

机器学习与计算力学的结合及应用初探

李想, 严子铭, 柳占立*

清华大学航天航空学院, 应用力学实验室, 北京 100084

* 联系人, E-mail: liuzhanli@tsinghua.edu.cn

2019-01-02 收稿, 2019-01-11 修回, 2019-01-14 接受, 2019-02-15 网络版发表

国家自然科学基金优秀青年科学基金(11722218)和科学挑战计划(TZ2018001)资助

摘要 自20世纪50年代以来, 随着计算机科学的不断进步, 机器学习和数据科学得到了长足发展. 这些技术一般依靠大量数据作支撑, 通过训练过程提取出蕴藏在数据内部的抽象映射关系, 目前已被成功应用于化学、生物等自然科学研究领域. 近年来, 这些技术也逐渐受到计算力学领域研究者的关注. 本文结合作者的相关研究成果介绍了机器学习、数据科学与计算力学相结合的3种形式: 第一种是与有限元方程求解方面的结合, 直接应用卷积神经网络算法求解线性有限元方程; 第二种方式结合有限元计算和机器学习预测复杂材料结构与力学性能的关系. 本文作者应用该方法基于细观页岩扫描照片和随机建模算法, 成功训练出可以有效预测细观页岩样本等效模量的卷积神经网络; 第三种方式是建立基于数据驱动的计算力学方法, 比如直接利用真实的材料实验数据代替材料本构模型. 这些工作显示了机器学习、数据驱动在处理材料的力学实验数据、设计新型材料以及创建更高效的计算力学模型方面的广阔前景. 随着计算力学的发展, 未来将可能出现更多将数据科学、机器学习与计算力学结合的应用场景, 进一步开发出更加强健、高效和保真的计算力学方法.

关键词 机器学习, 数据驱动, 人工神经网络, 计算力学, 有限单元法

在力学发展过程中, 研究者往往通过提出理想化的力学模型开展理论分析. 这些理想模型一般具有较为规则的几何形状、边界条件和基于显式表达式描述的材料响应函数, 将真实的复杂力学问题抽象为简单的数学模型, 进而寻求问题的解析解. 计算力学方法的出现使求解具有更复杂的几何模型和边界条件的问题成为可能. 例如, 有限单元法将连续求解区域划分为离散单元, 将求解连续模型上的偏微分方程转换为求解离散模型上的积分方程, 通过离散单元上的近似解逼近连续模型的真实解^[1,2].

随着计算机科学的不断发展, 机器学习和数据科学取得了长足进步. 常见的一些机器学习方法包括感知器^[3]、遗传编程^[4]、蒙特卡罗方法^[5]、K-近邻算法(K-nearest neighbor, K-NN)^[6]、线性回归^[7]、逻辑

回归^[8,9]、自适应增强^[10]和支持向量机^[11,12]等. 这些方法各具特点. 例如: 线性回归方法使用线性函数对数据集进行拟合, 通过梯度下降法优化最小二乘误差函数, 一般采用均方差函数描述误差. 该方法计算简单且易于实现, 但其固有缺点是无法拟合非线性数据^[7]. 逻辑回归是一种用于解决分类问题的判别式模型, 它与线性回归相比最大的区别在于引入非线性函数, 不要求训练数据满足线性关系, 一般使用分类交叉熵作为误差函数. 其缺点是对于多分类问题的拟合精度欠佳^[8,9]. K-近邻算法基于特征空间内与目标样本最相似(或最临近)的 k 个样本的类别对目标样本进行分类^[13]. 该方法需要记录所有训练数据, 导致计算成本较高, 且结果比较依赖于 k 的取值^[14]. 但该方法同时拥有以下优点: 简单易用、准确度高且

引用格式: 李想, 严子铭, 柳占立. 机器学习与计算力学的结合及应用初探. 科学通报, 2019, 64: 635–648

Li X, Yan Z M, Liu Z L. Combination and application of machine learning and computational mechanics (in Chinese). Chin Sci Bull, 2019, 64: 635–648, doi: 10.1360/N972019-00005

可以用于分类和回归问题. 支持向量机的目标是在高维度特征空间寻找最优的划分超平面. 该方法适用于处理高维度样本数据, 被广泛应用于文本分类和图像识别领域^[15]. 其性能依赖于核函数的选取^[16], 另外对于大量样本的计算性能还有待提高^[17]. 自适应增强方法在迭代训练时关注之前分类错误的样本, 通过不断修正之前的错误分类结果, 最终将每个迭代步训练得到的弱分类器基于权重叠加为一个强分类器. 该方法的分类精度高且不容易过拟合, 但缺点是模型训练时间较长^[18].

除了前述的各种方法之外, 人工神经网络(artificial neural network, ANN)已经逐步发展成机器学习领域的一个重要分支. 其相关研究工作可以追溯到20世纪40年代. 1943年, McCulloch和Pitts^[19]受到生物神经活性特点的启发, 首次提出了人工神经元的数学模型. 1958年, Rosenblatt^[3]建立了感知器(perceptron)的模型, 该模型也被普遍认为是现代人工神经网络的前身. 最早的人工神经网络模型是由斯坦福大学创建的名为“ADALINE”和“MADALINE”^[20]. 这些模型受制于其线性和相对简单的结构, 难以被用于解决复杂非线性问题. 1971年, Ivakhnenko^[21,22]首次提出了基于感知器的多层网络模型, 并在模型中应用了非线性多项式函数. 20世纪70~80年代, Rumelhart等人^[23]和Werbos^[24]分别独立地提出了反向传播算法. 该算法后来逐步发展为现代深度神经网络的核心组成部分. 2006年, Hinton和Salakhutdinov^[25]提出了基于受限玻尔兹曼机(restricted Boltzmann machine, RBM)的神经网络模型且首次明确地提出深度学习的概念. 该模型通过训练包括多个中间层的神经网络, 可以有效地提取高维度数据样本的特征, 实现了数据从高维度到低维度的变换.

对比上述各种机器学习方法, 人工神经网络模型使用了非线性激活函数, 利用多层网络结构引入了大量可调节参数, 通过梯度下降算法挖掘蕴含于大量训练数据中的非线性关系, 一般可获得良好的分类和回归精度, 且对噪音数据不敏感, 具有一定的概括和推广能力^[26~28]. 由于这些优点, 人工神经网络被逐步应用到信号处理、模式识别、自动控制、生物医药和金融等领域^[29]. 但是, 多层神经网络结构中包含大量待定参数, 往往需要长时间的训练过程. 网络结构的选择依赖于经验, 且存在过拟合问题^[30,31]. 另外, 人工神经网络模型的复杂结构使得其

解释性较差, 难以直接通过模型参数解读数据背后蕴藏的抽象规律. 不过, 近年来发展的反卷积网络等技术可以将网络中的低维抽象数据映射到对应的高维空间进行分析^[32~34].

近年来, 包括人工神经网络在内的各种机器学习方法被不断应用于工程和力学等领域. Jin等人^[35], Zhou等人^[36], Fan^[37], Ge^[38], Feng和Diao^[39]将人工神经网络用于岩石力学的研究; Wu和Zhao^[40], Gao和Li^[41], Zhang和Zhang^[42]利用人工神经网络进行结构优化设计; Xia和Xiong^[43], Feng和Wang^[44], Chen等人^[45]基于人工神经网络开展了边坡稳定性的相关研究; Lei等人^[46], Ohsaki^[47], Hajela和Lee^[48]将支持向量回归、K近邻算法等模型用于拓扑优化设计; Ghaboussi等人^[49], Jung和Ghaboussi^[50], Ji等人^[51], Furukawa和Yagawa^[52], Hashash等人^[53], Sun等人^[54]利用人工神经网络研究了固体材料的本构关系; Faller和Schreck^[55], Wang和Liao^[56], Yuhong和Wenxin^[57], Butz和Stryk^[58], Beigzadeh和Rahimi^[59], Mi等人^[60]借助数值计算方法和人工神经网络开展了流体力学的相关研究. 这些研究工作表明, 机器学习模型在处理材料的力学响应数据、揭示复杂或模糊的力学机理等方面具有潜在优势, 逐渐受到计算力学研究者的关注. 本文将通过3个例子来初步阐述数据科学、机器学习等与计算力学结合的应用, 并希望为该领域的进一步发展提供新的思路.

1 卷积神经网络在线性有限元求解中的应用

人工神经网络因其优良的数据拟合性能, 逐渐成为目前主流的机器学习技术之一. 其核心思想是构建一个包括大量待定参数的网络结构, 通过梯度下降算法寻找最优的网络参数组合, 使得误差函数达到极值. 学者们尝试将人工神经网络的算法应用于线弹性有限元方程和偏微分方程的求解中^[61~64]. 在本节中, 作者将目前广泛应用的卷积神经网络框架整合于线弹性有限元的求解过程中. 首先简要回顾传统人工神经网络的基本原理, 然后将其与卷积神经网络结构进行对比, 接着阐述本节所提出方法的实现过程.

