

# Some Developments of Estimation Procedures of Mechanical Model Parameter Based on Deep Learning Neural Network

Shouju Li

Department of Engineering Mechanics, Dalian University of Technology, Dalian Liaoning  
Email: lishouju@dlut.edu.cn

Received: Apr. 16<sup>th</sup>, 2020; accepted: May 1<sup>st</sup>, 2020; published: May 8<sup>th</sup>, 2020

---

## Abstract

In order to estimate mechanical model parameters of materials, the basic scheme and algorithm of parameter inversion based on classical BP neural network are discussed. Some drawbacks and improvement methods for classical BP neural network are analyzed. Combined two examples of estimating model parameters of electric-chemical model of Li-ion cells, the basic thinking of estimating model parameter procedure based on deep learning neural network is introduced. How to determine hyper parameters of deep learning neural network is discussed. How to improve learning efficiency and generalization ability of deep learning neural network is developed to estimate model parameters of materials.

## Keywords

Deep Learning Neural Network, Model Parameter Estimation Procedure, Some Development, Back-Propagation Neural Network, Hyper Parameter, Parameter Optimization

---

# 深度学习神经网络在力学模型参数估计中的应用研究进展

李守巨

大连理工大学工程力学系, 辽宁 大连  
Email: lishouju@dlut.edu.cn

收稿日期: 2020年4月16日; 录用日期: 2020年5月1日; 发布日期: 2020年5月8日

## 摘要

为了估计岩土材料的力学模型参数,讨论了基于经典BP神经网络的参数反演方法的基本框架和算法。分析了经典BP神经网络所存在的某些缺陷及其改进方法。结合两个基于深度学习神经网络估计锂离子电池电化学模型参数的例子,介绍了基于深度学习神经网络估计模型参数的基本思路。讨论了深度学习神经网络超参数确定方法,分析了提高深度学习神经网络学习效率和泛化能力的某些行之有效的策略。

## 关键词

深度学习神经网络, 模型参数估计, 研究进展, BP神经网络, 超参数, 参数优化

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在进行岩土边坡、地下隧道、地基基础、水电站大坝安全评估分析中,往往需要计算结构的应力场和变形场。有限元数值计算模拟方法为精确定量描述结构的应力分布提供了强有力的计算手段。当计算模型材料的本构模型确定之后,如何客观准确确定估计材料的本构模型参数是极为重要的。无论有限元数值计算方法多么精确,如果材料力学模型参数估计不准确,所得到的计算结果也无任何工程意义。Tarantola指出,最小二乘法是求解模型参数估计问题最普遍、最简易的方法,其缺点在于该方法缺少鲁棒性,也就是对观测数据的误差及其敏感,也就是反问题解的稳定性问题[1]。为了得到改性沙土的Duncan-Chang弹性非线性本构模型参数,Li采用遗传算法、基于不同围压条件下沙土三轴压缩应力-应变曲线,得到了沙土的力学模型参数[2]。李守巨采用经典BP神经网络,基于堆石料的三轴压缩实验曲线,进行了堆石料本构模型参数估计研究[3]。经典BP神经网络的缺陷在于泛化能力较差、其目标函数往往可能搜索到局部极小值,而不是全局最小值。为了解决这些问题,Lecun(2015)提出了深度学习神经网络模型,对神经网络的发展做出了创新性贡献[4]。深度学习神经网络是重叠了许多隐含层的一种神经网络,使得其具有更好的学习泛化能力。卷积神经网络是深度学习神经网络中的一个典型代表,其隐含层是由多个卷积层和隐含层组成的。山下隆义深入讨论了深度学习神经网络输入数据的预处理方法和提高神经网络泛化能力的方法[5]。涌井良幸推导了深度学习神经网络的权值计算数学模型[6]。斋藤康毅基于Python理论,深入研究了深度学习神经网络的算法和程序实现[7]。Chun基于卷积神经网络、采用数值模拟方法研究了锂电池电化学模型参数估计问题[8]。Shen基于深度学习神经网络研究了锂电池电化学模型参数在线估计问题[9]。Kazim基于深度学习神经网络研究了太阳能辐射模型参数估计问题[10]。本文的目的在于讨论基于经典BP神经网络的参数反演方法的基本框架和算法。总结基于经典BP神经网络进行模型参数估计的主要步骤。分析经典BP神经网络所存在的某些缺陷及其改进方法。介绍基于深度学习神经网络估计模型参数的基本思路。讨论了深度学习神经网络超参数确定方法。

