

轧制

基于深度神经网络的斜轧穿孔机调整参数预测

王清华¹, 孙继芸¹, 胡建华², 双远华², 赵铁琳³

(1. 太原科技大学 电子信息工程学院, 山西 太原 030024; 2. 太原科技大学 材料科学与工程学院, 山西 太原 030024;
3. 太原重工股份有限公司 技术中心, 山西 太原 030024)

摘要: 针对无缝钢管二辊斜轧穿孔生产工艺中轧机调整参数对钢管质量影响较大, 且传统机理公式计算的设定值精度不高的问题, 提出了基于深度神经网络的斜轧穿孔机调整参数预测模型。首先, 综合分析了传统的调整参数的数学模型, 并在此基础上确定了调整参数的主要影响因素。依据现场收集的数据集, 训练了二辊斜轧穿孔时轧机参数的深度神经网络预测模型。在训练深度神经网络时, 运用小批量梯度下降法和 Adam 算法相结合的方法进行了梯度估计修正, 优化了训练速度。仿真结果表明: 经深度神经网络模型预测的轧机调整参数与实测数据比较, 预测模型的 R -squared 值控制在 0.98 左右, 调整参数的相对误差均可控制在 5% 以内。相比于传统数学模型, 该预测模型具有更高的预测精度, 能够实现轧机调整参数高精度预测并用于指导生产。

关键词: 无缝钢管; 二辊斜轧穿孔; 轧辊间距; 导板间距; 顶头前伸量; 深度神经网络

DOI: 10.13330/j.issn.1000-3940.2023.11.012

中图分类号: TG355 **文献标志码:** A **文章编号:** 1000-3940 (2023) 11-0073-07

Prediction on adjusting parameters for skew rolling puncher based on deep neural network

Wang Qinghua¹, Sun Jiyun¹, Hu Jianhua², Shuang Yuanhua², Zhao Tielin³

(1. School of Electronic Information, Taiyuan University of Science and Technology, Taiyuan 030024, China;
2. School of Materials Science and Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China;
3. Technology Center, Taiyuan Heavy Industry Co., Ltd., Taiyuan 030024, China)

Abstract: Aiming at the problems that the adjustment parameters of rolling mill has a great influence on the quality of steel pipe in the two-roll skew rolling and punching production process of seamless steel pipes, and the accuracy of the set value calculated by the traditional mechanism formula is not high, a prediction model for adjustment parameters of skew rolling puncher based on deep neural network was proposed. Firstly, the traditional mathematical model of adjustment parameters was analyzed comprehensively, and the main influencing factors were determined on this basis. Then, based on the data set collected in the field, the deep neural network prediction model of rolling mill parameters during the two-roll skew rolling and punching was trained, and in the deep neural network training, the gradient estimation correction was conducted by using the combination of mini-batch gradient descent method and Adam algorithm to optimize the training speed. The simulation results show that the adjustment parameters of rolling mill predicted by the deep neural network model are compared with the measured data, the R -squared value of the prediction model is controlled at about 0.98, and the relative error of the adjustment parameters can be controlled within 5%. Compared with the traditional mathematical model, this prediction model has higher prediction accuracy, and can realize the high-precision prediction of rolling mill adjustment parameters and be used to guide production.

Key words: seamless steel pipe; two-roll skew rolling and punching; roll spacing; guide plate spacing; plug forward extension; deep neural network

斜轧穿孔是一种最为常用的无缝钢管制造工艺, 而二辊斜轧穿孔是目前热轧无缝钢管生产中常用的

方法之一, 利用穿孔设备将实心管坯穿轧为空心管坯。穿孔的变形区由轧辊、顶头、导盘(导板)构成。轧机的工具尺寸和形状设计以及轧辊间距、导板间距、顶头位置等决定了钢管质量的好坏^[1]。因此, 轧机工具设计完成后, 轧机调整参数成为对钢管的成形质量影响较大的因素。传统生产中, 多数情况下, 调整参数首先根据产品规格, 采用工艺参数计算公式计算得出^[1], 然后再进行调试。关键工