图1(a)是人工神经网络基本结构的示意图. 人工神经网络中一般包括一个输入层、一个输出层和若干中间隐层. 网络中的每一个圆代表一个神经元. 神经元的值称为“激活”. 两个神经元之间的连接线称为

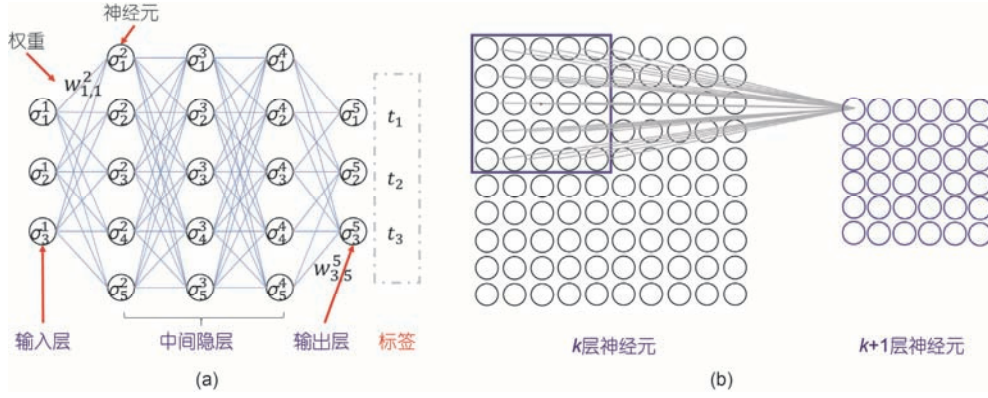


图1 (网络版彩色)人工神经网络结构示意图。(a) 传统神经网络; (b) 卷积神经网络

Figure 1 (Color online) The architecture of an artificial neural network. (a) Traditional network; (b) convolutional neural network

权重. 神经元的值由上一层神经元的值与两层神经元之间的权重计算得到^[65,66], 如式(1)所示. σ 代表神经元的激活, w 是权重, b 称为偏差. f 表示非线性函数, 常用的非线性函数包括ReLU和Sigmoid函数^[66,67]:

$$\sigma_j^l = f(z_j^l) = f\left\{\sum_{i=1}^{n_{l-1}}(w_{ji}^{l-1} \cdot \sigma_i^{l-1} + b_j^{l-1})\right\}. \quad (1)$$

卷积神经网络(convolutional neural network, CNN)是人工神经网络一个重要的分支或变种, 概念源自Fukushima和Miyake^[68]提出的“新认知机”(neocognitron)模型. 20世纪90年代, LeCun等人^[69,70]确立了卷积神经网络的现代结构, 并不断对其完善. 与传统人工神经网络相比, 卷积神经网络具有以下主要特点^[26,66,68]: 第一个特点称为共享参数, 表示同一个卷积层的所有神经元拥有相同的权重 w 和偏差 b , 能够有效减少网络参数数量, 提高训练效率. 第二个特点是局部接受域. 图1(a)所示的传统神经网络中, 相邻两层网络中的神经元是完全连接的. 而在如图1(b)所示卷积神经网络中, 每一层只有部分神经元与下一层的一个神经元相连. 第三个特点是引入池化层. 池化层通常施加在卷积层之后. 最大池化和平均池被广泛用于简化来自卷积层的输出信息^[66,67]. 该网络被成功运用在机器视觉、语音处理及棋类游戏等领域^[26,71~74].

卷积神经网络的神经元激活通过下式得到:

$$\sigma_j^l = f(z_j^l) = f\left(b + \sum_{k=0}^m \sum_{l=0}^m w_{k,l} a_{i+k,j+l}\right), \quad (2)$$

其中, w 是权重, b 称为偏差, m 是权重的尺寸. w 与 b 共同构成网络的参数集合 $\mathbf{v}$. 输出层神经元的值与数据标签之间往往存在一定的误差. 为方便阐述, 本文选

择均方差函数作为误差函数 $C(\mathbf{v})$:

$$C(\mathbf{v}) = \frac{1}{2n} \sum_{i=1}^n \|\sigma_i - t_i\|^2. \quad (3)$$

式中的 σ 代表输出层的神经元激活, t 代表数据的标签, n 代表训练样本的数量. 误差函数 $C(\mathbf{v})$ 的变化可以通过自变量 $\mathbf{v}$ 的变化来近似:

$$\Delta C \approx \nabla C \cdot \Delta \mathbf{v}, \quad (4)$$

其中, ∇C 是误差函数的梯度, $\Delta \mathbf{v}$ 是参数的微小变化量. $\Delta \mathbf{v}$ 的取值由下式确定:

$$\Delta \mathbf{v} = -\eta \nabla C, \quad (5)$$

η 是学习率, 其取值需谨慎选取以保证式(4)的近似关系成立. 权重和偏差在每一次迭代计算中按式(6)不断更新:

$$w_i = w_i - \eta \frac{\partial C}{\partial w_i}, \quad b_j = b_j - \eta \frac{\partial C}{\partial b_j}. \quad (6)$$

通过不断调整参数, 网络最终可建立起输入数据到标签的映射函数. 以上简述了人工神经网络的基本原理. 接下来简要回顾线弹性有限元的基本原理. 对于一个静态线弹性系统, 其系统总势能可以表示为内力功和外力功的组合^[75]:

$$\begin{aligned} E_p &= \int_{\Omega} \frac{1}{2} \boldsymbol{\varepsilon} : \boldsymbol{\sigma} d\Omega - \int_{\partial\Omega} \mathbf{f} \cdot \mathbf{u} d\partial\Omega \\ &= \sum_e \left(\int_{\Omega_e} \frac{1}{2} \mathbf{u}_e^T \cdot \mathbf{K}_e \cdot \mathbf{u}_e d\Omega_e \right) - \sum_e \left(\int_{\partial\Omega_e} \mathbf{f}_e \cdot \mathbf{u}_e d\partial\Omega_e \right) \\ &= \frac{1}{2} \mathbf{u}^T \cdot \mathbf{K} \cdot \mathbf{u} - \mathbf{f} \cdot \mathbf{u}, \end{aligned} \quad (7)$$

其中, $\mathbf{K}_e$, $\mathbf{u}_e$ 和 $\mathbf{f}_e$ 分别代表每个单元上的刚度和节点力矩阵. $\mathbf{K}$, $\mathbf{u}$ 和 $\mathbf{f}$ 分别代表经过组装得到的整体离散模型对应的刚度、节点位移自由度和节点力矩阵. 根据最小势能原理, 真实的位移解使得系统总势能达到最

小值, 即:

$$\frac{\partial E_p}{\partial \mathbf{u}} = 0. \quad (8)$$

由式(8)进一步可得到线弹性有限元的离散方程组:

$$\mathbf{K} \cdot \mathbf{u} = \mathbf{f}. \quad (9)$$

离散方程组可采用高斯消去法等数值方法进行求解^[75]. 从本质上说, 有限元方法求解出一组使系统总势能达到最小的节点自由度. 人工神经网络的算法则是通过迭代找到一组使得误差函数取极小值的 w 和 b . 由此可见, 有限元求解与人工神经网络算法之间存在共性. 是否可以利用人工神经网络来求解线弹性有限元问题呢?

为实现这一设想, 构造了图2所示的卷积神经网络结构. 该网络中仅包括一个输入层和一个输出层. 输入层是有限元模型的刚度矩阵 $\mathbf{K}$. 对于确定的有限

元模型, 刚度矩阵 $\mathbf{K}$ 为常数矩阵. 未知的节点位移 $\mathbf{u}$ 作为该网络的权重, 输出层是 $\mathbf{K}$ 和 $\mathbf{u}$ 依据式(10)运算得到的节点力 $\mathbf{f}$, 而标签则是有限元模型的真实节点力 $\bar{\mathbf{f}}$:

$$f_j = \sum_{i=1}^n K_{ji} \cdot u_i. \quad (10)$$

根据上述原理, 卷积神经网络算法在每次迭代中求出 $\mathbf{f}$ 与真实节点力 $\bar{\mathbf{f}}$ 的误差 $C(\mathbf{u})$. 神经网络算法根据误差函数的梯度 $\nabla C(\mathbf{u})$ 迭代更新权重 $\mathbf{u}$, 最终使得 $\mathbf{f}$ 与 $\bar{\mathbf{f}}$ 尽可能接近. 至此, 该算法求出了一组使得线弹性有限元方程(式(9))成立的位移解. 整个求解流程如图3所示.

