

基于自组织增量-图卷积神经网络的金相图半监督学习

李维刚* 湛竟成 谢璐 赵云涛

(武汉科技大学冶金自动化与检测技术教育部工程研究中心 武汉 430081)

摘要: 采用深度学习对钢铁材料显微组织图像分类, 需要大量带标注信息的训练集。针对训练集人工标注效率低下问题, 该文提出一种新的融合自组织增量神经网络和图卷积神经网络的半监督学习方法。首先, 采用迁移学习获取图像数据样本的特征向量集合; 其次, 通过引入连接权重策略的自组织增量神经网络(WSOINN)对特征数据进行学习, 获得其拓扑图结构, 并引入胜利次数进行少量人工节点标注; 然后, 搭建图卷积网络(GCN)挖掘图中节点的潜在联系, 利用Dropout手段提高网络的泛化能力, 对剩余节点进行自动标注进而获得所有金相图的分

关键词: 自组织增量神经网络; 图卷积神经网络; 自动标注; 钢材显微组织

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2021)11-3301-08

DOI: 10.11999/JEIT201029

Semi Supervised Learning of Metallographic Data Based on Self-organizing Incremental and Graph Convolution Neural Network

LI Weigang SHEN Jingcheng XIE Lu ZHAO Yuntao

(Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan 430081, China)

Abstract: It needs a large number of training sets with annotation information to classify microstructure images of steel materials by deep learning. To solve the problem of low efficiency of manual image annotation, a new semi-supervised learning method combining self-organizing incremental neural network and graph convolutional neural network is proposed. Firstly, it uses transfer learning to obtain the feature vector set of images. Secondly, it obtains the topology structure by adopting the Weighted Self-Organizing Incremental Neural Network(WSOINN) based on connection weight strategy to learn feature data, and manually annotates a small number of nodes which are selected by the number of victories of node. Then, a Graph Convolution Network (GCN) is built to mine the potential connections of nodes in the graph, dropout is used to improve the generalization ability of the network, and the remaining nodes are automatically annotated to obtain the classification results of the metallograph. Experiment on the metallographic data collected from a state key laboratory, the accuracy of automatic classification under different manual annotation ratio is compared. The results show when the image annotation amount is only 12% of the traditional model, and the classification accuracy of the proposed model can reach up to 91%.

Key words: Self-Organizing Incremental Neural Network(SOINN); Graph Convolution Network(GCN); Automatic annotation; Steel material microstructure

1 引言

钢的微观组织决定其使用性能, 对微观组织的

定性和定量研究一直都是钢铁材料领域的重要工作^[1-3]。将采用适当方法(如抛光、腐蚀)处理后的实验钢试样置于光学显微镜(Optical Microscope, OM)或电子显微镜(Electron Microscope, EM)下, 可观察到钢的显微组织形貌图像, 即金相图^[4,5]。传统上, 金相图的辨识由人工完成, 对人的专业经验依赖性较大, 即使是经验丰富的专家也会因为肉眼看不到的图像细节而分析失误。而现代钢材种类

收稿日期: 2020-12-07; 改回日期: 2021-03-21; 网络出版: 2021-04-09

*通信作者: 李维刚 liweigang.luck@foxmail.com

基金项目: 国家自然科学基金(51774219)

Foundation Item: The National Natural Science Foundation of China (51774219)

越来越多,其内部显微组织越来越复杂,人工辨识面临巨大挑战^[6,7]。

随着计算机视觉的深入发展,国内外研究者已开始将深度学习用于金相图的自动辨识问题。Pauly等人^[8]使用数据挖掘方法对金相图分类,但由于不同类别间提取到的特征差异性不够,在测试集上仅获得48%的精度。Chowdhury等人^[9]组合不同的特征提取和特征选择方法作用于金相图,并选择不同分类器,比较了不同组合间模型性能的差异。Azimi等人^[10]采用全卷积神经网络对金相图进行分类,在所收集的数据集上能够获取90%以上的精度。文献^[11]融合多种图像增强方法,并改进AlexNet, VGGNet, ResNet, GoogleNet适用于金相图,在所收集的数据集上能够获取95%以上的精度。上述研究所使用的一般均是小规模数据集,且由于人工标注困难、效率低问题,用于模型训练的已标记图像数量较少,模型泛化能力弱、实际应用困难。