收稿日期: 2022-11-10; 修订日期: 2023-02-13

基金项目: 山西省科技重大专项 (20191102009)

作者简介: 王清华 (1980-), 女, 博士, 讲师

E-mail: 2002043@tyust.edu.cn

通信作者: 孙继芸 (1996-), 女, 硕士研究生

E-mail: s18835387178@163.com

艺质量的预测与运行决策基于操作工人的知识储备和认知水平。调试周期长,且调试过程需要试生产,故该方式容易造成能源的浪费、降低生产效率。在无缝钢管生产过程中,穿孔是最重要的一道工序,如果因为调整参数调整不当而导致钢管产生缺陷,即会影响无缝钢管的质量。所以,在穿孔操作过程中需要根据产品的不同规格,对相应的调整参数进行设置,以确保无缝钢管质量的稳定性。因此,针对斜轧穿孔调整参数建立预测模型具有重要的意义。

近些年,各国研究学者在热轧带钢生产方面开展了大量的性能预测和控制研究,并建立了多种相关模型。为实现我国钢铁制造业的高质量发展,必须从传统生产方式向智能化生产方式转变。在斜轧穿孔过程中,许多专家学者将机器学习算法模型应用到热轧钢领域的研究。Lee D 等^[2]利用长短时记忆网络提高了热轧机轧制力的预测准确率,使得带材头部的厚度误差减小;刘欣玉等^[3]利用 Matlab 神经网络工具箱中的 BP (Back Propagation) 网络,通过构建轧制力预测模型,对轧制力进行了预测,从而实现了轧制过程中各关键参数的精确预测;杨景明等^[4]提出了一种基于改进粒子群算法的 RBF (Radial Basis Function) 神经网络模型用于实现轧制力的预测,较传统的粒子群算法的 RBF 神经网络模型的预测精度更高;陈鑫等^[5]提出将传统轧制理论与 BP 神经网络模型相结合来对轧制力进行预测,仿真结果表明,所建模型的预测精度较高,将理论与智能更好地结合起来,以便满足工业现场需求;文献 [6] 也提出了基于灰色关联分析的 PSO-BP (Particle Swarm Optimization-Back Propagation) 神经网络模型来解决斜轧穿孔的管形预测问题,并应用 Matlab 软件开发了穿孔工艺模型参数预测系统;文献 [7] 将轧制机理和深度学习模型相结合,实现了精轧宽度的高精度预测;马威等^[8]利用深度学习框架建立了深度前馈神经网络的热连轧轧制力模型,与传统模型相比,所建模型具有较高的预测精度。

上述研究在斜轧穿孔方面取得了一些较好的预测效果,但针对斜轧穿孔设备的调整参数的预测研究鲜有报道。目前,仍采用依据轧制机理并结合生产经验加以整理,从理论到实践进行分析验证和导出主要调整参数的理论公式以供小型无缝钢管生产使用^[9]。因此,本文以二辊斜轧穿孔设备调整参数为研究对象,通过对传统数学模型分析与调整参数相关的影响因素,现场采集某无缝钢管厂的生产数据,以无缝钢管的管坯直径、钢管外径、顶

头直径、送进角、壁厚、轧辊出口锥角作为输入,轧辊间距、导板间距、顶头前伸量 3 个调整参数作为输出,建立了预测模型。由于 BP 神经网络属于浅层神经网络,对于复杂函数的表达能力和泛化能力受到一定的制约。而深度学习方法近年来经过大量实践并取得很好成果的一种通用方法,也是近年来人工智能领域中最活跃的分支之一。深度神经网络具有较强的抗噪、复杂函数表达能力与更强的特征学习能力。因此,采用深度神经网络算法结合数据训练集建立斜轧穿孔调整参数预测模型,并融合小批量梯度下降法 (Mini-batch Gradient Descent, MBGD) 和 Adam 优化算法来提高模型的精度和训练速度。