## 2. 经典BP神经网络在力学模型参数估计中应用

在岩土工程数值模拟分析中,首先需要确定岩土材料本构模型中的力学参数,经典BP神经网络某

些为解决这些问题提供了强有力的技术手段。岩土材料是一种典型的非线性变形材料，其常用的岩土本构模型包括剑桥粘土模型(Cam-Clay)，此模型非常适用于正常固结粘土和弱固结粘土。Duncan-Chang 双曲线模型是一种非线性弹性模型，包括 E-B 模型和 E- $\mu$  模型。Lade-Duncan 弹塑性模型是根据对砂土的真三轴试验结果，提出的适用于表征砂土类的真三轴压缩实验数据。清华弹塑性模型是另一种弹塑性模型。基于 Janbu (1963)提出的经验公式，采用 Mohr-Coulomb 破坏准则，Duncan 和 Chang (1970)提出一个土体非线性双曲线本构模型，即著名的 Duncan-Chang 弹性非线性本构模型，与围压和土体抗剪指标相关的土体切线模量为：

$$E_t = kP_a \left( \frac{\sigma_3}{P_a} \right)^n \left[ 1 - \frac{R_f (\sigma_1 - \sigma_3)(1 - \sin \phi)}{2c \cos \phi + 2\sigma_3 \sin \phi} \right]^2 \quad (1)$$

式中  $c$  为土体的粘聚力， $\phi$  为土体的内摩擦角， $R_f$  为破坏比， $P_a$  是用与  $\sigma_3$  和  $E_t$  相同的单位表示的大气压强， $K$ 、 $n$  为由实验确定的常数。因此，被估计的与土体变形模量相关的参数有 5 个，即  $K$ ， $n$ ， $R_f$ ， $c$ ， $\phi$ ，5 个参数。土体非线性弹性本构模型中含有两个独立的变量，一般经常采用切线弹性模量和和切线泊松比进行表征。因此，土体的另一个参数就是切线泊松比，它也是围压的函数。1980 年 Duncan 通过用切线体积模量代替切线泊松比，提出了 E-B 模型：

$$B_t = K_b P_a \left( \frac{\sigma_3}{P_a} \right)^m \quad (2)$$

与土体切线体积模量相关的参数有 2 个，即  $m$ ， $R_b$ ，它们也是由实验确定。这样，与 Duncan-Chang 双曲线模型相关的被估计参数一共有 7 个。

