步骤 5: 数据划分与模型训练

数据划分: 用 Python 中 Scikit-learn 库提供的 k 折交叉验证函数, 把预处理后的数据划分为 5 折, 分别用作训练子集与验证子集。

模型训练: 以 TensorFlow 为底层框架, 通过 Keras API 训练卷积神经网络模型。

结构调优: 针对具体任务, 对卷积神经网络的结构作了细致的定制与优化, 过程中还一并考虑了卷积层与池化层的数量、滤波器尺寸、全连接层的神经元个数以及学习率等超参数。

优化器与损失函数: 模型训练采用 Adam 优化器与均方误差 (MSE) 损失函数。均方误差适用于回归任务, 有助于缩小渗透率预测值与实测值之间的差距。

训练终止: 训练持续到验证损失不再下降、且达到预先设定的最大训练轮数为止, 此时模型已达到最优状态, 可避免过度训练。

(六) 步骤 6: 模型评价与预测

模型训练完成之后, 还要对卷积神经网络的渗透率预测能力作全面评价。

测试集评价: 把训练好的卷积神经网络模型放到一个与训练集、验证集完全分离的独立测试集上评估, 以了解模型面对未见数据时的表现。

性能指标: 用决定系数 (R^2)、均方根误差 (RMSE)、绝对误差 (MAE)、标准差 (SD)、绝对偏差 (AD)、平均绝对偏差 (AAD)、平均绝对百分比相对误差 (AAPRE) 与最大误差 (Emax) 等一系列指标衡量模型性能。这些指标可以全面反映模型的预测精度, 以及它捕捉预测值与实测值之间差异的能力。

图 3 给出了本研究中卷积神经网络模型的工作流程。在此之后, 还用不同的机器学习算法建立模型, 并用

不同的性能指标考察各模型的表现。

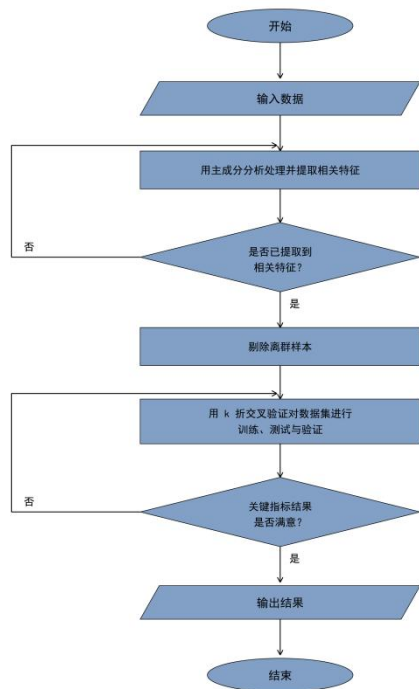


Figure 3 Workflow for the CNN Model

图 3 卷积神经网络模型的工作流程

实际渗透率预测：把训练好的卷积神经网络模型用于由新的核磁共振测井曲线预测渗透率，再把预测值与真实渗透率对照，以检验模型在实际场景中的精度与可靠性。

（七）步骤 7：模型解释与可视化

卷积神经网络训练与评估完成之后，还需要深入其内部机理，以更好地理解模型的决策过程与它所学到的特征。这一步包括模型解释与可视化。

特征图：把卷积神经网络在输入数据中检测到的模式与特征可视化，直观呈现卷积层学到的滤波器，进而揭示哪些空间特征与渗透率预测最为相关。

激活最大化：优化输入图像，使卷积神经网络中特定神经元或特征图的激活值最大，从而生成能强烈激活某些神经元的输入图像。

梯度加权类激活映射（Grad-CAM）：引入 Grad-CAM 技术，突出显示输入图像中对模型预测贡献最大的区域。它生成的热力图可以指示输入图像的哪些部分在模型决策过程中起主导作用。

随后，把卷积神经网络的渗透率预测结果与既往研究中人工神经网络、支持向量机与高斯过程回归的预测结果作对比，以确定由核磁共振测井预测渗透率的最佳方法。表现最好的模型进入实施阶段：该阶段以最优模型为核心，接收前端传入并在此阶段处理的输入，再把输出结果回送前端；前端则与数据库相连，支持用户注册与登录。

四、结果与讨论

（一）结果讨论

为检验卷积神经网络模型的泛化能力与稳定性，用训练好的模型预测了尼日尔三角洲盆地 X 油田一口井

的渗透率，测试数据共 89 个数据点。模型的训练与验证划分为五折，并以神经网络领域常用的性能指标加以衡量。结果显示：第 1 折训练过程的 RMSE 为 2.390876，测试过程为 2.2772；第 2 折训练为 2.7297054，测试为 1.956999；第 3 折训练为 2.7777839，测试为 2.672934517；第 4 折训练为 2.307744885，测试为 3.42918388；第 5 折训练与测试的 RMSE 分别为 2.50440248 与 2.322125687。