1 调整参数计算模型及主要因素确定

在斜轧穿孔形变过程中,很多因素会影响钢管尺寸的精度,并且会导致钢管的尺寸出现偏差。应该根据实际情况调整这些参数,以确保钢管尺寸的准确性和稳定性。其中,轧机的调整参数中轧辊间距、导板间距和顶头前伸量为影响钢管的椭圆度、壁厚和外径的重要因素。如轧辊间距的大小是决定钢管的壁厚的主要因素;钢管的外径主要取决于导板间距的大小,在实际应用中需根据不同的情况选择合适的导板间距,以确保钢管的外径不受影响;钢管外径和壁厚受顶头前伸量的影响较大。轧件在变形区的椭圆度由导板间距和轧辊间距的比值决定^[10]。由文献 [10] 可知,调整参数——轧辊间距、导板间距、顶头前伸量的计算公式如式 (1) ~ 式 (4) 所示。

$$b = (1 - \delta)d_z \quad (1)$$

式中: b 为轧辊间距; δ 为根据经验数据选择的压缩带处压缩量; d_z 为管坯直径。轧辊间距的大小与压缩带处的压缩量和管坯直径有关。

$$\tau = b \times \left[1 + \frac{0.75 \times (d_z - b)}{d_z} \times \frac{d_0}{d_z} \right] \quad (2)$$

式中: τ 为导板间距; d_0 为钢管外径。导板间距的计算取决于轧辊间距、管坯直径和钢管外径。

$$D_1 = d_m - (0.075 - 0.00135s_0)d_0 \quad (3)$$

$$y = (0.8 \sim 1.0)d_1 + (1.0 \sim 2.5)D_1 + \frac{(1.50 \sim 1.75)\pi}{2} \times \frac{\eta_0}{\eta_r} d_0 \tan \beta - \frac{D_1 + 2s_0 - b}{2 \tan \alpha_2} \quad (4)$$

式中: D_1 为顶头直径, mm; d_m 为钢管内径, mm; s_0 为壁厚, mm; y 为顶头前伸量, mm; d_1 为顶头鼻

部直径, mm; β 为送进角, ($^{\circ}$); α_2 为轧辊出口锥角, ($^{\circ}$); η_0 为轴向滑移系数; η_T 为切向滑移系数。

由式 (3) 和式 (4) 可以看到, 顶头前伸量的计算较复杂, 影响因素较多。

根据上述轧辊间距、导板间距、顶头前伸量的计算公式可知, 轧辊间距由压缩带处的压缩量和管坯直径决定; 导板间距和顶头前伸量均要考虑到轧辊间距。从计算模型中总结可以得出影响 3 个调整参数的可测量因素有管坯直径、钢管外径、顶头直径、送进角、壁厚、轧辊出口锥角。而压缩量和滑移系数则需要根据经验和调试确定。

2 改进的深度神经网络模型

2.1 神经网络的前向传播

深度神经网络 (Deep Neural Networks, DNN) 是一种基于感知机的、多层无监督的神经网络。根据神经元的位置其分为输入层、隐含层和输出层 3 部分, 深度神经网络由输入层、隐含层和输出层的神经元之间的全连接构成, 通过误差反向传播算法训练来更新权重和阈值^[11]。深度神经网络模型的基本结构如图 1 所示。