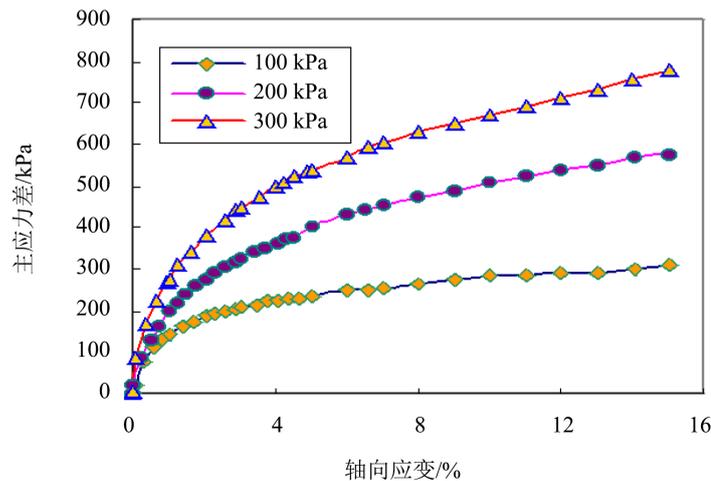


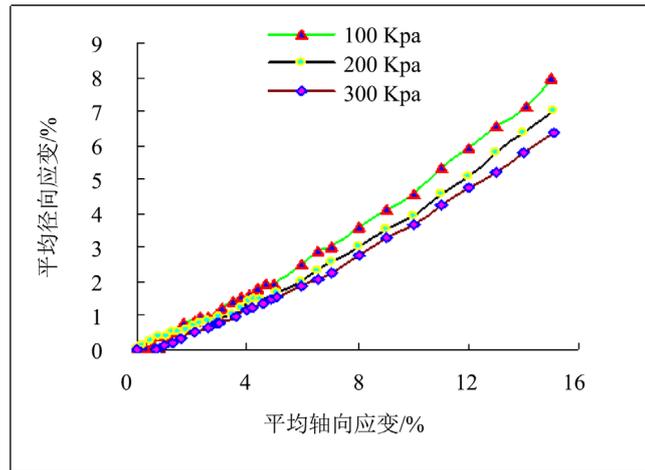
Figure 1. Relationship between stress and axial strain in different confined pressure

图 1. 不同围压条件下沙土三轴压缩应力 - 应变曲线

改性沙土在不同围压条件下三轴压缩应力 - 应变实验曲线、三轴压缩径向应变 - 轴向应变曲线，如图 1 和图 2 所示[2]。这两组曲线是进行沙土 Duncan-Chang 弹性非线性本构模型参数估计的基础。

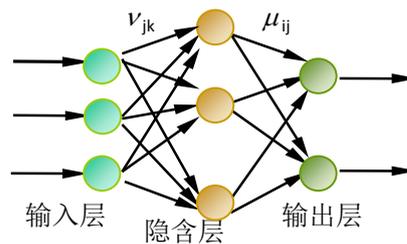
基于图 1 和图 2 的沙土应力和应变观测数据，建立了如图 3 所示的神经网络模型，用于反演估计沙土的 Duncan-Chang 弹性非线性本构模型中的 7 个参数。其中，神经网络输入层神经元的个数与图 1 和图 2 离散点的个数相一致， $N_i = 30 \times 3 \times 2 = 180$ 。隐含层神经元的个数取为  $N_h = 360$  个，隐含层神经元的作

用函数采用 Sigmoid 函数。输出层神经元的个数与被估计力学模型参数的个数相一致,  $N_o = 7$ 。输出层神经元的作用函数采用线性函数。



**Figure 2.** Relationship between radial and axial strains in different confined pressure

**图 2.** 不同围压条件下沙土三轴压缩径向应变 - 轴向应变曲线



**Figure 3.** Topology structure of BP neural network for estimating model parameters

**图 3.** BP 神经网络的拓扑结构

基于前人已有的关于沙土力学模型参数研究的先验信息, 初步估计 Duncan-Chang 本构模型中的 7 个参数可能的区间范围。采用 MATLAB 软件, 基于沙土的 Duncan-Chang 弹性非线性本构模型, 以主应力差和轴向应变增量形式给出, 随机建立了 90 个训练样本, 即神经网络样本集; 10 个测试样本, 即测试集。建立训练样本时可以认为, 力学模型参数在可能区间符合均匀分布或者状态分布。神经网络训练时, 定义如下均方误差(mean squared error)为目标函数

$$E(w) = \frac{1}{2} \sum_{p=1}^P \left( \sum_{i=1}^{N_o} (y_i^p - o_i^p)^2 \right) \quad (3)$$

式中  $y_i^p$  和  $o_i^p$  分别为神经网络输出层单元的实际输出和期望输出,  $P$  是训练样本的数目。在神经网络模型中经常采用的作用函数, 包括 Sigmoid 作用函数

$$f(\text{Sigma}) = 1 / [1 + \exp(-\text{Sigma})] \quad (4)$$

式中  $\text{Sigma}$  表示神经元的输入求和,  $f$  表示神经元的输出。

**ReLU 作用函数:** 深度学习神经网络除了传统的 sigmoid 函数之外, 使用了 ReLU (Rectified Linear Unit) 函数。当输入大于 0 时, 该函数的输出就是输入; 而当输入小于 0 时, 该函数的输出是 0。

$$f(\text{Sigma}) = \max(\text{Sigma}, 0) \quad (5)$$

双曲正切函数(Tanh)作用函数

$$f(\text{Sigma}) = \frac{\exp(\text{Sigma}) - \exp(-\text{Sigma})}{\exp(\text{Sigma}) + \exp(-\text{Sigma})} \quad (6)$$