为说明该方法的正确性, 采用了图S1所示的模型进行验证. 模型采用线弹性本构关系, 下边界位移固定, 上边界施加水平方向的均布载荷. 整个模型划

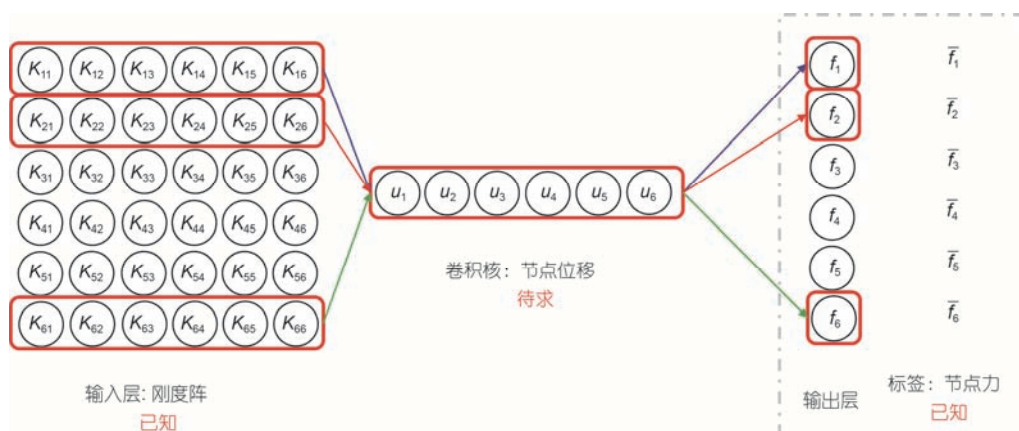


图2 (网络版彩色)用于求解线弹性有限元的卷积神经网络结构示意图

Figure 2 (Color online) The scheme of the convolutional neural network for solving a linear elastic finite element problem

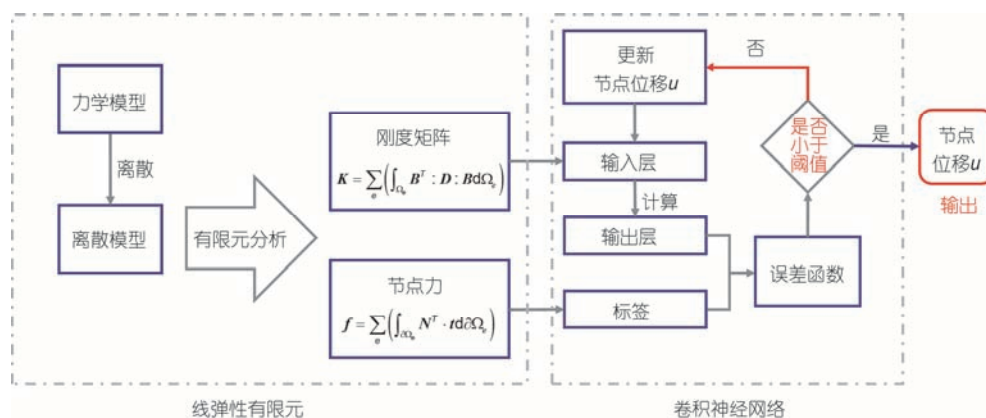


图3 (网络版彩色)使用卷积神经网络求解线弹性有限元的流程图

Figure 3 (Color online) The scheme of the convolutional neural network for solving a linear elastic finite element problem

分为100个平面应变单元。

如图3所示,根据模型几何与边界条件计算刚度矩阵 $\mathbf{K}$ 和节点力 $\bar{\mathbf{f}}$ 。结合模型的位移边界条件指定一组初始结点位移。由于节点位移的初始值不是该问题的正确解,因此初始误差较大,需要通过逐步计算误差函数对节点位移 $\mathbf{u}$ 的偏导数,并依照式(6)更新 $\mathbf{u}$ 。如图S2(a)所示,误差函数的值随着迭代次数增加不断收敛。当误差趋近于0时,即可求得问题的位移解。图S2(b)是通过若干迭步后得到的位移解所对应的变形构型。随着迭代次数增加,位移解对应的变形构型不断逼近正确解。

图S3对比了本方法和标准有限元的计算结果。图S3(a)和(b)分别描述了水平和竖直方向位移解的相对误差。本方法位移解的最大相对误差低于0.05%,从而验证了其正确性。

本节将卷积神经网络的框架与核心算法应用于线弹性有限元问题的求解中,并通过一个标准算例验证了方法的正确性。但是,仍需意识到该方法尚存的局限性。首先,该方法将方程的求解转换为优化问题,通过梯度下降算法寻找使目标函数满足的最优自由度集合。该优化过程需要经历较多次迭代,虽可通过GPU加速,其求解效率仍低于主流数值求解方法。另外,线弹性有限元方程的刚度矩阵被转换为卷积神经网络的输入层数据。卷积神经网络的框架限制了输入层数据必须为常数矩阵,这对非线性有限元方程的求解(可变刚度阵)提出了挑战。以上均为值得后续深入研究的问题。

2 基于细观页岩扫描照片的模量预测方法

在上一节中,基于算法的相似性,将人工神经网络的核心算法用于求解线弹性有限元问题。本节着重探讨卷积神经网络模型在非均匀材料研究中的应用。Radlinski等人^[22], Sahimi^[76], Milliken和Laubach^[77], Blair等人^[78], Antonellini等人^[79], Al-Raoush和Willson^[80]采用解析和数值模型研究了材料微观结构和等效材料性质之间的联系。近年来,机器学习方法被逐步引入非均质材料的研究中。Sundararaghavan和Zabaras^[81]利用支持向量机进行了3维非均质材料的分类和重建。Liu等人^[82]基于支持向量机预测了非均质材料3维模型中弹性应变场分布。Kondo等人^[83]采用卷积神经网络建立起微观结构与陶瓷离子电导率之间的映射。Cang等人^[84]基于变分自动编码器发展了一种非均质

材料的随机重建方法。Bessa等人^[85]提出了一种基于数据驱动思想的计算力学框架,涵盖了实验设计、数据库建立和机器学习模型训练3部分。该框架可以进一步用于设计非均质材料的力学响应和微观结构等。

作者基于卷积神经网络提出一个用于预测非均质材料等效力学性质的框架。为了便于阐述,考虑了细观尺度的页岩材料,使用该框架建立起页岩样本的细观结构与等效模量之间的映射关系。使用扫描电子显微镜(scanning electron microscope, SEM)获得页岩的细观结构的图像,并引入一个简化模型将SEM图像转换为5相材料样本。接下来,使用随机重建算法基于5相简化样本生成大量随机样本,再利用有限元方法计算出随机样本的等效模量。基于随机样本的细观结构和等效模量,训练了一个多层卷积神经网络,建立起细观结构到等效模量的映射关系。完成训练的网络可被用于预测真实页岩样本的等效模量。完整的流程如图4所示。

页岩约占沉积岩总量的50%^[86,87]。宏观尺度的页岩是一种深灰色致密沉积岩。图S4(a)是一块用于实验室研究的圆柱形页岩样品。自然状态下的页岩往往具有明显的层理状结构,如图S4(b)所示页岩露头的层状结构。

从细观角度看,页岩则是包含多种矿物成分的复杂异质材料,通常由石英、方解石、蒙脱石、黄铁矿、黏土和有机物等成分组成^[88]。能量色散X射线光谱(energy-dispersive X-ray spectroscopy, EDX)、聚焦离子束研磨(focused ion beam, FIB)和扫描电子显微镜等方法常被用于识别矿物成分及表征页岩样品的异质性^[89-93]。

研究发现,宏观页岩的变形、破坏等力学性质与其包含的矿物组分密切相关^[94-96]。因此,理解其细观结构与宏观力学性质之间的联系具有重要工程意义。

作者使用SEM研究了页岩样本的细观结构。页岩样品的尺寸为1.5 mm×1.5 mm。样本扫描图像如图5(a)所示。样本中包含石英、长石、黄铁矿、方解石、白云石、高岭石、伊利石和干酪根等矿物组分。各矿物组分的分布形态具有较强的随机性,造成宏观页岩样本具有显著的异质性和各向异性。