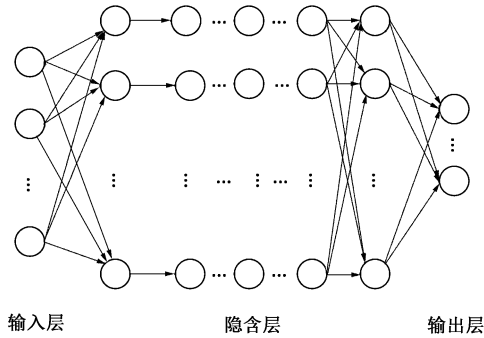


图 1 深度神经网络模型的基本结构

Fig. 1 Basic structure of deep neural network model

对于第 l 层第 i 个神经元来说, 其输入 z_i^l 和输出 g_i^l 的表达分别如式 (5) 和式 (6) 所示。

$$z_i^l = \sum_{j=1}^{n'} w_{ji}^l z_j^{l-1} + b_i^l \quad (5)$$

$$g_i^l = \sigma(z_i^l) \quad (6)$$

式中: w_{ji}^l 为连接第 l 层的第 i 个神经元与第 $l-1$ 层的第 j 个神经元之间的连接权重值; $i=1, 2, \dots, n$, n 为 DNN 第 l 层的神经元个数; $j=1, 2, \dots, n'$, n' 为 DNN 第 $l-1$ 层的神经元个数; z_j^{l-1} 为第 $l-1$ 层的第 j 个神经元的输入; b_i^l 为第 l 层的第 i 个神经元的输入偏置; $\sigma(\cdot)$ 为激活函数。本文的输入层和

隐含层采用 Relu 函数, 输出层采用 Sigmoid 函数。

2.2 深度神经网络的反向传播

在深度神经网络算法中, 反向传播将输出层的结果与目标值进行比较, 若误差达不到设计要求, 在经过对应的权重修改之后, 将其反向传播至输入层, 并对每一层的权重和偏置进行修改。实现逐层向前进行反向处理的过程, 通过反复计算误差, 直至误差满足设计需求。为了使预测输出值和目标值愈加接近, 文中采用均方误差损失函数来度量训练样本的输出损失, 其损失函数定义如下:

$$J(w, b) = \frac{1}{m} \sum_{i'=1}^m \frac{1}{2} \|\hat{y}_{i'} - y_{i'}\|^2 \quad (7)$$

式中: $J(w, b)$ 为损失函数; w 为权重; b 为偏置; $\hat{y}_{i'}$ 为第 i' 个样本的模型预测值; $y_{i'}$ 为第 i' 个样本的目标值; $i'=1, 2, \dots, m$, m 为样本数。

反向传播的目的是通过调整权重 w 、偏置 b 使损失函数 $J(w, b)$ 最小化。

在深度神经网络的训练过程中, 由于所需的训练样本数量通常很多。所以在传统梯度下降算法中, 每次迭代均需要对全部的训练数据进行梯度计算, 这将耗费大量的计算资源, 而且大规模训练集数据通常会非常冗余, 导致收敛速度缓慢^[12]。此外, 当目标函数 (损失函数) 存在多个局部最小值时, 传统梯度下降算法极易陷入局部最小值。为避免上述问题, 本文采用小批量梯度下降法, 而梯度估计是影响该方法性能的主要因素之一。Adam 算法是一种随机目标函数一阶梯度优化算法, 以低阶自适应矩估计为基础, 该方法使用动量作为参数更新方向, 收敛速度快、学习效果更佳。为了更有效地训练深度神经网络, 本文采用小批量梯度下降法和 Adam 算法相结合的方式以修正梯度估计。MBGD 是在每次迭代中, 均从大小为 m 的训练集中随机取出 1 个小批量, 然后用这个小批量来计算梯度。Adam 算法通过指数移动平均方法估计网络梯度的均值 (一阶矩) 和方差 (二阶矩), 以此为不同参数提供独立的自适应学习率^[13]。其主要参数包括: 学习率 η , 一阶矩估计的指数衰减率 β_1 , 二阶矩估计的指数衰减率 β_2 , 通常取 $\beta_1=0.900$, $\beta_2=0.999$ 。MBGD 和 Adam 算法相结合的优化算法在第 t 次训练迭代的更新具体流程如下:

Step 1: 初始化参数 w 、 b ; 初始化参数一阶矩估计 u , 二阶矩估计 v 。

Step 2: 令输入训练集 $D = \{ (x_{i'}, y_{i'}) \}_{i'=1}^m$, 其中 $x_{i'}$ 为第 i' 个样本的输入。从输入训练集中随机

采集一个小批量 $D' = \{(x_{i''}, y_{i''})\}_{i''=1}^{m'}$, 其中, $x_{i''}$ 为第 i'' 个样本的输入, $y_{i''}$ 为第 i'' 个样本的目标值, $i''=1, 2, \dots, m'$; m' 为样本数; $i'' \leq i', m' \leq m, (x_{i''}, y_{i''}) \subset (x_{i'}, y_{i'})$ 。

Step 3: 计算 t 次迭代时 m' 个样本的平均梯度:

$$\begin{cases} \delta_w = \frac{1}{m'} \nabla_w \sum J(w, b) \\ \delta_b = \frac{1}{m'} \nabla_b \sum J(w, b) \end{cases} \quad (8)$$

式中: δ_w 和 δ_b 分别为参数 w 和 b 的平均梯度; ∇_w 和 ∇_b 分别为参数 w 和 b 的梯度。

Step 4: 更新有偏一阶矩估计:

$$\begin{cases} u_w(t) = \beta_1 \cdot u_w(t-1) + (1 - \beta_1) \cdot \delta_w \\ u_b(t) = \beta_1 \cdot u_b(t-1) + (1 - \beta_1) \cdot \delta_b \end{cases} \quad (9)$$

式中: $u_w(t)$ 和 $u_b(t)$ 分别为 t 次迭代时关于参数 w 和 b 的一阶矩估计; $u_w(t-1)$ 和 $u_b(t-1)$ 分别为 $t-1$ 次迭代时关于参数 w 和 b 的一阶矩估计。

Step 5: 更新有偏二阶矩估计:

$$\begin{cases} v_w(t) = \beta_2 \cdot v_w(t-1) + (1 - \beta_2) \cdot \delta_w^2 \\ v_b(t) = \beta_2 \cdot v_b(t-1) + (1 - \beta_2) \cdot \delta_b^2 \end{cases} \quad (10)$$

式中: $v_w(t)$ 和 $v_b(t)$ 分别为 t 次迭代时关于参数 w 和 b 的二阶矩估计; $v_w(t-1)$ 和 $v_b(t-1)$ 分别为 $t-1$ 次迭代时关于参数 w 和 b 的二阶矩估计。

Step 6: 修正一阶矩偏差:

$$\begin{cases} \hat{u}_w = \frac{u_w(t)}{1 - \beta_1} \\ \hat{u}_b = \frac{u_b(t)}{1 - \beta_1} \end{cases} \quad (11)$$

式中: \hat{u}_w 和 \hat{u}_b 分别为修正偏差后的关于参数 w 和 b 的一阶矩估计。

Step 7: 修正二阶矩偏差:

$$\begin{cases} \hat{v}_w = \frac{v_w(t)}{1 - \beta_2} \\ \hat{v}_b = \frac{v_b(t)}{1 - \beta_2} \end{cases} \quad (12)$$

式中: \hat{v}_w 和 \hat{v}_b 分别为修正偏差后的关于参数 w 和 b 的二阶矩估计。

Step 8: 计算参数更新:

$$\begin{cases} w(t) = w(t-1) - \eta \frac{\hat{u}_w}{\sqrt{\hat{v}_w} + \varepsilon} \\ b(t) = b(t-1) - \eta \frac{\hat{u}_b}{\sqrt{\hat{v}_b} + \varepsilon} \end{cases} \quad (13)$$