线性作用函数

$$f(\text{Sigma}) = \text{Sigma} \quad (7)$$

输出层神经网络任意神经元的输出表示为

$$y_i = f_o \left( \sum_{j=1}^{N_h} (\mu_{ij} z_j + \lambda_i) \right) \quad (8)$$

式中  $m_{ij}$  为连接隐含层与输出层的权值;  $l_i$  为输出层的阈值,  $f_o$  为输出层神经元的作用函数;  $z$  为隐含层的输出,  $N_h$  为隐含层神经元的个数。隐含层任意神经元的输出表示为

$$z_j = f_h \left( \sum_{k=1}^{N_i} (v_{jk} x_k + \eta_k) \right) \quad (9)$$

式中  $v_{jk}$  为连接隐含层与输入层的权值;  $x_k$  为隐含层的阈值,  $f_h$  为隐含层神经元的作用函数。基于经典 BP 神经网络进行模型参数估计的步骤如下:

1) 基于先验信息, 确定被估计力学模型参数的大致范围; 随机产生训练样本集和测试集。当采用 sigmoid 作用函数时, 输入数据需要进行归一化处理。输入数据的正规化(归一化 normalization)处理, 使得输入数据的值介于[-1,1]之间。

$$x_{kn} = 2 \left\{ \frac{(x_k - x_{\min})}{(x_{\max} - x_{\min})} \right\} - 1 \quad (10)$$

式中  $x_{kn}$  为归一化后的神经网络输入。

2) 基于 Duncan-Chang 弹性非线性本构模型, 确定将要模拟的围压值, 解析计算在不同力学模型参数随机组合条件下的应力 - 应变曲线, 类似于图 1 和图 2;

$$\Delta \sigma_d = E_t \Delta \varepsilon_a \quad (11)$$

式中  $\Delta \sigma_d$  为偏应力增量,  $\Delta \varepsilon_a$  为轴向应变增量。径向应变增量与轴向应变的关系为

$$\Delta \varepsilon_r = \mu_t \Delta \varepsilon_a \quad (12)$$

式中  $\Delta \varepsilon_r$  为径向应变增量。

3) 建立 BP 神经网络模型, 确定神经网络的拓扑结构, 随机给神经网络权值和阈值赋初始值, 设定神经网络收敛准则;

4) 将训练样本集中应力 - 应变数据作为神经网络的输入, 而力学模型中的参数作为神经网络的期望输出。基于 MATLAB 软件或者其他软件平台, 采用优化方法训练神经网络, 例如 Levenberg-Marquardt 优化算法, Quasi-Newton 优化算法, 直至神经网络收敛, 记录神经网络的权值和阈值;

5) 将测试集输入到已经训练好的神经网络, 检验神经网络的泛化能力和预测精度。

6) 室内进行土体三轴压缩实验, 记录不同围压条件下的应力 - 应变曲线。将这些观测数据代入到训练好的神经网络模型, 神经网络的输出就是力学模型的参数, 具体的实施步骤和方法详见参考文献[3]。

尽管经典 BP 神经网络在模型参数估计问题中得到了广泛的应用, 但是其存在着许多固有的缺陷。其一, 目标函数可能无法收敛到全局最优解。根据神经网络目标函数的定义可以看出, 实际上, 神经网络目标函数寻优过程, 就是一个关于权值极小化的广义最小二乘法。基于梯度下降法的神经网络训练学

习过程，无法避开局部极小值问题。其二，超参数的选取方法缺少理论依据。其三，隐含层的层数和单元个数确定缺少理论依据。其四，较差的神经元之间初始权值，可能无法得到理想的结果。其五，在某些情况下，会出现梯度消失问题，无法继续进行迭代。