工程和地质力学通常将页岩内部具有相似材料特性的矿物组分归为一类。一种广泛接受的分类方法将各种矿物组分划分为4大类^[97]。第一类称为硅酸

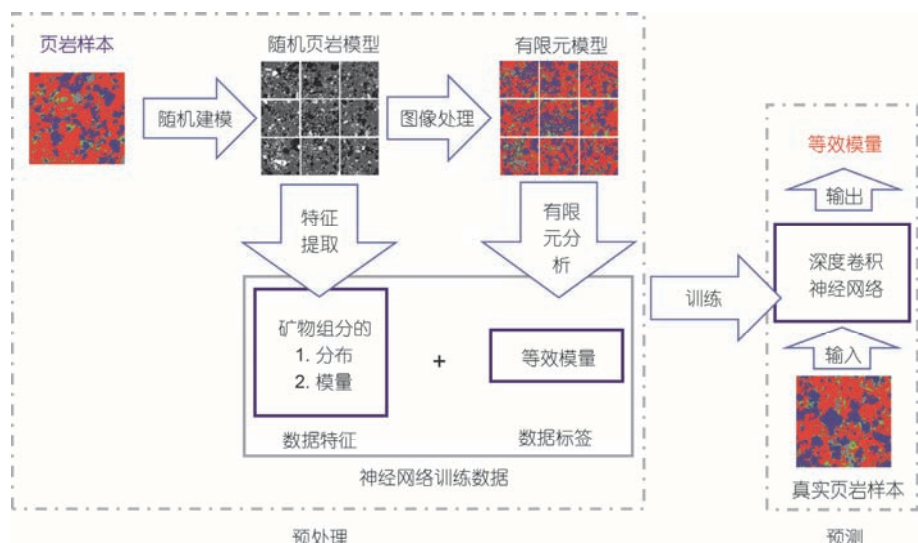


图 4 (网络版彩色)建立卷积神经网络预测页岩等效模量的工作流程

Figure 4 (Color online) The workflow of establishing a deep neural network for calculating the shale modulus

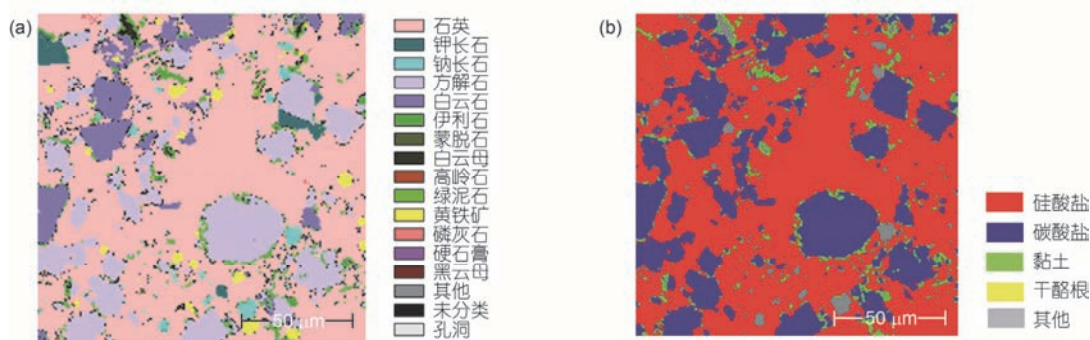


图 5 (网络版彩色)细观页岩结构图. (a) SEM照片; (b) 简化数字模型

Figure 5 (Color online) Mesoscale shale structure presentations. SEM image (a) and the corresponding simplified model (b)

盐, 主要包括石英、长石和黄铁矿等. 硅酸盐的模量和硬度最高. 第二类是黏土, 包含高岭石、伊利石、绿泥石和蒙脱石. 第三类是有机物质或干酪根, 其模量和硬度最低, 页岩气通常被吸附于其中. 最后一类包含上面未提到的其他矿物组分, 常称为基质.

考虑到中国西南地区页岩的特点, 本文增加了碳酸盐这类主要成分. 碳酸盐由方解石和白云石构成. 因此, 本文采用如表1所示的5相简化页岩模型, 将页岩的SEM扫描照片转换为如图5(b)所示的简化数字模型.

页岩的变形和破坏等力学性质与其细观矿物组分的分布方式以及各矿物组分的力学性质有关. 纳米压入实验常被用于研究页岩细观矿物组分的模量和硬度等力学性质^[98~100]. 本文使用纳米压入实验测

表 1 页岩各主要成分中包含的矿物组分

Table 1 The mineral constituents contained in the main phases of shale

主要成分	矿物组分
硅酸盐	石英、长石、黄铁矿
碳酸盐	方解石、白云石
黏土	高岭石、伊利石、绿泥石、蒙脱石
干酪根	存储页岩气的有机物成分
其他	包括其他矿物成分

定了各主要成分的模量.

图S5是各主要成分模量的测量曲线. 随着压入深度增加, 测量值受到探针周围其他成分的干扰程度不断增加^[101], 导致各成分的模量逐渐趋向于某个

中间值. 取最初测量值作为各主要成分的等效模量, 则硅酸盐、碳酸盐、黏土、干酪根和基质的等效模量分别为89.6, 65.8, 22.3, 9.2和12.3 GPa.

本文基于卷积神经网络建立页岩样本细观结构和等效模量之间的映射关系. 目前, 卷积神经网络已被成功应用于机器视觉领域. 图S6是一个基于卷积神经网络的人脸识别应用. 研究发现, 该网络中的低层神经元能够捕捉到细节面部特征; 中间层神经元可以捕捉到较大的局部面部特征; 高层神经元则能够捕捉到整体面部特征, 即高层的神经元可以提取更抽象的数据特征.

受上述机理的启发, 本文利用一个多层卷积神经网络建立起页岩细观结构与等效模量之间的映射关系, 如图6所示. 较低层的神经元可提取页岩模型中细小的矿物分布特征; 中间层神经元可提取中等尺寸的矿物分布特征; 高层神经元则记录了样本的整体结构. 最终, 通过全连接层得到样本的等效模量.

卷积神经网络的训练需要大量的图像样本, 但是, 真实页岩样本的数量相对有限. 为此, 作者利用随机建模方法创建了12000个随机样本^[102~104]. 使用其中10000个样本训练网络, 使用剩余2000个样本进行交叉验证. 图S7是2000个随机样本的等效模量分布.

训练误差和迭代次数之间的关系如图S8(a)所示. 由于权重和偏差的初始值是随机分配的, 首次迭代的误差较大. 经过几次迭代后, 训练误差迅速下降. 图S8(b)是交叉验证误差与迭代次数的关系, 误差在

50次迭代后基本收敛. 图S9是交叉验证误差的分布, 大部分误差在2%以内, 平均误差为0.55%.

完成训练后, 该网络被用来预测真实页岩样本的等效模量. 本文共使用500个真实页岩样本. 图7是部分真实页岩样本的细观结构, 真实样本中矿物成分的分布和占比具有很强的随机性. 该500个真实样本被输入训练好的网络, 网络输出其对应的等效模量. 作者对比了该网络的预测结果与有限元计算结果, 预测误差分布如图S10所示. 大部分预测误差在3%以内, 平均预测误差为0.97%.

本节提出了一种预测页岩等效模量的新方法. 该方法基于页岩SEM图像生成大量随机样本, 并将样本转换为有限元模型, 再基于有限元分析获得其等效模量. 随后, 该方法将样本的细观结构和等效模量组合为训练数据, 训练一个深度卷积神经网络. 经验证, 该网络可以高效且准确地预测页岩样本的等效模量. 另外, 该方法不局限于页岩材料, 还可以进一步用于预测其他异质材料的等效力学性质.