图^[12-14] $G = (V, E)$ 能表达非欧空间复杂数据关系^[15],图卷积(Graph Convolutional Network, GCN)可用于超维关联数据的挖掘和分析^[16-18]。基于已有拓扑学习模型^[19],本文提出一种基于自组织增量-图卷积神经网络的半监督学习方法。在原自组织增量神经网络(Self-Organizing Incremental Neural Network, SOINN)中引入连接权重概念来表示两节点相似性,得到引入连接权重策略的自组织增量神经网络(Weighted SOINN, WSOINN),并引入节点胜利次数以挑选少量节点进行人工标注;进而,搭建GCN学习拓扑图中高阶特征来预测节点的类别信息,达到用较少的图像标注量获取较高模型分类准确率的目的。本文用WSOINN获取拓扑图结构描述图像数据的空间分布,用GCN将WSOINN拓展至半监督学习,实现钢铁材料金相图的自动分类。针对收集到的贝氏体、低碳板条

马氏体、高碳片状马氏体、铁素体、下贝氏体、珠光体6种类型的金相图进行实验,结果表明,本方法具有较高的准确性和适应性。

2 连接权重策略的自组织增量-图卷积神经网络(WSOINN-GCN)

本节首先给出WSOINN-GCN的整体框架;然后,分小节介绍各个模块:2.1节介绍图像数据特征提取;2.2节给出WSOINN算法步骤;2.3节结合金相图特征设计GCN;2.4节给出WSOINN-GCN的算法步骤。

WSOINN-GCN模型框架如图1所示,它由3部分组成:第1部分基于迁移学习获得图像数据的特征向量集合;第2部分采用引入连接权重策略的自组织增量神经网络(WSOINN)提取特征数据的拓扑图结构,并按照节点胜利次数选择少量节点进行人工标注;第3部分搭建图卷积网络(GCN),采用交叉熵损失函数、Adam算法优化网络参数,自动标注剩余节点,最后基于欧氏距离来分类所有图像数据。

2.1 迁移学习获取特征向量集合

图2展示了收集自某国家重点实验室场发射扫描电子显微镜所拍摄的不同钢铁材料的微观组织图片,依次为铁素体、珠光体、贝氏体、下贝氏体、板条马氏体、片状马氏体,共2342张,图片像素大小均为 221×221 。

如图3,本文采用在ImageNet数据集上已训练好的VGG16卷积模块提取每张金相图的特征,并对从每张金相图获得的512张特征图作全局均值池化,每张图输出一个512维的特征向量,从而得到所有金相图特征提取后的数据特征集合。

2.2 引入连接权重策略的自组织增量神经网络(WSOINN)