### 3. 深度学习神经网络在模型参数估计中应用

卷积神经网络模拟了生物的视觉功能，是一类包含卷积计算，具有深度结构的前馈神经网络，采用有监督和非监督学习方法完成网络的学习训练，利用逐层推理方法实现自我学习和推理的功能，克服深层结构训练学习的困难[4]。卷积神经网络是深度神经网络的一个典型代表。深度神经网络已经被广泛应用于科学建模[11]、模型参数估计[12]、时变系统的荷载参数识别[13]、工业产品特征提取[14]、故障诊断[15]等工程领域。卷积神经网络主要是针对二维图像识别而开发的，它的输入常常是二维图片。而岩土材料力学模型参数估计用到的数据是主应力差随轴向应变变化的一维数据系列，如图1和图4所示。因此，需要研究一维卷积神经网络的特性和处理方法。金列俊基于一维卷积神经网络模型，提出了钻杆故障诊断方法，实验验证了所提出方法的有效性、适应性和抗噪能力[16]。朱锡祥开发了基于一维卷积神经网络的辨识方法，用于车载语音识别研究[17]，韩林洁研究了基于一维卷积神经网络预测模型，进行轴承剩余寿命预测，并且进行了实验验证[18]。

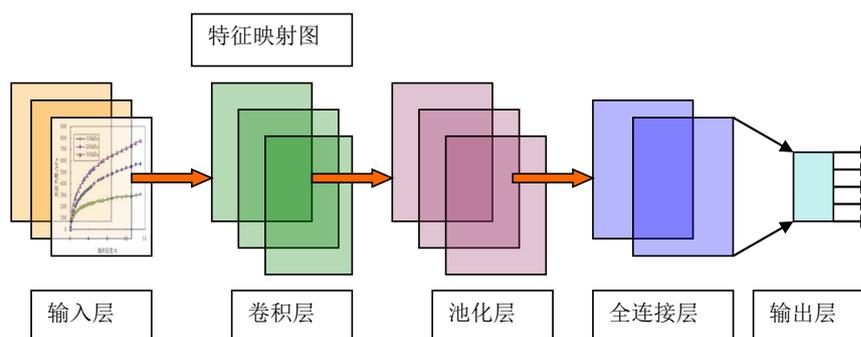


Figure 4. Topology structure of convolutional neural network for estimating model parameters

图4. 用于模型参数估计的卷积神经网络的拓扑结构

如图4所示，用于模型参数估计的卷积神经网络包括输入层、卷积层、池化层、全连接层和输出层组成。卷积神经网络是通过模仿动物或者人类的视觉原理来实现图像分类、识别的目的。其中，输入层输入的是多个图像信息，例如一幅  $256 \times 256$  像素的图像。输入层可以处理多维数据，例如一维数组的时间序列，平面上的二位像素点。卷积层可以看作作为输入样本与卷积核的内积运算，它应用了一个核函数，然后，经过作用函数计算后，得到特征映射图。卷积神经网络常用的作用函数包括 sigmoid、tanh、ReLU 等。卷积不是将函数作用在整个图像上，而是一次扫描图像的一个小窗口。卷积核的功能是对输入数据进行特征提取。池化层的作用在于减少卷积层产生特征图的尺寸。常用的池化方法包括，最大池化、平均池化以及 LP 池化。其中，LP 池化是通过突出图像区域内的中央值而计算出新的特征图。最大值池化操作是在扫描区域中从每个通道中选取最大值。平均值池化操作使用平均值代替整个扫描区域。全连接层与多层感知器类似，通过作用函数计算出全连接层各单元的输出值。全连接层的作用是对提取的特征进行非线性回归，并传递给输出层。其输出层与经典的 BP 神经网络类似，通过作用函数计算出输出层各单元的输出值，输出层神经元的数目与被估计模型参数的数目相一致。