3 利用材料实验数据代替经验本构模型

模拟仿真及实验技术的不断发展使材料响应数据库的建立成为可能. 在这个背景下, 学者们提出了基于“数据驱动”的计算力学框架^[85, 105~109]. 例如, Yvonnet等人^[105~108]提出了基于数据驱动的材料均匀化模型; Bessa等人^[85]提出了一个数据驱动的计算框架用于设计结构和材料. 该框架提取用于描述微观

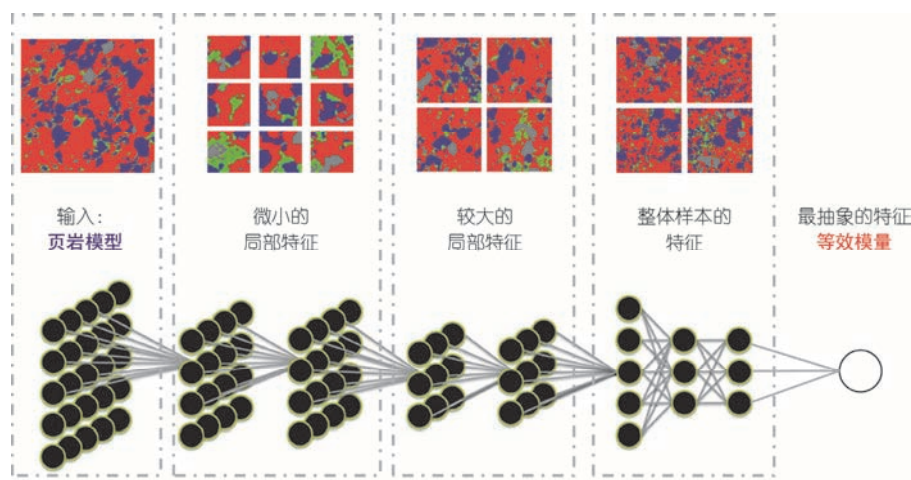


图6 (网络版彩色)用于提取细观尺度页岩样品特征的卷积神经网络结构

Figure 6 (Color online) The convolutional neural network to establish the implicit mapping between the mesoscale structure of a shale sample and its effective modulus

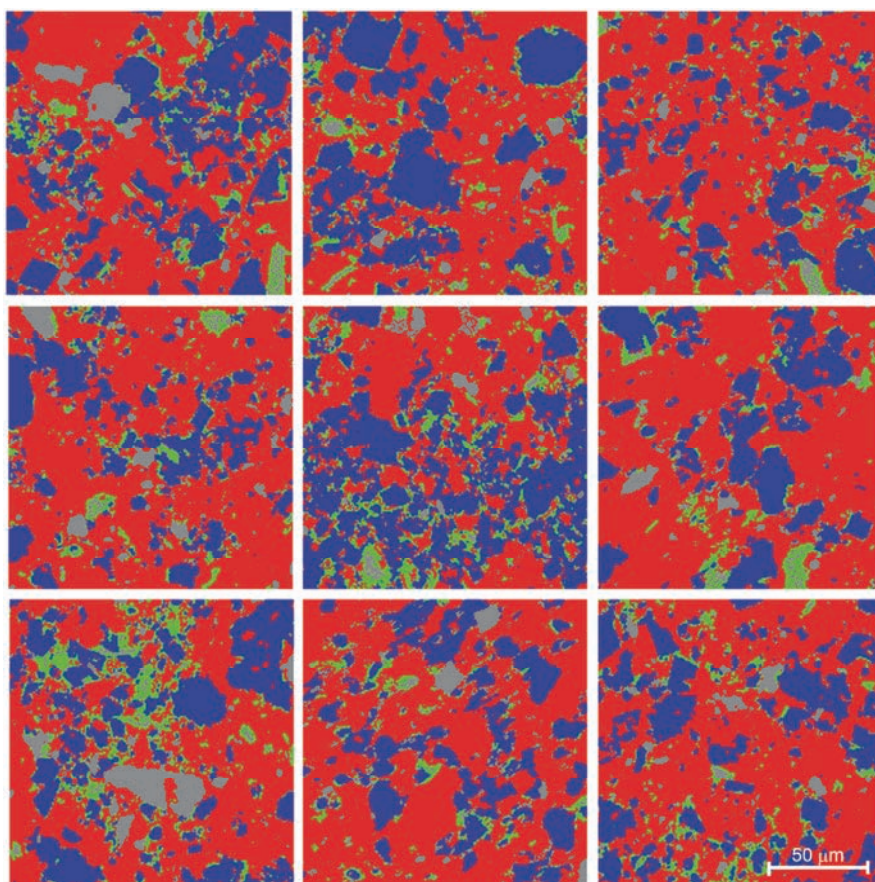


图7 (网络版彩色)部分真实页岩样本的微观结构扫描图

Figure 7 (Color online) Several real mesoscale shale samples employed for modulus prediction

结构、材料属性和边界条件的样本数据；通过计算建立其材料响应数据库；结合机器学习算法构建起样本数据的描述符与关注的材料属性之间的联系。该框架可以进一步用于设计新的结构或响应模型。Liu等人^[109]发展了一套基于数据驱动的框架，可高效准确地描述包含复杂微观结构的非弹性材料响应。

在力学研究中，一般基于实验数据建立起材料变形与内力之间的唯象本构关系。在建立本构模型时，研究者往往提出经验参数，通过构造显式函数拟合实验数据。但是，拟合出的本构函数将不可避免地丢失数据中蕴含的部分信息，引入误差和不确定性。鉴于此，Kirchdoerfer和Ortiz^[110–113]发展的连续介质计算力学框架逐步受到关注。该框架基于材料响应实测数据，结合约束条件与守恒方程进行求解，回避了传统计算力学对经验性本构关系的依赖。

该框架的核心是一个由材料响应数据驱动的求解器。该求解器在大量材料应力-应变实测数据中进

行搜索，在每个单元上迭代寻找一组满足总体优化方程应力-应变组合。

以一维桁架结构为例，该方法在每个单元上定义了如下的罚函数：

$$F_e(\varepsilon_e, \sigma_e) = \min_{(\varepsilon'_e, \sigma'_e) \in E_e} (W_e(\varepsilon_e - \varepsilon'_e) + W_e^*(\sigma_e - \sigma'_e)). \quad (11)$$

集合 F_e 代表给定的应力-应变数据的全集。 W_e 和 W_e^* 代表每个桁架单元的应变能和应变余能：

$$W_e(\varepsilon_e) = \frac{1}{2} C_e \varepsilon_e^2, \quad W_e^*(\sigma_e) = \frac{1}{2} \frac{\sigma_e^2}{C_e}. \quad (12)$$

结合平衡方程和协调约束，可以得到对于整个桁架结构的优化方程：

$$\begin{aligned} \min : & \sum_{e=1}^m w_e F_e(\varepsilon_e, \sigma_e), \text{ subject to: } \varepsilon_e \\ & = \sum_{i=1}^n B_{ei} u_i \text{ and } \sum_{e=1}^m w_e B_{ei} \sigma_e = f_i. \end{aligned} \quad (13)$$

求解器在整个桁架结构上找到一组满足优化方

程的最优应力-应变组合. 如图8所示, Kirchdoerfer和Ortiz^[110~112]利用一维和三维模型验证了该框架的正确性, 并将该框架进一步推广到处理带有噪音的数据及求解动力学问题.

Kirchdoerfer和Ortiz^[110~113]从全新的视角提出了基于实验数据的计算力学框架. 该框架不再依赖具有显式表达式的材料本构关系, 而将实验测定的材料响应数据与传统的协调方程和平衡方程等相结合进行求解, 避免了提取经验性材料本构过程中引入的不确定性, 具有很强的原创性和前沿性. 同时, 该方法也存在值得进一步研究之处. 例如, 该方法使用了类似于模拟退火的迭代求解算法, 其计算效率值得进一步探讨. 另外, 在拥有大量实验数据的情况下, 可以采用各种机器学习方法拟合出隐式的本构关系, 并结合传统的计算力学方法求解. 相比之下, 本节所述方法的优势值得进一步讨论.

4 小结与展望

机器学习和数据科学一般依靠大量数据作支撑, 通过训练提取出数据内部蕴藏的抽象映射关系, 近年来得到了长足发展, 受到计算力学领域研究者的

关注. 本文结合作者的相关研究成果, 初步介绍了机器学习、数据科学与计算力学相结合的3种形式. 这些结合形式从不同角度展现了大数据技术背景下计算力学的发展方向. 例如, 第一种结合方式直接应用卷积神经网络算法求解线性有限元方程, 但该方法的求解效率上还存在一定的提升的空间; 第二种结合方式是通过深度学习建立起材料细观结构和等效力学性质之间的映射关系. 作者应用该方法成功建立了预测细观页岩样本等效模量的卷积神经网络; 第三种结合方式是建立基于数据驱动的计算力学方法. 例如, 本文介绍了一种利用材料力学响应的实测数据代替传统本构模型的计算力学方法. 该方法提供了一种将真实材料响应数据引入计算力学模型的全新思路, 但其计算效率有待进一步检验.

虽然尚存不足, 以上工作仍在一定程度上反映出机器学习、数据驱动与计算力学方法相结合的广阔前景. 这些结合将给复杂材料和结构设计、材料本构模拟以及高效计算力学求解等领域带来巨大机遇. 随着研究者的不断开拓, 我们有理由相信未来将会出现更多的结合和应用场景, 进而发展出更加强健、高效和保真的计算力学方法.

