

[文章编号] 1004- 0609(2001) 03- 0481- 04

基于人工神经网络的凝固组织晶粒尺寸预测^①

訾炳涛¹, 崔建忠², 巴启先², 姚可夫¹

(1. 清华大学 机械工程系, 北京 100084; 2. 东北大学 材料与冶金学院, 沈阳 110006)

[摘 要] 通过分析强脉冲电磁场作用下铝合金凝固组织晶粒尺寸实验数据, 并结合人工神经网络, 建立了强脉冲电磁场作用下铝合金凝固组织晶粒尺寸的人工神经网络 BP 算法模型。研究表明, 用该神经网络模型进行模拟得到的计算结果和实验数据吻合得较好, 因此这一方法可用来对强脉冲电磁场作用下的凝固组织晶粒尺寸进行预测和控制, 为优化实验设计提供了简便、实用的方法和手段。

[关键词] 凝固组织; 晶粒尺寸; 人工神经网络; BP 算法模型

[中图分类号] TM 15; TN 15

[文献标识码] A

对金属凝固过程进行强脉冲电磁场处理是近几年才发展起来的一种新颖独特的凝固组织细化法^[1~3], 它具有细化效果显著、无污染(即不改变合金原有成分)、操作简便等特点。作者在文献[1]中给出了强脉冲电磁场作用下 LY12 铝合金的晶粒尺寸随强脉冲电流密度和强脉冲磁感应强度的变化曲线。

传统的晶粒度测定是先制备金相试样, 再在显微镜下用不同的方法计算出晶粒尺寸或数量, 然后按一定算法求出晶粒度^[4~5]。该办法沿用了多年, 但仍存在缺点: 1) 一般的晶粒度分级均是定性或半定量的; 2) 操作过程繁复、不便, 而且人为影响因素较多, 误差大; 3) 晶粒度不仅是一个统计概念, 而且还与放大倍数、所取视场的大小和位置、腐蚀剂的类型、晶粒的形状和取向、人们对晶界的理解程度等因素有关。为了减少测量中的麻烦和各种误差, 作者提出了用人工神经网络^[6~11]对强脉冲电磁场作用下的 LY12 铝合金凝固组织晶粒尺寸进行预测, 并通过样本数据检验该模型的可行性和抗干扰能力。

1 BP 算法模型

1.1 BP 神经网络的特点

采用的误差反向传播的 BP 型神经网络是一个有导师指导学习的前向人工神经网络。该神经网络一般有三层神经元连接, 同一层的神经元之间

无任何连接, 而层与层的神经元之间采用全连接方式。其连接强度或记忆强度用权重或权值表示, 正的权值表示影响的增加, 负的权值表示影响的减弱。BP 模型在学习训练过程中, 存在 2 种相反的信息传递路线: 正向传播和反向传播。在正向传播过程中, 其输入信息从输入层经过隐层逐层处理, 并传向输出层, 这时, 每一层的神经元状态只影响下一层的神经元状态。若在输出层不能得到期望的输出, 则转入反向传播, 将误差信号沿原来的连接通路返回, 并通过修改各层神经元的权值和阈值, 使输出层误差达到最小, 从而完成网络的学习过程。当权值和阈值调整到适应于一组样本后, 则该网络模型就已经建立, 而且该模型还可以对其它样本继续学习, 以使其具有更广泛的适应性, 而不必重新构造一个新的网络结构。

BP 神经网络把一组样本的输入/输出(I/O)问题转变为一个非线性优化问题, 并使用优化方法中最普遍采用的梯度下降法, 用迭代运算求解权值, 加入隐层节点使优化问题的可调参数增加, 从而得到更精确的题解。它可直接从实验数据进行学习, 而对所建立的网络模型未做任何假设; 而且, BP 网络模型可对实验数据建立起由输入矢量空间到输出矢量空间的复杂非线性关系的直接映射。如果输入矢量空间的输入层节点数为 m , 输出层空间的输出层节点数为 n , 则该 BP 网络可看作是从矢量空间 R^m 到矢量空间 R^n 的映射, 即有以下关系成立:

① [基金项目] 国家重点基础研究发展规划基金资助项目(G1999064900- 05); 教育部电磁过程重点实验室访问学者基金资助项目
[收稿日期] 2000- 09- 21; [修订日期] 2001- 01- 08 [作者简介] 訾炳涛(1965-), 男, 博士后。

$$F: R^m \rightarrow R^n, Y = F(X) \quad (1)$$

对于样本集合 X 和输出 Y , 可认为存在某一映射 G , 使

$$Y_i = G(X_i), i = 1, 2, \dots, k \quad (2)$$

式中 (X_i, Y_i) 为样本的输入-输出矢量对。

1.2 凝固晶粒尺寸模型

本文采用的 BP 网络模型为 $3 \times 4 \times 1$ 型结构, 共有 3 层神经元连接, 即输入层、中间隐层和输出层, 如图 1 所示。输入层有 3 个节点, 即在强脉冲电磁场细化金属凝固组织工艺^[1,3]中, 主要有 3 个影响因素 X_1, X_2 和 X_3 , 分别表示放电电压、电容量和强脉冲电流密度(或强脉冲磁感应强度); 输出层只有一个节点, 表示凝固组织晶粒尺寸 D ; 中间层为隐含层, 本文取 4 个节点 $Y_j (j = 1, 2, 3, 4)$ 。

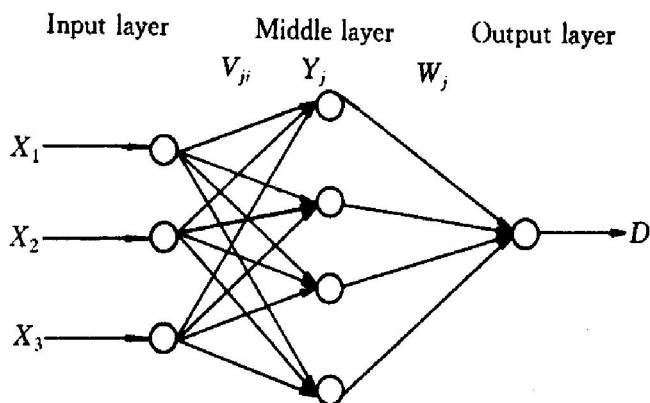


图 1 $3 \times 4 \times 1$ 结构 BP 人工神经网络模型

Fig. 1 $3 \times 4 \times 1$ structure model of BP artificial neural network

输入层采用线性单元; 中间隐层采用非线性单元, 其第 j 个单元的输入 N_j 可通过把输入层各单元的输入加权求和得到, 其输出 Y_j 是由其输入 N_j 经过非线性函数 $f(x)$ 求得; 输出层则采用一个非线性单元, 其输入 N 可由隐层各单元的输入 Y_j 加权求和得到, 其输出 D (即晶粒尺寸) 亦由其输入 N 经过同一个非线性函数 $f(x)$ 求得。输入层-隐层及隐层-输出层之间的连接权值分别表示为 V_{ji} 和 W_j 。

鉴于神经网络的特点, 对所有的输入层和输出层的参数均做了归一化处理, 使所有参数的取值落在 $[0, 1]$ 范围内。该 BP 型人工神经网络的正向、反向传播的学习与训练过程如下:

①置 V_{ji} 和 W_j 为 $[-0.5, 0.5]$ 之间的随机数。

②选取学习率 η , 动量项系数 α , 并置 $X_4 = Y_5 = -1$ 。

③从输入层到输出层, 逐层计算每个单元的输入。

隐层:

$$N_j = \sum_{i=1}^4 V_{ji} X_i, j = 1, 2, 3, 4 \quad (3)$$

$$Y_j = f(N_j) \quad (4)$$

输出层:

$$N = \sum_{j=1}^5 W_j Y_j \quad (5)$$

$$D = f(N) \quad (6)$$

这里, 取激励函数

$$f(N) = \frac{1 - e^{-N}}{1 + e^N} \quad (7)$$

式(3)~(7)中将阈值归入 V_{ji} 和 W_j , 它们与固定为 -1 的输入相连接。

④计算该网络的系统误差。设有 P 组学习样本, 则该系统的均方误差

$$E = \frac{1}{2P} \sum_{n=1}^P (D_n - D'_n)^2 \quad (8)$$

式中 D_n 表示第 n 个学习样本的实验值, 通常被称为导师信号; D'_n 则表示相应网络的实际输出值。

⑤如果该网络的系统误差 E 满足设计要求或达到指定的学习次数, 则结束学习, 否则进行以下步骤。

⑥逐层计算网络中各单元的学习信号:

$$\delta = (D - D') f'(N) \quad (9)$$

$$\delta_j = W_j \delta' (N_j), j = 1, 2, 3, 4 \quad (10)$$

⑦按梯度下降方向自修正权值:

$$W_j(t+1) = W_j(t) + \eta \delta Y_j + \alpha [W_j(t) - W_j(t-1)] \quad (11)$$

$$V_{ji}(t+1) = V_{ji}(t) + \eta \delta X_i + \alpha [V_{ji}(t) - V_{ji}(t-1)] \quad (12)$$

⑧转至 ②继续执行(必要时, 可适当调整学习率和动量项系数的大小)。

2 模拟结果及分析

所有模拟计算工作均在 AMD K6(tm)/233 CPU 的 PC 机上进行。当迭代进行到 2 万次后, 系统误差基本上稳定在 4.4×10^{-4} , 已能够完全满足设计或实际要求, 如图 2 所示。图 3 则给出了模拟与实验^[1]的比较结果。

通过模拟运算发现, 经过 2 万次的迭代学习与训练, 该网络的实际输出值与实验值很接近。所有样本的模拟计算结果的相对误差均在 11% 以内, 甚至个别样本的相对误差为 0, 如表 1 所示。所以, 用该学习与训练过程得到的网络连接权值即可作为强