基于卷积神经网络模型，Chun 用于锂电池电化学模型参数估计的训练数据包括：电池电压 - 时间曲

线、电池电流 - 时间曲线、电池温度 - 时间曲线和 SOC (state of charge)数据, 这些数据都是基于锂电池正演计算模型(P2D, pseudo two-dimensional)得到的, 数据的采样时间持续 100 s, 它们是待估计锂电池电化学模型参数的函数, 被估计的 P2D 模型参数一共 6 个。训练样本数目 72,000, 测试样本数目 1728。研究表明, 模型参数的估计最大相对误差 12.25%, 最小的相对误差 2.214% [8]。Shen 研究了基于卷积神经网络模型锂电池电化学模型参数估计方法, 目标函数采用了带有罚函数的正则化形式, 训练数据包括: 电池电压 - 时间曲线、电池电流 - 时间曲线、电池蓄电量 - 时间曲线。锂电池充电(放电)时间离散为 25 个时间步。Shen 所建立的用于锂电池电化学模型参数估计的卷积神经网络共有 28,000 个参数, 6000 个神经元, 包括 5 个卷积层, 3 个全连接层。Shen 设定的卷积神经网络超参数值如下: 动量项系数 0.9, 学习率 0.01, 目标函数中带有惩罚项网络权值正则化因子 0.0001。数据中 70% 用作训练样本, 30% 用作测试样本。研究表明, 预测的电池蓄电量最大相对误差 3.52% [9]。

以图 1 所示的应力 - 应变曲线, 在围压 200 KPa 时的一段数据作为卷积神经网络输入数据为例, 讨论如何进行力学模型参数反演中卷积神经网络的卷积和池化运算。如图 5(a)所示, 卷积核初选如图 5(b)所示, 其中, 卷积核类似于传统神经网络中的权值一样, 在卷积神经网络训练过程中不断被优化; 如果采用 ReLU 函数作为卷积层的作用函数, 则经过一维卷积运算输出的特征图如图 5(c)所示。

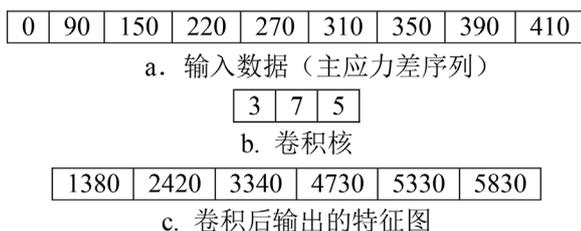


Figure 5. One dimensional Convolutional operation

图 5. 一维卷积运算

根据图 5(c)卷积后输出的特征图, 作为池化层的输入, 如图 6(a)所示, 然后进行池化运算, 即对于  $2 \times 1$  的区域进行池化操作, 得到的池化层输出如图 6(b)所示。经过池化运算后的输出作为全连接层的输入, 乘以池化层与全连接层的权值求和, 再经过作用函数运算得到全连接层的输出。同样, 池化层与全连接层的权值也是在卷积神经网络训练过程中不断被优化。最后, 全连接层与输出层之间进行权值连接, 输出层的输出就是所要估计的力学模型的参数。



Figure 6. One dimensional maximum pooling operation

图 6. 一维最大池化运算

基于卷积神经网络估计力学模型参数的主要步骤如下:

1) 选定卷积神经网络的拓扑结构, 确定卷积层、池化层的层数; 选定核函数和池化方法; 选取神经元的作用函数;

2) 选取神经网络学习的优化算法和目标函数, 例如随机梯度搜索算法; 确定神经网络中的超参数、收敛准则和最大迭代次数;

- 3) 基于解析或者数值模型, 建立神经网络训练样本集和测试集;
- 4) 输入神经网络训练样本集, 训练神经网络。卷积神经网络常常采用 BP 框架的随机梯度下降 (stochastic Gradient descent) 方法, 完成神经网络的训练[19]。
- 5) 输入神经网络测试集, 测试神经网络;
- 6) 室内进行土体三轴压缩实验, 记录不同围压条件下的应力 - 应变曲线。
- 7) 将实验数据输入神经网络, 估计力学模型参数。将基于神经网络反演得到力学模型参数, 代入到已知的力学模型, 预测岩土材料的变形曲线, 验证方法的有效性。

为了提高深度神经网络的性能, 前人提出了一些针对神经网络固有缺陷的算法和方法, 总结如下:

1) 学习率衰减方法。如果学习率过大, 神经网络训练时, 可能不收敛; 相反, 如果学习率过小, 会使得神经网络的收敛速度变慢。该方法随着神经网络学习的进行(迭代), 学习率逐渐减小。该方法中典型的代表是 AdaGrad 算法, 此算法会在学习过程中自适应调整学习率, 并且使得权值增量较大的, 学习率变小。