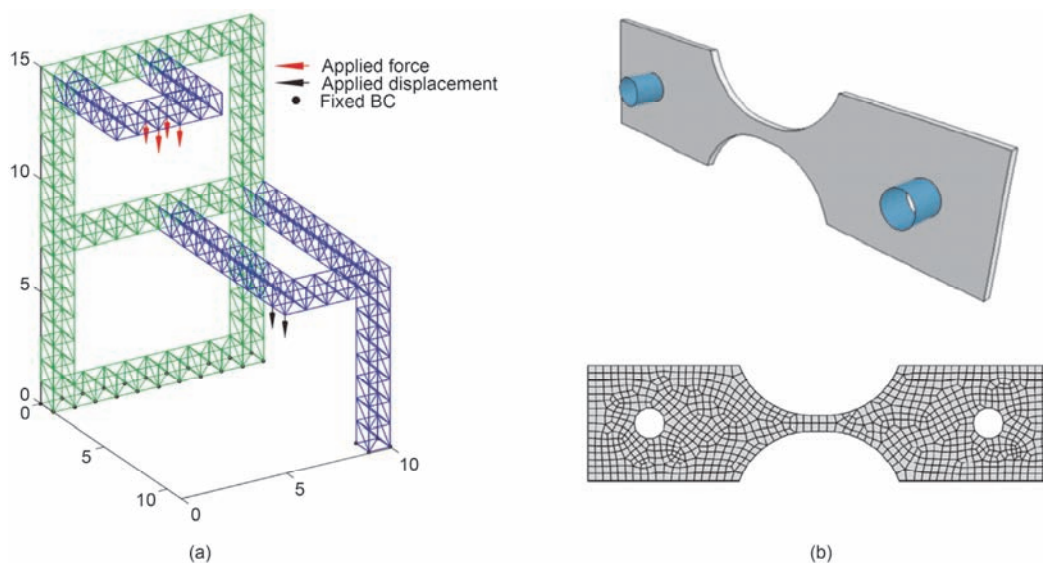


图8 (网络版彩色)基于数据驱动的计算力学方法的验证算例^[110]. (a) 一维线弹性桁架模型; (b) 三维线弹性模型

Figure 8 (Color online) The verification examples of the data-driven method^[110]. (a) One-dimensional elastic truss model; (b) three-dimensional elastic model

参考文献

- 1 Bathe K J. Finite element method. In: Wiley Encyclopedia of Computer Science and Engineering. Hoboken: John Wiley & Sons, 2007. 1–12

- 2 Belytschko T, Liu W K, Moran B, et al. *Nonlinear Finite Elements for Continua and Structures*. Hoboken: John Wiley & Sons, 2013
- 3 Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol Rev*, 1958, 65: 386
- 4 Banzhaf W, Nordin P, Keller R E, et al. *Genetic Programming: An Introduction*. San Francisco: Morgan Kaufmann, 1998
- 5 Nasrabadi N M. Pattern recognition and machine learning. *J Electron Imag*, 2007, 16: 049901
- 6 Altman N S. An introduction to kernel and nearest-neighbor nonparametric regression. *Am Stat*, 1992, 46: 175–185
- 7 Neter J, Kutner M H, Nachtsheim C J, et al. *Applied Linear Statistical Models*. Chicago: Irwin, 1996
- 8 Cox DR. The regression analysis of binary sequences. *J Royal Stat Soc B*, 1958. 215–242
- 9 Walker S H, Duncan D B. Estimation of the probability of an event as a function of several independent variables. *Biometrika*, 1967, 54: 167–179
- 10 Freund Y, Schapire R E. A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci*, 1997, 55: 119–139
- 11 Boser B E, Guyon I M, Vapnik V N. A training algorithm for optimal margin classifiers. In: *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*. New York: Association for Computing Machinery, 1992. 144–152
- 12 Cortes C, Vapnik V. Support-vector networks. *Mach Learn*, 1995, 20: 273–297
- 13 Dudani S A. The distance-weighted k -nearest-neighbor rule. *IEEE T Syst Man Cy*, 1976, 17: 325–327
- 14 Keller J M, Gray M R, Givens J A. A fuzzy k -nearest neighbor algorithm. *IEEE T Syst Man Cy*, 1985. 580–585
- 15 Wang L. *Support Vector Machines: Theory and Applications*. Netherlands: Springer Science & Business Media, 2005
- 16 Burges C J. A tutorial on support vector machines for pattern recognition. *Data Min Knowl Discov*, 1998, 2: 121–167
- 17 Suykens J A. *Advances in Learning Theory: Methods, Models, and Applications*. Leuven: IOS Press, 2003
- 18 Hastie T, Rosset S, Zhu J, et al. Multi-class adaboost. *Stat Its Interface*, 2009, 2: 349–360
- 19 McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bull Math Biophys*, 1943, 5: 115–133
- 20 Widrow B. Adaline and madaline-1963. In: *Proceedings of the IEEE First International Conference on Neural Networks*. Piscataway: IEEE, 1987, 1(Suppl): 143–157
- 21 Ivakhnenko A G. Polynomial theory of complex systems. *IEEE T Syst Man Cy*, 1971, 1: 364–378
- 22 Radlinski A, Ioannidis M, Hinde A, et al. Angstrom-to-millimeter characterization of sedimentary rock microstructure. *J Colloid Interface Sci*, 2004, 274: 607–612
- 23 Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors. *Nature*, 1986, 323: 533–536
- 24 Werbos P. *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. Doctor Dissertation. Cambridge: Harvard University, 1974
- 25 Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. *Science*, 2006, 313: 504–507
- 26 Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. In: *Proceedings of the Advances in Neural Information Processing Systems*. New York: Red Hook, 2012. 1097–1105
- 27 Schmidhuber J. Deep learning in neural networks: An overview. *Neural Netw*, 2015, 61: 85–117
- 28 Goodfellow I, Bengio Y, Courville A, et al. *Deep Learning*. Cambridge: MIT Press, 2016
- 29 Zhu D Q, Shi H. *The Principle and Application of Artificial Neural Network (in Chinese)*. Beijing: Science Press, 2006 [朱大奇, 史慧. 神经网络原理及应用. 北京: 科学出版社, 2006]
- 30 Tu J V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *J Clin Epidemio*, 1996, 49: 1225–1231
- 31 Basheer I A, Hajmeer M. Artificial neural networks: Fundamentals, computing, design, and application. *J Microbiol Meth*, 2000, 43: 3–31
- 32 Zeiler M D, Taylor G W, Fergus R. Adaptive deconvolutional networks for mid and high level feature learning. In: *Computer Vision (ICCV), 2011 IEEE International Conference on*. Piscataway: IEEE, 2011. 2018–2025
- 33 Zeiler M D, Fergus R. Visualizing and understanding convolutional networks. In: *European Conference on Computer Vision*. Cham: Springer, 2014. 818–833
- 34 Zeiler M D, Krishnan D, Taylor G W, et al. Deconvolutional networks. In: *Proceedings of the Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. Piscataway: IEEE, 2010. 2528–2535
- 35 Jin C Y, Ma Z Y, Zhang Y L, et al. Application of neural network to back analysis of mechanical parameters and initial stress field of rock masses (in Chinese). *Rock Soil Mech*, 2006, 27: 1263–1266 [金长宇, 马震岳, 张运良, 等. 神经网络在岩体力学参数和地应力场反演中的应用. 岩土力学, 2006, 27: 1263–1266]
- 36 Zhou J C, Wei Q, Liu G D. Back analysis on rock mechanics parameters for highway tunnel by BP neural network method (in Chinese). *Chin J Rock Mech Eng*, 2004, 23: 941–945 [周建春, 魏琴, 刘光栋. 采用 BP 神经网络反演隧道围岩力学参数. 岩石力学与工程学报, 2004, 23: 941–945]