Table 1 Metrics for CNN folds Results

表 1 卷积神经网络各折的指标结果

指标	第 1 折	第 2 折	第 3 折	第 4 折	第 5 折
RMSE	2.2772	1.9567	2.6729	3.4292	2.3221
R ²	0.9544	0.9267	0.8846	0.904	0.8879
MAE	1.5369	1.5081	1.8855	2.1877	1.6928
E _{max}	110.8197	188.4445	2087.9464	948.6314	1026.4923
SD	1.0511	1.3293	1.7674	1.5181	1.7107
AD	1.1873	1.5081	1.8855	1.4615	1.6928
AAD	0.15147	0.4299	0.97211	0.85700	0.4093
AAPRE	2.7586	24.7136	149.4856	79.1761	74.9408

Table 2 Statistical Comparison with ANN, SVM, GPR, TC, and SDR Models

表 2 与 ANN、SVM、GPR、TC 及 SDR 模型的统计对比

指标	本研究 (CNN)	(SVM)	(GPR)	Olayiwola (2017) (ANN)	Timur-Coates 模型	SDR 模型
AD	1.1873	0.3159	-4.955e-12	0.020	-1.3169	-1.5724
AAD	0.15147	0.6816	0.1636	0.620	2.7953	3.0740
RMSE	2.2772	1.0573	0.1961	1.510	3.5021	3.8619
SD	1.0511	26.563	4.4842	6.100	66.991	209.05
E _{max}	110.8197	77.155	12.014	6.090	139.24	637.88
AAPRE	2.7586	14.467	3.1058	6.600	49.949	124.23

(二) 第 1 折至第 5 折的性能评价结果

卷积神经网络第 1 折至第 5 折的性能评价结果为：AD 依次为 1.1873、1.5081、1.8855、1.4615、1.6928；AAD 依次为 0.15147、0.4299、0.97211、0.85700、0.4093；RMSE 依次为 2.2772、1.9567、2.6729、3.4292、2.3221；R² 依次为 0.9544、0.9267、0.8846、0.904、0.8879。其余结果见表 1，与人工神经网络、支持向量机、高斯过程回归、Timur-Coates 模型及 SDR 模型的统计对比见表 2。图 4 给出了卷积神经网络预测值与岩心实测渗透率的散点图，其中图 4 (a) 至图 4 (e) 依次对应第 1 折至第 5 折。

五、结论与展望

本文提出用卷积神经网络由核磁共振测井预测渗透率。结果表明，卷积神经网络在这一任务上效果显著，可推荐用于储层管理与开发。不过需要注意的是，卷积神经网络模型通常建立在来自特定地区的大规模数据集之上，数据量对其预测性能影响很大，因此在构建新的卷积神经网络模型时，宜采用规模足够大的数据集。就 R²、E_{max}、AD、AAD 与 AAPRE 而言，本文建立的关联式优于常规模型与其他人工智能模型。后续工作建议采用其他深度学习 / 机器学习模型对本研究作进一步验证。

参考文献

- [1] 吴一戎. 合成孔径雷达成像优化算法. 雷达学报, 2022, 11(02): 201-212.
- [2] 黄维和. 天然气管网动态调度优化. 油气储运, 2024, 41(04): 361-368.
- [3] Ajagbe S A, Adegun A A, Mudali P, et al. Performance of machine learning models for pandemic detection using COVID-19 dataset//2023 IEEE AFRICON. 2023: 1-6.
- [4] 叶列平. 再生混凝土高性能化改性技术. 建筑材料学报, 2022, 25(02): 305-314.
- [5] 李清泉. 车载激光雷达高精度定位技术. 测绘学报, 2024, 51(04): 437-446.
- [6] 赵金洲. 页岩油体积压裂增产技术. 石油学报, 2022, 43(04): 511-520.
- [7] Zhang Y, Nishizawa O, Park H, et al. Relative permeability of CO₂ in a low-permeability rock: implications for CO₂ flow behavior in reservoirs with tight interlayers. Energy Procedia, 2017, 114: 4822-4831.
- [8] 苑世剑. 大型构件液态模锻成型工艺. 塑性工程学报, 2020, 29(04): 1-8.
- [9] 周绪红. 钢混组合桥梁受力优化设计. 中国公路学报, 2022, 35(04): 1-14.
- [10] 戴琼海. 工业视觉缺陷检测深度学习模型. 自动化学报, 2020, 48(06): 1105-1116.