2) 目标函数的正则化方法[5] [14], 该方法受 Tikhonov 正则化算法的启发, 它也类似于罚函数方法, 在均方误差(mean squared error)后面增加了一项

$$E(w) = \frac{1}{2} \sum_{p=1}^P \left( \sum_{i=1}^{N_o} (y_i^p - o_i^p)^2 \right) + \alpha W^T W \quad (13)$$

式中  $W$  为神经网络的权值矩阵,  $\alpha$  为正则化因子, 它也是一个超参数, 该参数越大, 权重的惩罚就越大。

3) 神经网络训练目标函数除了均方误差之外, 还有交叉熵误差作为目标函数[6], 该方法可以提高基于梯度下降法搜索最优网络权值的收敛速度。

$$E(w) = -\sum_{p=1}^P \sum_{i=1}^{N_o} o_{pk} \ln y_{pi} \quad (14)$$

4) 神经网络超参数优化方法。超参数, 例如学习率(learning rate), 动量项(momentum), 隐含层神经元的个数。超参数往往需要进行多次尝试, 以便找到一组合理的超参数[7]。神经网络超参数(hyper parameter)优化问题, 与其说是一门科学方法, 倒不如说是一门艺术。为了获得比较合理的一组超参数, 往往采用试错的方法, 经过多次试凑得到。将数据集分为训练数据、验证数据和测试数据。其中, 训练数据用于神经网络学习, 确定神经网络的权值和阈值, 而验证数据用于评估神经网络的超参数, 而测试数据用于检验神经网络的泛化能力。超参数的优化问题, 可以采用贝叶斯优化方法(Bayesian optimization) [7]。这可能是目前限制神经网络更广泛应用的一个瓶颈。

5) 神经网络初始权值选取的问题。较差的初始权值可能会导致神经网络学习失败。同时, 为了防止神经网络初始权值均一化, 初始权值应该随机产生。Xavier 提出了使用标准差为  $1/N1/2$  的高斯分布进行权值初始化的方法,  $N$  表示神经元的个数。

6) 抑制过拟合、提高泛化能力的问题。为了解决该问题, Srivastava 提出了随机删除神经网络单元方法(Dropout 方法) [20]。在神经网络训练过程中, 随机选择隐含层的神经元, 将其删除(杀死), 使其不再传递信号。但是, 随机删除某些神经单元方法, 删除多少个, 删除原则, 缺少理论依据。

7) 深度学习输入数据的预处理方法, 包括归一化方法、均值减法、均一化方法、白化方法等[5] [21]。

## 4. 结论

深度学习是一种机器学习方法, 能够根据输入数据的特征, 进行分类和递归学习。神经网络从已知的数据中获得规律或者模型, 然后根据学习获得的知识对其他数据进行预测分析。深度学习起源于图像识别领域, 经过对其改进和补充, 能够应用到岩土材料力学模型参数估计中。随着神经网络研究的不断

深入,已经开发了许多深度学习研发平台,例如:基于MATLAB的卷积神经网络开发平台[22][23],基于Theano的深度学习平台,基于Pylearn2的卷积神经网络平台,基于TensorFLOW的深度学习框架训练卷积神经网络。尽管深度学习神经网络取得了突破性进展,仍然还存在许多需要改进和完善的方面。由于岩土力学模型参数估计实验获得的是多条一维序列,而不是二维图像,因此,需要应用一维卷积神经网络进行力学模型参数估计研究;尽管一维卷积神经网络与二维卷积神经网络在本质上是相同的。