- 37 Fan K. Inverse analysis of rock soil mechanics parameters based on neural network (in Chinese). *J Hohai University (Nat Sci)*, 1998, 26: 98–102 [樊琨. 基于人工神经网络的岩土工程力学参数反分析. 河海大学学报: 自然科学版, 1998, 26: 98–102]
- 38 Ge H W, Liang Y C, Liu W, et al. Application of artificial neural network and genetic algorithms to rock mechanics (in Chinese). *Chin J Rock Mech Eng*, 2004, 23: 1542–1550 [葛宏伟, 梁艳春, 刘玮, 等. 人工神经网络与遗传算法在岩石力学中的应用. 岩石力学与工程学报, 2004, 23: 1542–1550]
- 39 Feng X T, Diao X H. Intelligent rock mechanics (1)—Introduction (in Chinese). *Chin J Rock Mech Eng*, 1999, 18: 222 [冯夏庭, 刁心宏. 智能岩石力学(1)——导论. 岩石力学与工程学报, 1999, 18: 222]
- 40 Wu J G, Zhao L P. Neural network method for structural engineering optimization (in Chinese). *Chin J Comput Mech*, 1998, 15: 69–74 [吴剑国, 赵丽萍. 工程结构优化的神经网络方法. 计算力学学报, 1998, 15: 69–74]
- 41 Gao X S, Li H D. Neuro computing in structural analysis and design: An overview (in Chinese). *Chin J Comput Mech*, 2000, 17: 223–228 [高行山, 李红达. 结构分析和设计中神经网络计算研究评述. 计算力学学报, 2000, 17: 223–228]
- 42 Zhang Y M, Zhang L. Reliability-based structural optimization using neural network (in Chinese). *Chin J Comput Mech*, 2005, 22: 257–261 [张义民, 张雷. 结构系统可靠性优化设计的神经网络方法. 计算力学学报, 2005, 22: 257–261]
- 43 Xia Y Y, Xiong H F. Sensibility analysis of slope stability based on artificial neural network (in Chinese). *Chin J Rock Mech Eng*, 2004, 23: 2703–2707 [夏元友, 熊海丰. 边坡稳定性影响因素敏感性人工神经网络分析. 岩石力学与工程学报, 2004, 23: 2703–2707]
- 44 Feng X T, Wang Y J. Neural network estimation of slope stability (in Chinese). *J Eng Geol*, 1995, 3: 54–61 [冯夏庭, 王泳嘉. 边坡稳定性的神经网络估计. 工程地质学报, 1995, 3: 54–61]
- 45 Chen C Y, Wang S J, Shen X K. Predicting models to estimate stability of rock slope based on artificial neural network (in Chinese). *Chin J Geotech Eng*, 2001, 23: 157–161 [陈昌彦, 王思敬, 沈小克. 边坡岩体稳定性的人工神经网络预测模型. 岩土工程学报, 2001, 23: 157–161]
- 46 Lei X, Liu C, Du Z, et al. Machine learning-driven real-time topology optimization under moving morphable component-based framework. *J Appl Mech*, 2019, 86: 011004
- 47 Ohsaki M. Genetic algorithm for topology optimization of trusses. *Comput Struct*, 1995, 57: 219–225
- 48 Hajela P, Lee E. Genetic algorithms in truss topological optimization. *Int J Solids Struct*, 1995, 32: 3341–3357
- 49 Ghaboussi J, Pecknold D A, Zhang M, et al. Autoprogressive training of neural network constitutive models. *Int J Numer Meth Eng*, 1998, 42: 105–126
- 50 Jung S, Ghaboussi J. Neural network constitutive model for rate-dependent materials. *Comput Struct*, 2006, 84: 955–963
- 51 Ji G, Li F, Li Q, et al. A comparative study on arrhenius-type constitutive model and artificial neural network model to predict high-temperature deformation behaviour in aermet100 steel. *Mat Sci Eng A*, 2011, 528: 4774–4782
- 52 Furukawa T, Yagawa G. Implicit constitutive modelling for viscoplasticity using neural networks. *Int J Numer Meth Eng*, 1998, 43: 195–219
- 53 Hashash Y, Jung S, Ghaboussi J. Numerical implementation of a neural network based material model in finite element analysis. *Int J Numer Meth Eng*, 2004, 59: 989–1005
- 54 Sun Y, Zeng W, Zhao Y, et al. Development of constitutive relationship model of ti600 alloy using artificial neural network. *Comp Mater Sci*, 2010, 48: 686–691
- 55 Faller W E, Schreck S J. Unsteady fluid mechanics applications of neural networks. *J Aircraft*, 1997, 34: 48–55
- 56 Wang D, Liao W. Modeling and control of magnetorheological fluid dampers using neural networks. *Smart Mater Struct*, 2004, 14: 111
- 57 Zeng Y H, Huai W X. Application of artificial neural network to predict the friction factor of open channel flow. *Commun Nonlinear Sci*, 2009, 14: 2373–2378
- 58 Butz T, Von Stryk O. Modelling and simulation of electro- and magnetorheological fluid dampers. *J Appl Math Mech*, 2002, 82: 3–20
- 59 Beigzadeh R, Rahimi M. Prediction of heat transfer and flow characteristics in helically coiled tubes using artificial neural networks. *Int Commun Heat Mass*, 2012, 39: 1279–1285
- 60 Mi Y, Ishii M, Tsoukalas L. Flow regime identification methodology with neural networks and two-phase flow models. *Nucl Eng Des*, 2001, 204: 87–100
- 61 Dissanayake M, Phan-Thien N. Neural-network-based approximations for solving partial differential equations. *Commun Numer Meth Eng*, 1994, 10: 195–201
- 62 Takeuchi J, Kosugi Y. Neural network representation of finite element method. *Neural Netw*, 1994, 7: 389–395
- 63 Lagaris I E, Likas A, Fotiadis D I. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Trans Neural Netw*, 1998, 9: 987–1000
- 64 Ramuhalli P, Udpa L, Udpa S S. Finite-element neural networks for solving differential equations. *IEEE Trans Neural Netw*, 2005, 16: 1381–1392

- 65 Schalkoff R J. Artificial Neural Networks. New York: McGraw-Hill, 1997
- 66 Venkatesan R, Li B. Convolutional Neural Networks in Visual Computing: A Concise Guide. Boca Raton: CRC Press, 2017
- 67 Jain A K, Mao J, Mohiuddin K M. Artificial neural networks: A tutorial. *Computer*, 1996, 29: 31–44
- 68 Fukushima K, Miyake S. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In: *Competition and Cooperation in Neural Nets*. Berlin: Springer, 1982. 267–285
- 69 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, 521: 436–444
- 70 LeCun Y, Bengio Y. Convolutional networks for images, speech, and time series. *The Handbook of Brain Theory and Neural Networks*. Cambridge: MIT Press, 1995
- 71 Lawrence S, Giles C L, Tsoi A C, et al. Face recognition: A convolutional neural-network approach. *IEEE Trans Neural Netw*, 1997, 8: 98–113
- 72 Karpathy A, Toderici G, Shetty S, et al. Large-scale video classification with convolutional neural networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Washington D C: IEEE Computer Society, 2014. 1725–1732
- 73 Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of go without human knowledge. *Nature*, 2017, 550: 354–359
- 74 Silver D, Hubert T, Schrittwieser J, et al. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 2018, 362: 1140–1144
- 75 Zienkiewicz O C, Taylor R L, Zienkiewicz O C, et al. *The Finite Element Method*. London: McGraw-hill, 1977
- 76 Sahimi M. *Flow and Transport in Porous Media and Fractured Rock: From Classical Methods to Modern Approaches*. Hoboken: John Wiley & Sons, 2011
- 77 Milliken K L, Laubach S E. Brittle deformation in sandstone diagenesis as revealed by scanned cathodoluminescence imaging with application to characterization of fractured reservoirs. In: *Cathodoluminescence in Geosciences*. Berlin: Springer, 2000. 225–243
- 78 Blair S C, Berge P A, Berryman J G. Using two-point correlation functions to characterize microgeometry and estimate permeabilities of sandstones and porous glass. *J Geophys Res Solid Earth*, 1996, 101: 20359–20375
- 79 Antonellini M, Aydin A, Pollard D, et al. Petrophysical study of faults in sandstone using petrographic image analysis and X-ray computerized tomography. *Pure Appl Geophys*, 1994, 143: 181–201
- 80 Al-Raoush R, Willson C. Extraction of physically realistic pore network properties from three-dimensional synchrotron X-ray microtomography images of unconsolidated porous media systems. *J Hydrol*, 2005, 300: 44–64
- 81 Sundararaghavan V, Zabaras N. Classification and reconstruction of three-dimensional microstructures using support vector machines. *Comp Mater Sci*, 2005, 32: 223–239
- 82 Liu R, Yabansu Y C, Agrawal A, et al. Machine learning approaches for elastic localization linkages in high-contrast composite materials. *Integr Mater Manuf Innov*, 2015, 4: 13
- 83 Kondo R, Yamakawa S, Masuoka Y, et al. Microstructure recognition using convolutional neural networks for prediction of ionic conductivity in ceramics. *Acta Mater*, 2017, 141: 29–38
- 84 Cang R, Li H, Yao H, et al. Improving direct physical properties prediction of heterogeneous materials from imaging data via convolutional neural network and a morphology-aware generative model. *Comput Mater Sci*, 2018, 150: 212–221
- 85 Bessa M, Bostanabad R, Liu Z, et al. A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. *Comput Method Appl M*, 2017, 320: 633–667
- 86 Pettijohn F J. *Sedimentary Rocks*. New York: Harper & Brothers, 1957
- 87 Boggs S. *Petrology of Sedimentary Rocks*. Cambridge: Cambridge University Press, 2009
- 88 Lee S, Hyder L, Alley P. Microstructural and mineralogical characterization of selected shales in support of nuclear waste repository studies. In: *Microstructure of Fine-Grained Sediments*. Berlin: Springer, 1991. 545–560
- 89 Bernard S, Wirth R, Schreiber A, et al. FIB-SEM and TEM investigations of an organic-rich shale maturation series from the lower toarcian posidonia shale, germany: Nanoscale pore system and fluid-rock interactions. In: *Electron Microscopy of Shale Hydrocarbon Reservoirs: AAPG Memoir*. Tulsa: American Association of Petroleum Geologists, 2013. 102: 53–66
- 90 Ohkouchi N, Kuroda J, Okada M, et al. Why cretaceous black shales have high c/n ratios: Implications from SEM-EDX observations for livello bonarelli black shales at the cenomanian-turonian boundary. *Front Res Earth Evol*, 2003, 1: 239–241
- 91 Abedi S, Slim M, Hofmann R, et al. Nanochemo-mechanical signature of organic-rich shales: A coupled indentation-EDX analysis. *Acta Geotech*, 2016, 11: 559–572
- 92 Kelly S, El-Sobky H, Torres-Verdín C, et al. Assessing the utility of FIB-SEM images for shale digital rock physics. *Adv Water Resour*, 2016, 95: 302–316
- 93 Tahmasebi P, Javadpour F, Sahimi M. Three-dimensional stochastic characterization of shale SEM images. *Transport Porous Med*, 2015, 110: 521–531