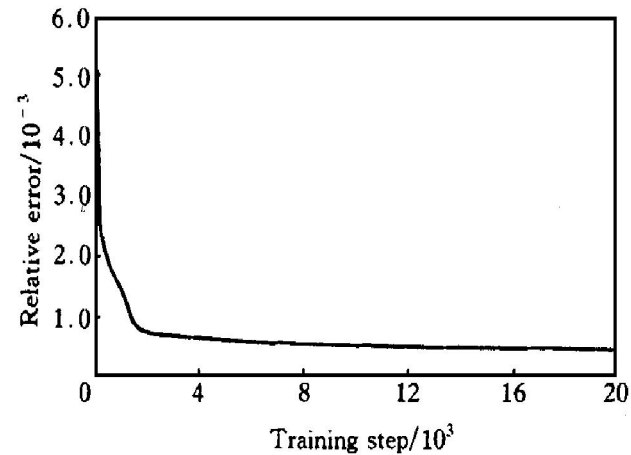


图 2 相对误差随训练次数的变化曲线

Fig. 2 Curve of relative error vs training step

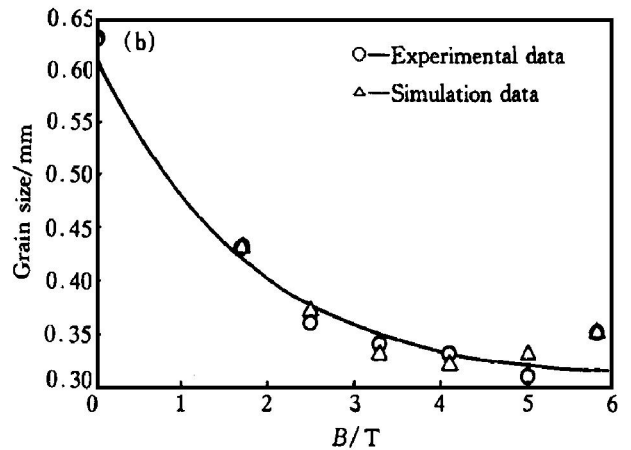
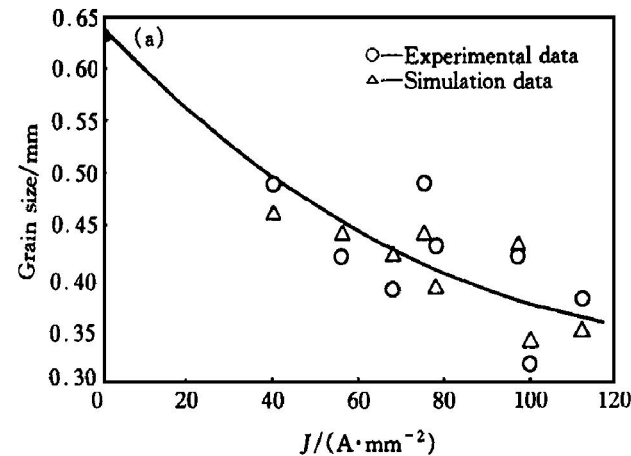


图 3 模拟计算结果与实验结果^[1]的比较

Fig. 3 Comparison between simulation data and experimental data^[1]

- (a) —Effect of strong pulsed electric current
- (b) —Effect of strong pulsed magnetic field

脉冲电磁场作用下的 LY12 铝合金凝固组织晶粒尺寸的预测和控制模型, 并可用它对别的实验样本和新的实验数据进行学习和训练, 以及对后期的实验结果进行预测。

为了检验该人工神经网络 BP 算法模型的抗干扰能力, 分别对 3 个检验样本的输入数据进行了改变, 然后再模拟, 所得的模拟计算结果如表 2

表 1 模拟计算结果及相对误差

Table 1 Simulation data and relative errors

X_1 /kV	X_2 /μF	X_3 /(A·mm ⁻²)	X_3 /T	D' /mm	D /mm	Relative error/%
0	0	0		0.63	0.63	0
3	80	40		0.46	0.49	6
4	80	56		0.44	0.42	5
5	80	68		0.42	0.39	8
6	80	75		0.42	0.47	11
7	80	97		0.44	0.42	5
4	160	78		0.39	0.43	9
4	240	100		0.35	0.32	9
4	320	112		0.35	0.38	8
2	80		1.7	0.43	0.43	0
3	80		2.5	0.37	0.36	3
4	80		3.3	0.33	0.34	3
5	80		4.1	0.32	0.33	3
6	80		5.0	0.33	0.31	6
7	80		5.8	0.35	0.35	0

表 2 模拟干扰结果及相对误差

Table 2 Disturbed data and relative errors

X_1 /kV	X_2 /μF	X_3 /(A·mm ⁻²)	X_3 /T	D' /mm	D /mm	Relative error/%
7 [#]	80	68		0.38	0.39	3
4	240 [#]	112		0.35	0.38	8
2	80		3.3 [#]	0.44	0.43	2

Data labeled[#] were changed to predict the anti-disturbance and stability of the model.

所示。可见, 虽然改变了其中个别数据, 但并不影响模拟计算结果的准确性。这表明该网络模型具有较强的稳定性和抗干扰能力。

需要说明的是, 在实际应用中, 学习率 η 或学习步长的选取是一个至关重要的因素, 它对学习效果影响很大。 η 取值越大, 虽然学习速度加快, 每次权值的改变也越激烈, 但大的 η 值容易使学习过程陷于局部极小, 或容易使权重系数产生反复震荡而不能收敛。而 η 取值越小, 则学习速度越慢, 占用机时越多, 学习效率越慢。一般地, 在学习的初期, 可使 η 值取大一些, 这样可加快学习速度, 提高学习效率; 但在接近最佳点时, 必须使 η 的取值小些, 以避免使权重产生震荡造成不能收敛。