## 基金项目

国家重点基础研究发展计划“973”项目(2015CB057804);国家自然科学基金资助项目(11572079)。

## 参考文献

- [1] Albert Tarantola. 模型参数估计的反问题理论与方法[M]. 北京: 科学出版社, 2009.
- [2] Li, S.-J. and Shao, L.-T. (2012) Inverse Procedure for Determining Model Parameter of Soils Using Real-Coded Genetic Algorithm. *Journal of Central South University*, **19**, 1764-1770. <https://doi.org/10.1007/s11771-012-1203-2>
- [3] 李守巨, 于申, 孙振祥. 基于神经网络的堆石料本构模型参数反演[J]. 计算机工程, 2014, 46(4): 267-271.
- [4] Lecun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. *Nature*, **521**, 436-444. <https://doi.org/10.1038/nature14539>
- [5] 山下隆义. 图解深度学习[M]. 北京: 中国工信出版社, 2018.
- [6] 涌井良幸. 深度学习的数学[M]. 北京: 中国工信出版社, 2019.
- [7] 斋藤康毅. 深度学习入门[M]. 北京: 中国工信出版社, 2018.
- [8] Chun, H.Y. (2019) Parameter Identification of an Electrochemical Lithium-Ion Battery Model with Conventional Neural Network. *IFAC Paperonline*, **52**, 129-134. <https://doi.org/10.1016/j.ifacol.2019.08.167>
- [9] Shen, S. (2019) A Deep Learning Method for Online Capacity Estimation of Lithium-Ion Batteries. *Journal of Energy Storage*, **25**, 1-13.
- [10] Kaba, K. (2018) Estimation of Daily Global Solar Radiation Using Deep Learning Model. *Energy*, **162**, 126-135. <https://doi.org/10.1016/j.energy.2018.07.202>
- [11] Radoslaw, M.C. and Aniel, K. (2019) Deep Neural Networks as Scientific Models. *Trends in Cognitive Sciences*, **23**, 305-317.
- [12] Taesic, L., Juwon, K., Young, U. and Hyunju, L. (2019) Deep Neural Network for Estimating Low Density Lipoprotein in Cholesterol. *Clinica Chimica Acta*, **489**, 35-40. <https://doi.org/10.1016/j.cca.2018.11.022>
- [13] Cui, M.J. and Khodayar, M. (2019) Deep Learning Based Time-Varying Parameter Identification for System-Wide Load Modeling. *IEEE Transactions on Smart Grid*, **10**, 6102-6114. <https://doi.org/10.1109/tsg.2019.2896493>
- [14] Weimer, D. and Scholz-Reiter, B. (2016) Design of Deep Convolutional Neural Network Architectures for Automated Feature Extraction in Industrial Inspection. *CIRP Annals-Manufacturing Technology*, **65**, 417-420. <https://doi.org/10.1016/j.cirp.2016.04.072>
- [15] Yu, H.Y. and Khan, F. (2015) Nonlinear Gaussian Belief Network Based Fault Diagnosis for Industrial Processes. *Journal of Process Control*, **35**, 178-200. <https://doi.org/10.1016/j.jprocont.2015.09.004>
- [16] 金列俊, 詹建明, 陈俊华. 基于一维卷积神经网络的钻杆故障诊断[J]. 浙江大学学报: 工学版, 2020, 54(3): 1-8.
- [17] 朱锡祥, 刘凤山. 基于一维卷积神经网络的车载语音识别研究[J]. 微电子学与计算机, 2017, 34(11): 21-25.
- [18] 韩林洁, 石春鹏, 张建超. 基于一维卷积神经网络的轴承剩余寿命预测[J]. 制造自动化, 2020, 42(3): 10-13.
- [19] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <https://doi.org/10.1109/5.726791>
- [20] Srivastava, N. and Hinton, G. (2014) Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *The Journal of Machine Learning Research*, **15**, 1929-1958.
- [21] Wang, L., Scott, K.A., Xu, L. and Clausi, D.A. (2016) Sea Ice Concentration Estimation during Melt from Dual-Pol SAR Scenes Using Deep Convolutional Neural Networks: A Case Study. *IEEE Transactions on Geoscience and Remote Sensing*, **54**, 4524-4533. <https://doi.org/10.1109/tgrs.2016.2543660>
- [22] Vedaldi, A. and Lenc, K. (2015) Matconvnet: Convolutional Neural Networks for MATLAB. *Proceedings of the*

23rd ACM International Conference on Multimedia, Brisbane, Australia, 26-30 October 2015, 689-692.  
<https://doi.org/10.1145/2733373.2807412>

- [23] Ahmad Radzi, S. (2016) A MATLAB-based Convolutional Neural Network Approach for Face Recognition System. *Journal of Bioinformatics and Proteomics Review*, **2**, 1-5. <https://doi.org/10.15436/2381-0793.16.009>