SOINN可获取特征数据的空间拓扑图结构,

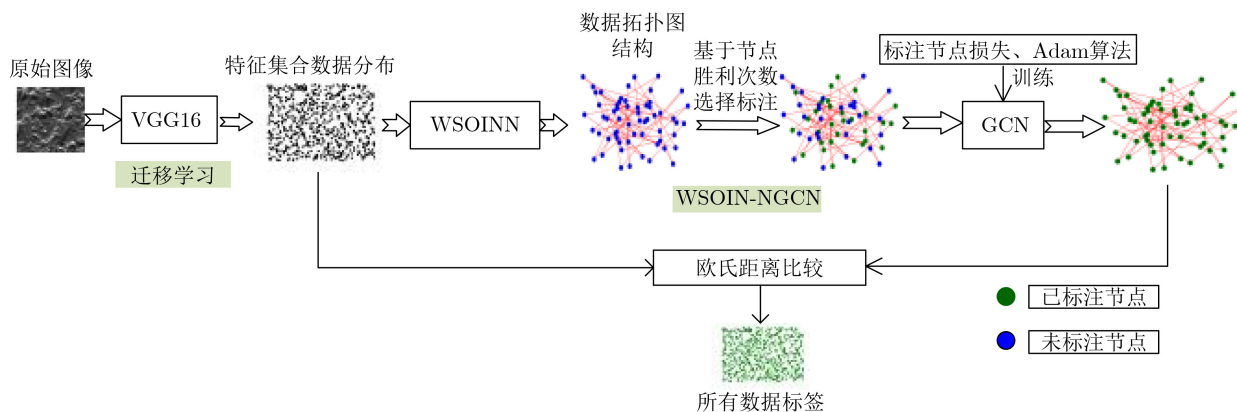


图1 WSOINN-GCN模型框架

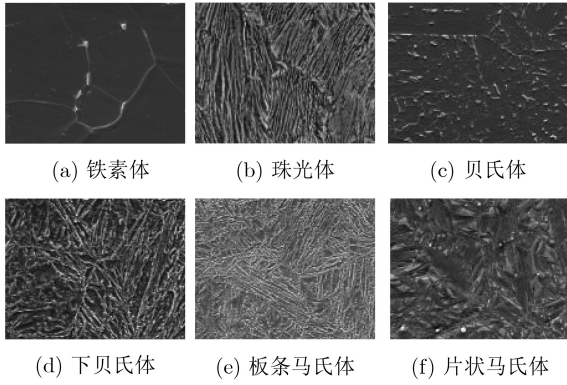


图 2 金相图样本

而GCN可用于挖掘巨量、稀疏、超维关联图数据的关系。为融合SOINN与GCN，本文提出引入连接权重数的自组织增量神经网络(WSOINN)，并引入节点胜利次数以挑选少量节点进行人工标注。WSOINN的算法步骤如下：

(1) 初始化节点集合 $V = \{v_1, v_2\}$, $v_1, v_2 \in R^d$; 连接 $E \subseteq N \times N$ 为空集; 节点胜利次数 $\text{win_times} = \{t_{v_1} = 0, t_{v_2} = 0\}$ 。

(2) 接收新的输入样本 $\xi \in R^d$, 根据欧几里得范数查找 V 中与 ξ 最接近的节点 s_1 和 s_2 , 即 $s_1 = \underset{v_n \in V}{\operatorname{argmin}} \|\xi - v_n\|$, $s_2 = \underset{v_n \in V \setminus \{s_1\}}{\operatorname{argmin}} \|\xi - v_n\|$, 胜利次数加 1, 即 $t_{s_1} = t_{s_1} + 1$, $t_{s_2} = t_{s_2} + 1$ 。

(3) 计算节点 s_1, s_2 的相似度阈值 T_{s_1}, T_{s_2} , 对节点 $v \in V$, 将与 v 连接的节点集合记为 P , 若 $P = \emptyset$, $T_v = \min \|v - v_n\|, v_n \in V \setminus \{v\}$, 若 $P \neq \emptyset$, $T_v = \max \|v - v_n\|, v_n \in P$ 。

(4) 如果 $\|\xi - s_1\| > T_{s_1}$ 或 $\|\xi - s_2\| > T_{s_2}$, $V = V \cup \{\xi\}$, 否则丢弃样本 ξ , 同时修正节点 s_1 和 s_2 , $s_1 = s_1 + \varepsilon(t)(\xi - s_1)$, $s_2 = s_2 + \varepsilon'(t)(\xi - s_2)$, 其中 $\varepsilon(t) = 1/t_{s_1}$, $\varepsilon'(t) = t_{s_2}/100$ 。

(5) 若 s_1 和 s_2 无连接, $E = E \cup \{(s_1, s_2)\}$, $w_{(s_1, s_2)} = 1$; 若有连接, $w_{(s_1, s_2)} = w_{(s_1, s_2)} + 1$ 。

如果 $w_{(s_1, s_2)} > W_{\max}$, $E \setminus \{(s_1, s_2)\}$, W_{\max} 为预定义值。

(6) 每学习完总样本的百分比例 p , 删除孤立节点, 及对应的胜利次数。若样本输入未结束, 则返回步骤 (2)。否则, 输出图节点集合 V 、连接集合 E 及各个节点胜利次数。

由上述WSOINN算法过程可知, 其节点间连接权重 w 代表两节点间的相似性, 连接权重越大则表示两节点越相似; 节点胜利次数越多, 表明该节点越具有代表性, 越重要。