3 结论

用人工神经网络方法建立了强脉冲电磁场作用下的凝固组织晶粒尺寸 BP 算法模型。用该模型得到的模拟计算结果和实验结果基本吻合。该模型预测精度高、误差小, 且具有较强的抗干扰能力和稳定性。用该网络模型可以较好地对金属凝固组织

晶粒尺寸进行预测和控制。

[REFERENCES]

- [1] ZI Bing-*tao*(訾炳涛), BA Qí-xian(巴启先), CUI Jian-zhong(崔建忠), et al. 强脉冲电磁场对金属凝固组织的影响[J]. *Acta Physica Sinica*(物理学报), 2000, 49 (5): 1010– 1014.
- [2] ZI Bing-*tao*, BA Qí-xian, CUI Jian-zhong, et al. Study on axial changes of as-cast structures of Al-alloy sample treated by the novel SPMF technique [J]. *Scripta Materialia*, 2000, 43(4): 377– 380.
- [3] ZI Bing-*tao*(訾炳涛). LY12 铝合金凝固组织的脉冲电磁细化新工艺研究[D]. Shenyang: Northeastern University, 2000.
- [4] TU Shì-run(屠世润), GAO Yue(高 越). Theory and Practice of Metallography (金相原理与实践) [M]. Beijing: Mechanical Industry Press, 1990.
- [5] QIN Guo-you(秦国友). Quantitative Metallography(定量金相) [M]. Chengdu: Sichuan Science and Technology Press, 1987.
- [6] SHI Hong-bao(施鸿宝). Neural Network and Its Application(神经网络及其应用) [M]. Xí'an: Xí'an Jiaotong University Press, 1993.
- [7] Dobrzański L A, Sitek W. Application of a neural network in modelling of hardenability of constructional steels [J]. *Journal of Materials Processing Technology*, 1998, 78: 59– 66.
- [8] Korczak P, Dyja H, Labuda E. Using neural network models for predicting mechanical properties after hot plate rolling processes [J]. *Journal of Materials Processing Technology*, 1998, 80– 81: 481– 486.
- [9] YU Lì-juan(于丽娟), ZHU Chang-chun (朱长纯). 用人工神经网络预测场发射开启电场[J]. *Acta Physica Sinica*(物理学报), 2000, 49(1): 170– 172.
- [10] SHUANG Yuan-hua(双远华), FAN Jian-cheng(樊建成), LAI Ming-dao(赖明道). 人工神经网络对管材张减精度预测[J]. *Iron and Steel*(钢铁), 2000, 35 (2): 28– 31.
- [11] QI Le-hua(齐乐华), HOU Jun-jie (侯俊杰), YANG Fang (杨 方), et al. 金属液体凝固中直接挤压工艺的神经网络[J]. *The Chinese Journal of Nonferrous Metals*(中国有色金属学报), 1999, 9(3): 586– 589.

Application of artificial neural network in predicting grain size of solidification structure

ZI Bing-*tao*¹, CUI Jian-zhong², BA Qí-xian², YAO Ke-fu¹

(1. Department of Mechanical Engineering, Tsinghua University, Beijing 100084, P. R. China;

2. School of Materials & Metallurgy, Northeastern University, Shenyang 110006, P. R. China)

[Abstract] After analyzing the experimental data of the solidified grain size in Al-alloy under strong pulsed electromagnetic field, a BP algorithmic model of artificial neural network was established. The results show that the simulating results are better in agreement with the experimental results. Therefore, this BP algorithmic model of artificial neural network can be used to control the parameters and predict solidified grain size under applied strong pulsed electromagnetic field. It provides an easy and practical method and means for optimizing experimental design.

[Key words] solidification structure; grain size; artificial neural network; BP arithmetic model

(编辑 杨 兵)