式中: $w(t)$ 和 $b(t)$ 分别为 t 次迭代时的参数 w 和

b ; $w(t-1)$ 和 $b(t-1)$ 分别为 $t-1$ 次迭代时的参数 w 和 b ; ε 为用于数值稳定的小常数。

3 基于深度神经网络的斜轧穿孔调整参数预测模型

根据二辊斜轧穿孔机理和传统数学模型的分析得出, 斜轧穿孔机调整参数有轧辊间距、导板间距和顶头前伸量, 将这 3 个参数作为预测模型的输出, 而预测模型的输入为钢管外径、管坯直径、顶头直径、送进角、壁厚和轧辊出口锥角。确定了模型为 6 个输入 3 个输出的预测模型。然后以某钢厂无缝钢管的生产数据为基础进行建模。为了真实地反映模型的泛化能力和预测精度, 将采集的 1093 组样本中随机选取 993 组样本作为模型的训练集, 100 组样本作为测试集。对原始数据进行归一化处理, 再采用主成分分析 (Principal Component Analysis, PCA) 白化来降低输入数据特征之间的相关性。

神经网络模型中, 模型结构决定了模型性能, 在构建神经网络模型的过程中, 存在一些超参数, 如隐含层层数、激活函数、学习率、隐含层节点数、目标优化算法等, 需要手动设置, 这些超参数和模型性能有着密切的关系。

3.1 隐含层层数

本文通过固定其他所有参数, 设置如表 1 所示的对比实验, 并得到了实验结果。本文通过选取均方误差 MSE 和平均绝对百分误差 $MAPE$ 作为模型准确性的评价指标, 来验证模型的准确性。

表 1 不同隐含层层数时的模型性能指标

Table 1 Model performance indexs with different hidden layers

隐含层层数	MAPE	MSE
2	5.57	0.0903
3	3.30	0.0524
4	4.25	0.0754
5	6.63	0.0972
6	6.61	0.1056
7	7.92	0.0777
8	7.69	0.0724
9	10.07	0.0937

对隐含层层数设定一个初始值, 在此基础上逐渐增加。通过比较该网络模型的性能指标, 选择性能指标最佳时所对应的隐含层层数。根据表 1 可知,

当隐含层层数为 3 时，模型的性能最优，而后模型的性能随着隐含层层数的增加逐渐变差。且结合理论经验可知，隐含层层数的增加也会引起过拟合、训练难度的增加。且结合理论经验可知，隐含层层数的增加也会引起过拟合、训练难度的增加。

3.2 隐含层节点数

在隐含层中节点数太少将导致欠拟合，过多则可能会导致过拟合。最佳的节点数需要不断进行实验获得。在 3.1 节中已经确定模型的最优隐含层层数为 3 层，因此，需要确定每个隐含层的节点数。在调试第 1 层节点数时，固定其他两个隐含层节点数；在选择好第 1 层最佳的节点数后，继续固定第 3 层节点数，来调试第 2 层的节点数，以此类推^[14]。具体实验结果如表 2 所示。

根据表 2 的实验结果可知，并非隐含层节点数越多，模型的性能越好。当第 1 层节点数为 10、第 2 层节点数为 40、第 3 层节点数为 30 时（表 2 中黑体），模型的预测效果最佳。

4 预测结果与分析

本文目标函数选取均方误差，学习速率为 0.001、训练次数为 2000，批量大小为 100。图 2 和图 3 分别为基于深度神经网络模型对轧辊间距、导

表 2 不同隐含层节点数时的模型性能指标			
Table 2 Model performance indexs with different hidden layer nodes			
隐含层层数	隐含层节点数	MAPE	MSE
第 1 层	5	7.22	0.1044
第 1 层	10	3.47	0.0610
第 1 层	20	4.67	0.0791
第 1 层	30	5.68	0.0811
第 1 层	40	6.04	0.0987
第 1 层	50	8.05	0.1256
第 2 层	5	7.44	0.1249
第 2 层	10	9.14	0.0944
第 2 层	20	7.01	0.0707
第 2 层	30	5.67	0.0864
第 2 层	40	4.01	0.0664
第 2 层	50	9.17	0.0894
第 3 层	5	9.60	0.1891
第 3 层	10	6.11	0.0988
第 3 层	20	5.31	0.0841
第 3 层	30	4.33	0.0683
第 3 层	40	6.32	0.0845
第 3 层	50	7.87	0.1095