2.3 图卷积神经网络(GCN)

针对金相图的半监督学习分类问题, 搭建具有 3 层图卷积网络的 GCN 模型, 如图 4 所示。 N 表示图结构中节点数量, 每层图卷积后均接 ReLU 激活函数, 其中第 1、第 2 层用于特征整合与降维, 输出维度分别为 $N \times 512$, $N \times 256$, 参数量分别为 512×512 , 512×256 , 第 3 层结合 Softmax 层用于分类, 第 3 层输出维度为 $N \times 6$, 参数量为 256×6 。

基于 WSOINN 输出的含有 N 个节点的图 $G = (V, E)$, 用 $A \in R^{N \times N}$ 表示图节点连接关系矩阵, 定义归一化的图拉普拉斯矩阵为

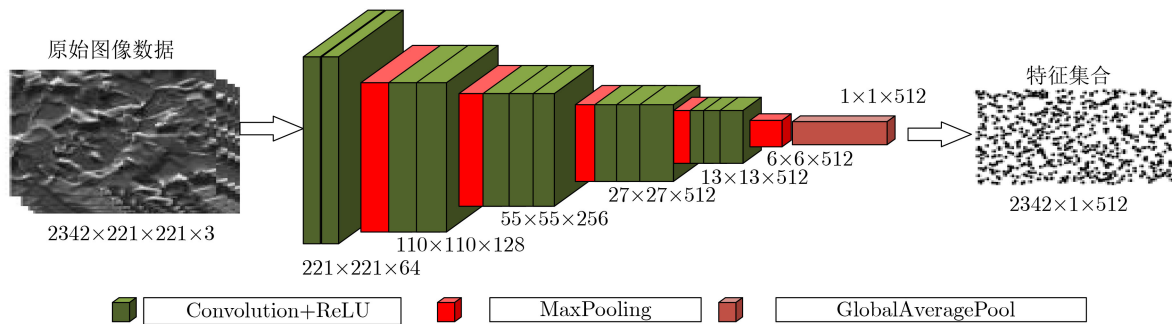


图 3 利用VGG16卷积模块提取金相图的特征

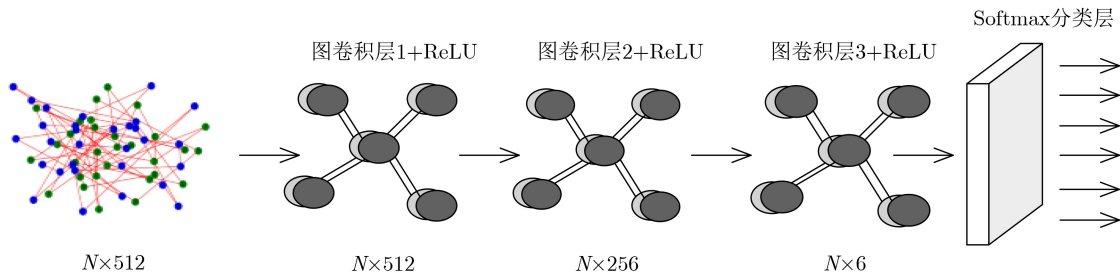


图 4 3层图卷积网络结构

$$\mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \quad (1)$$

其中, \mathbf{I} 为单位矩阵, \mathbf{D} 为对角矩阵, $D_{ii} = \sum_j A_{ij}$ 可知 \mathbf{L} 为对称矩阵, 对 \mathbf{L} 奇异值分解

$$\mathbf{L} = \mathbf{U} \mathbf{A} \mathbf{U}^T \quad (2)$$

其中, $\mathbf{A} = \text{diag}([\lambda_1, \lambda_2, \dots, \lambda_n])$ 为特征值对角阵, $\mathbf{U} = [\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_n]$ 为特征向量正交矩阵。对于图中单个节点 $\mathbf{x} \in \mathbf{X}$, 定义图傅里叶变换为 $\widehat{\mathbf{x}} = \mathbf{U}^T \mathbf{x}$, 逆变换为 $\mathbf{x} = \mathbf{U} \widehat{\mathbf{x}}$, 图信号频域滤波响应为 $\widehat{\mathbf{z}} = g_\theta(\mathbf{A}) \widehat{\mathbf{x}}$, 进行傅里叶逆变换即

$$\mathbf{z} = \mathbf{U} g_\theta(\mathbf{A}) \mathbf{U}^T \mathbf{x} \quad (3)$$