- 94 Foster C D, Nejad T M. Embedded discontinuity finite element modeling of fluid flow in fractured porous media. *Acta Geotech*, 2013, 8: 49–57
- 95 Barthélemy J F, Souque C, Daniel J M. Nonlinear homogenization approach to the friction coefficient of a quartz-clay fault gouge. *Int J Numer Anal Meth*, 2013, 37: 1948–1968
- 96 White J A. Anisotropic damage of rock joints during cyclic loading: Constitutive framework and numerical integration. *Int J Numer Anal Meth*, 2014, 38: 1036–1057
- 97 Kohli A H, Zoback M D. Frictional properties of shale reservoir rocks. *J Geophys Res Solid Earth*, 2013, 118: 5109–5125
- 98 Deirieh A, Ortega J, Ulm F J, et al. Nanochemomechanical assessment of shale: A coupled wds-indentation analysis. *Acta Geotech*, 2012, 7: 271–295
- 99 Kumar V. Geomechanical characterization of shale using nano-indentation. Doctor Dissertation. Norman: University of Oklahoma, 2012
- 100 Gathier B. Multiscale strength homogenization: Application to shale nanoindentation. Doctor Dissertation. Cambridge: Massachusetts Institute of Technology, 2008
- 101 Bennett K C, Berla L A, Nix W D, et al. Instrumented nanoindentation and 3d mechanistic modeling of a shale at multiple scales. *Acta Geotech*, 2015, 10: 1–14
- 102 Torquato S. Statistical description of microstructures. *Annu Rev Mater Res*, 2002, 32: 77–111
- 103 Yeong C, Torquato S. Reconstructing random media. *Phys Rev E*, 1998, 57: 495
- 104 Cule D, Torquato S. Generating random media from limited microstructural information via stochastic optimization. *J Appl Phys*, 1999, 86: 3428–3437
- 105 Le B, Yvonnet J, He Q C. Computational homogenization of nonlinear elastic materials using neural networks. *Int J Numer Meth Eng*, 2015, 104: 1061–1084
- 106 Yvonnet J, Monteiro E, He Q C. Computational homogenization method and reduced database model for hyperelastic heterogeneous structures. *Int J Multiscale Com*, 2013, 11
- 107 Clément A, Soize C, Yvonnet J. Computational nonlinear stochastic homogenization using a nonconcurrent multiscale approach for hyperelastic heterogeneous microstructures analysis. *Int J Numer Meth Eng*, 2012, 91: 799–824
- 108 Yvonnet J, Gonzalez D, He Q C. Numerically explicit potentials for the homogenization of nonlinear elastic heterogeneous materials. *Comput Method Appl M*, 2009, 198: 2723–2737
- 109 Liu Z, Bessa M, Liu W K. Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials. *Comput Method Appl M*, 2016, 306: 319–341
- 110 Kirchdoerfer T, Ortiz M. Data-driven computational mechanics. *Comput Method Appl M*, 2016, 304: 81–101
- 111 Kirchdoerfer T, Ortiz M. Data driven computing with noisy material data sets. *Comput Method Appl M*, 2017, 326: 622–641
- 112 Kirchdoerfer T, Ortiz M. Data-driven computing in dynamics. *Int J Numer Meth Eng*, 2018, 113: 1697–1710
- 113 Kirchdoerfer T, Ortiz M. Data-driven computing. In: *Advances in Computational Plasticity*. Berlin: Springer, 2018. 165–183

补充材料

- 图 S1 验证算例: 平面应变线弹性模型
- 图 S2 误差和变形构型分析
- 图 S3 水平和竖直方向位移的相对误差分析
- 图 S4 实验室宏观页岩样本和页岩露头的层状结构
- 图 S5 各主要矿物成分的测定模量曲线
- 图 S6 用于人脸识别应用的深度卷积神经网络结构
- 图 S7 2000 个随机页岩样品的模量分布
- 图 S8 训练误差和交叉验证误差与迭代次数的关系
- 图 S9 交叉验证的误差分布图
- 图 S10 真实页岩样品的测试误差分布图

本文以上补充材料见网络版 csb.scichina.com. 补充材料为作者提供的原始数据, 作者对其学术质量和内容负责.

Summary for “机器学习与计算力学的结合及应用初探”

Combination and application of machine learning and computational mechanics

Xiang Li, Ziming Yan & Zhanli Liu*

Applied Mechanics Laboratory, School of Aerospace Engineering, Tsinghua University, Beijing 100084, China

* Corresponding author, E-mail: liuzhanli@tsinghua.edu.cn

With the steady development of computer science, machine learning and data science have made significant progress in recent decades. These techniques generally rely on a substantial amount of data samples to extract the abstract mapping hidden within the data. Hence, these technologies have gradually attracted the attention of researchers in the field of computational mechanics. Combining the recent studies of the authors and other researchers, this paper aims to interpret several forms of applications that integrate machine learning and data science with computational mechanics. In the first application, the core algorithm of the convolutional neural network is implemented to solve the linear elastic finite element problem. A standard finite element equation is transformed into an optimization problem in this method. The method is verified by a plane strain linear elastic finite element problem. The method demonstrates promising accuracy by comparing the results obtained by traditional finite element solver. However, some limitations of this method need to be addressed. First, though the optimization process can be accelerated by GPU, the efficiency of the proposed method is still lower than most mainstream numerical solvers. And, the framework of convolutional neural networks requires that the input layer data should be a constant matrix. This is a major challenge for solving nonlinear finite element equations whose stiffness matrices contain variables. These are the issues worth considerations in future studies. In the second application, a method is proposed to establish the implicit mapping between the effective mechanical property and the mesoscale structure of heterogeneous materials. Shale is employed in this paper as an example to illustrate the method. At the mesoscale, a shale sample is a complex heterogeneous composite that consists of multiple mineral constituents. The mechanical properties of each mineral constituent vary significantly, and mineral constituents are distributed in an utterly random manner within shale samples. Large quantities of shale samples are generated based on mesoscale scanning electron microscopy images using a stochastic reconstruction algorithm. Image processing techniques are employed to transform the shale sample images to finite element models. Finite element analysis is utilized to evaluate the effective mechanical properties of the shale samples. A convolutional neural network is trained based on the images of stochastic shale samples and their effective moduli. The trained network is validated to be able to predict the effective moduli of real shale samples accurately and efficiently. Not limited to shale, the proposed method can be further extended to predict effective mechanical properties of various heterogeneous materials. In the third application, the authors discuss a data-driven computational mechanics framework proposed by Kirchdoerfer and Ortiz. The most outstanding feature of the framework is that explicit material constitutive equations are no longer required. More specifically, experimental material response data are employed in the framework to replace constitutive equations. Combined with traditional compatibility and equilibrium equations, the framework is able to find the optimal stress-strain combination from a material response dataset to best fit the current element. With this framework, the errors and uncertainties induced by the empirical constitutive functions of traditional computational mechanics approaches can be avoided. The aforementioned applications are only the tip of an iceberg in the recent advancement of computational mechanics. Hence, researchers have reasons to believe that there would be more application scenarios that integrate data science and machine learning with computational mechanics in the future. Hopefully, computational mechanics methods with more robustness, efficiency, and fidelity will be developed.

machine learning, data-driven, artificial neural network, computational mechanics, finite element method

doi: 10.1360/N972019-00005