板间距、顶头前伸量的预测值与目标值结果对比图和相对误差折线图。

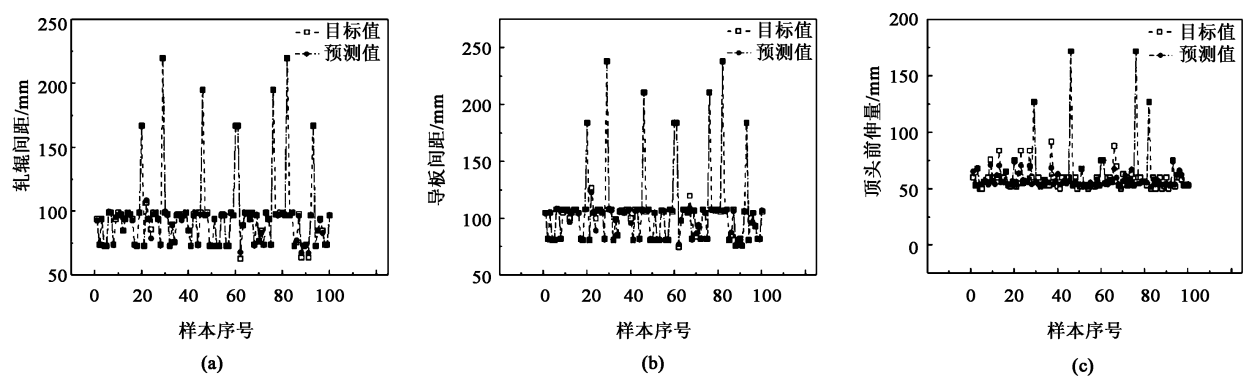


图 2 调整参数预测值与目标值对比图

(a) 轧辊间距 (b) 导板间距 (c) 顶头前伸量

Fig. 2 Comparison diagrams of adjusting parameters between predicted and target values

(a) Roll spacing (b) Guide plate spacing (c) Plug forward extension

由图 2 和图 3 中的预测结果可得，轧辊间距的预测值与目标值的平均相对误差为 0.9925%、导板间距的预测值与目标值的平均相对误差为 1.0999%、顶头前伸量的预测值与目标值的平均相对误差为 4.8015%，表明该深度神经网络模型对于轧辊间距、导板间距及顶头前伸量的预测效果精度

较高，预测误差较小。

表 3 为采用传统数学模型和深度神经网络模型得到的轧辊间距、导板间距和顶头前伸量 3 个调整参数的预测结果的均方根误差 *RMSE*、平均绝对误差 *MAE* 和决定系数 *R-square* 的比较。从对比结果可以看出，在预测轧辊间距和导板间距时，深度神经