式(3)即为图卷积的基本形式, 其中 $g_\theta(\mathbf{A})$ 对应为可学习的滤波函数。

用 λ_{\max} 表示为 \mathbf{L} 的最大特征值, 令 $\widehat{\mathbf{A}} = \frac{2}{\lambda_{\max}} \mathbf{A} - \mathbf{I}_N$, 采用切比雪夫多项式 T_k 来近似 $g_\theta(\mathbf{A})$, 即

$$g_\theta(\mathbf{A}) \approx \sum_{k=0}^K \theta_k T_k(\widehat{\mathbf{A}}) \quad (4)$$

$\widehat{\mathbf{L}} = \frac{2}{\lambda_{\max}} \mathbf{L} - \mathbf{I}$ 为对应拉普拉斯矩阵缩放, 结合恒等式 $(\mathbf{U} \mathbf{A} \mathbf{U}^T)^k = \mathbf{U} \mathbf{A}^k \mathbf{U}^T$, 则式(3)可近似为

$$\mathbf{z} = \sum_{k=0}^K \theta_k T_k(\widehat{\mathbf{L}}) \mathbf{x} \quad (5)$$

本文中取 $K = 1$, 即只考虑节点的1阶连接信息, 令归一化的 \mathbf{L} 最大特征值 $\lambda_{\max} = 2$, 则式(5)可展开并简化为

$$\mathbf{z} = \theta_0 \mathbf{x} + \theta_1 (\mathbf{L} - \mathbf{I}_N) \mathbf{x} = \theta_0 \mathbf{x} - \theta_1 \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \mathbf{x} \quad (6)$$

防止模型过拟合, 采用 $\theta = \theta_0 = -\theta_1$ 进一步约束参数数量可得

$$\mathbf{z} = \theta (\mathbf{I} + \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}) \mathbf{x} \quad (7)$$

$\mathbf{I} + \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}$ 的特征值范围为 $[0, 2]$, 实际使用发现该算子易发生消失梯度, 采用 $\widehat{\mathbf{D}}^{-\frac{1}{2}} \widehat{\mathbf{A}} \widehat{\mathbf{D}}^{-\frac{1}{2}}$ 近似算子效果更佳, 其中 $\widehat{\mathbf{A}} = \mathbf{A} + \mathbf{I}$, $\widehat{D}_{ii} = \sum_j \widehat{A}_{ij}$ 。

针对一个含有 N 个特征长度为 C 的节点的图 $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in R^{N \times C}$ 表示图节点信号矩阵, 则对于全图 \mathbf{G} 的卷积操作可定义为

$$\mathbf{Z} = \widehat{\mathbf{D}}^{-\frac{1}{2}} \widehat{\mathbf{A}} \widehat{\mathbf{D}}^{-\frac{1}{2}} \mathbf{X} \Theta \quad (8)$$

其中, $\Theta \in R^{C \times F}$ 为卷积核参数矩阵, \mathbf{Z} 为输出节点信号, 特征长度为 F 。

基于式(8), 设第 j 层图卷积和 ReLU 的操作为 f_{θ_j} , 则针对图4所示的3层图卷积网络

$$\text{output} = \text{Softmax}(f_{\theta_3}(f_{\theta_2}(f_{\theta_1}(\mathbf{X})))) \quad (9)$$

采用交叉熵作为损失函数, 图卷积网络训练过程中只计算已标注的节点损失, 即

$$\text{LOSS} = - \sum_{i=1}^{n=N_i} y_i \text{output}_i \quad (10)$$

其中, N_i 表示已标注节点数量, y_i 表示该节点真实标签。

2.4 WSOINN-GCN的算法步骤

结合图1中给出的模型框架, 本文提出的 WSOINN-GCN 对图像数据自动标注及分类的算法步骤如下:

设需要分类的 N 张图片集合为 $\mathbf{Pictures} = \{\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_N\}$ 。

(1) 对所有 $\mathbf{I}_i \in \mathbf{Pictures}$ 特征提取, 得到特征数据集 $\mathbf{Ft} = \{\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_N\}$, 其与 $\mathbf{Pictures}$ 具有一一对应关系。

(2) 对原始特征数据分布进行预估, 初始化 WSOINN 超参数 p, W_{\max} 。

(3) 随机从 \mathbf{Ft} 选择两个特征数据初始化 WSOINN 节点集合 $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2\}$, 并初始化节点连接集合 \mathbf{E} 和节点胜利次数集合 win_times 为空集。

(4) 将所有 $\mathbf{F}_i \in \mathbf{Ft}$ 按顺序输入至 WSOINN 学习, 得到图结构 \mathbf{G} , 输出 $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$, 及对应的 \mathbf{E} , $\text{win_times} = \{t_1, t_2, \dots, t_m\}$, 由 WSOINN 算法可知, \mathbf{V} 与 win_times 一一对应, 对于任意 $\mathbf{v}_i \in \mathbf{V}$, 存在 $\mathbf{F}_j \in \mathbf{Ft}$ 与之对应。

(5) 根据胜利次数 t_i 值大小, 按照对应关系, 对所有 $\mathbf{v}_i \in \mathbf{V}$ 进行排序, 得到 $\mathbf{V}_{\text{order}}$ 。

(6) 选出 $\mathbf{V}_{\text{order}}$ 排序靠前一定比例的节点集合 \mathbf{V}_l , 根据 $\mathbf{V}, \mathbf{Ft}, \mathbf{Pictures}$ 对应关系, 查看原始图片 \mathbf{I}_i , 人工标注所有 $\mathbf{v}_i \in \mathbf{V}_l$, 得到 \mathbf{V}_l 的标签集合 $\mathbf{L}^{\text{labeled}}$ 。

(7) 根据节点数量、特征维度, 选择正则化方法, 搭建合适多层图卷积网络 GCN, 其中 GCN 末层接 Softmax 用于预测节点信息。

(8) 将 WSOINN 输出的图 $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ 输入至 GCN 预测所有节点信息, 根据 $\mathbf{V}, \mathbf{V}_l, \mathbf{L}^{\text{labeled}}$ 统计已标注节点的误差, 反向传播误差, 选择合适算法优化网络参数, 反复迭代直至训练误差不再降低。

(9) 将图 $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ 输入至已训练好的 GCN, 得到 \mathbf{V} 的标签集合 \mathbf{L} 。

(10) 对任意 $\mathbf{I}_i \in \mathbf{Pictures}$, \mathbf{I}_i 的标签为 $\mathbf{v}_j = \underset{\mathbf{v}_m \in \mathbf{V}}{\text{argmin}} \|\mathbf{F}_i - \mathbf{v}_m\|$ 的标签 L_j 。

由上述算法步骤可知, WSOINN 在原 SOINN 上引入了边连接权重表示两节点相似性, 从而使得图卷积神经网络 GCN 能够挖掘金相图之间的关

系，通过引入节点胜利次数 t_i 来选择少许具有代表性的重要节点进行人工标注，避免随机选择造成的模型不稳定，从而有机地将WSOINN与GCN结合起来，在减少人工标注的同时，实现了图像数据的高效分类。

3 实验结果及分析

本节先给出WSOINN-GCN模型参数的优选方法，再比较在不同节点标注率下模型的节点标注精度及金相图分类精度，最后给出其他常见方法的对比实验结果。

3.1 实验环境及模型指标

本实验硬件支持有CPU为i5-7500，4核4线程，主频3.41 GHz，内存12 GB，GPU为NVIDIA GeForce GTX 1060，显存6 GB，操作系统为win10，编程环境为spyder,Python3.7，框架平台为tensorflow。对于图卷积网络结构，采用Adam算法优化参数，初始学习率为0.01，dropout神经元失活的概率为0.5，Glorot_normal初始化参数，采用Early Stopping提前终止。统计精确率与召回率2个指标。