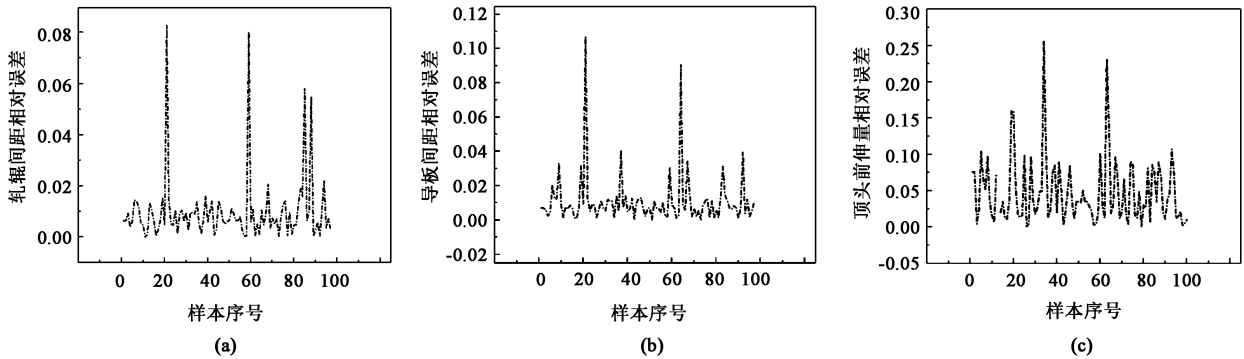


图 3 调整参数预测值与目标值的相对误差折线图

(a) 轧辊间距 (b) 导板间距 (c) 顶头前伸量

Fig. 3 Relative error polygraphs of adjusting parameters between predicted and target values

(a) Roll spacing (b) Guide plate spacing (c) Plug forward extension

表 3 传统数学模型与预测模型结果对比

Table 3 Comparison of results between traditional mathematical model and prediction model

参数	模型	RMSE	MAE	R-square
轧辊间距	传统数学模型	2.8505	2.2690	0.8995
	深度神经网络	1.2816	2.6433	0.9847
导板间距	传统数学模型	3.1700	1.8938	0.9696
	深度神经网络	1.9330	2.7673	0.9907
顶头前伸量	传统数学模型	49.0102	34.4652	<0.0000
	深度神经网络	9.5521	4.9154	0.8989

网络预测的结果要稍优于传统数学模型的结果。在预测顶头前伸量时，由于传统数学模型的计算公式中使用了轧辊间距的计算值，并且有些系数采用经验值，使得误差累积变大，预测结果明显差于深度神经网络。表明基于深度神经网络的调整参数预测模型的性能优于传统数学模型。

利用搭建好的深度神经网络模型框架，针对 RMSprop 算法和 Adam 算法的选择进行了对比实验，具体对比实验结果如表 4 所示。

表 4 各优化模型预测下正确率对比

Table 4 Comparison of prediction accuracy for each optimization model

优化算法	测试集正确率/%		
	轧辊间距	导板间距	顶头前伸量
无优化 (迭代 1000 次)	97.80	95.90	65.00
MBGD+Adam (迭代 1000 次)	99.35	97.50	74.70
MBGD+RMSprop (迭代 1000 次)	95.35	93.80	68.15
MBGD+RMSprop (迭代 2000 次)	99.35	96.94	71.41

根据表 4 可知，在迭代训练 1000 次的情况下，选择 MBGD 和 Adam 结合优化算法，其正确率较高。而且，当迭代训练为 2000 次的情况下，RMSprop 算法的学习效果依然不及 Adam 优化算法。由此可知，在 MBGD 和 Adam 优化算法模型下，其收敛速度更快，学习效果更为有效。

5 结论

(1) 结合实际生产过程以及工业现场采集数据，将轧辊间距、导板间距、顶头前伸量作为关键调整参数，根据传统数学模型分析得出，影响较大的工艺参数有钢管外径、管坯直径、顶头直径、壁厚、送进角、轧辊出口锥角。

(2) 使用深度神经网络建立调整参数预测模型，以小批量梯度下降法为基础结合 Adam 优化算法避免在训练过程中陷入局部最优值。根据实例验证仿真，采用深度神经网络建立的穿孔机调整参数预测模型具有较高的预测精度，可在生产前预测穿孔机设备的调整参数，对高效、高质量生产具有指导意义。

参考文献:

[1] 李连诗, 韩观昌. 小型无缝钢管生产 (上册) [M]. 北京: 冶金工业出版社, 1989.
Li L S, Han G C. Production of Small Seamless Steel Pipe (Volume 1) [M]. Beijing: Metallurgical Industry Press, 1989.
[2] Lee D, Lee Y. Application of neural-network for improving accuracy of roll force model in hot-rolling mill [J]. IFAC Proceedings Volumes, 2000, 33 (22): 227-231.

(下转第 103 页)