3.2 WSOINN-GCN模型参数分析

WSOINN在每输入样本的百分比例 p 后，会删除孤立节点，会影响最终节点输出数量。节点数过多可能含有噪声节点，节点数过少不能全面反映所有样本分布，从而间接影响自动标注精度。图5表示为列举了不同 p, W_{max} 值下，WSOINN获取拓扑图节点的数量情况，颜色越深代表产生的节点数越少，节点数最大值为865，最小值256。图6代表连接矩阵的稀疏程度(非0元素所占比例)，颜色越深表示连接矩阵越稀疏，由图5、图6可知，随着 p, W_{max} 增大，节点数随之增大，连接矩阵越稠密。

实验收集到的金相图总样本数有2432张，用 n, a 分别代表WSOINN输出图的节点数和连接矩阵中

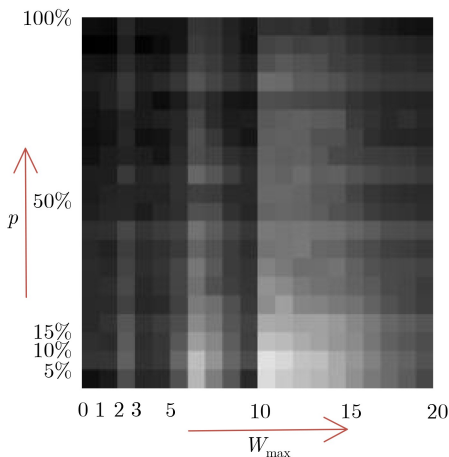


图 5 不同 p, W_{max} 下节点数

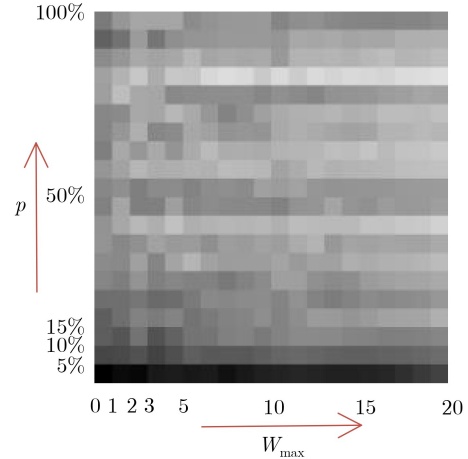


图 6 不同 p, W_{max} 下连接矩阵稀疏程度

非0元素个数，为保证精度同时加快运算，选择节点数为原数据量的1/10~1/6的WSOINN进一步分析。

表1列举了节点标注率为0.3、不同 p, W_{max} 值时模型对剩余节点的自动标注精度，其中Acc_w是按照胜利次数选择标注的结果，Acc_r是随机选择节点标注的结果。可见：

(1)与按照节点胜利次数选择节点标注相比，随机选择节点标注导致剩余节点自动标注精度时高时低，且自动标注精度未超过前者，按照节点胜利次数选择节点标注具有稳定的优势。

(2)无论哪种标注方式，随着节点数增多，剩余节点标注精度呈下降趋势，且在同等规模节点下，连接矩阵越稀疏(即越小)，节点标注精度越高。

根据表1，选取适合金相图分类的网络参数 $p = 10\%, W_{max} = 2$ ，此时节点数为 $n = 294, a = 324$ 。按照节点胜利次数选择节点标注继续完成后续实验。

表 1 标注率为0.3时，不同 p, W_{max} 值下剩余节点标注精度

	W_{max}, p					
	20, 100%	2, 10%	15, 45%	10, 5%	4, 30%	17, 70%
n	286	294	300	308	440	474
a	720	324	734	688	640	946
Acc_w(%)	85	93	84	90	89	88
Acc_r(%)	70	81	80	75	88	85

3.3 节点自动标注精度

为便于比较，搭建与WSOINN-GCN具有相同参数量的自组织增量-全连接神经网络(WSOINN-MLP)，其将WSOINN-GCN中GCN模块替换成多层感知机(MultiLayer Perceptron, MLP)，相应的图结构输入变为单节点输入，神经元激活函数采用ReLU，其中MLP结构如图7所示。

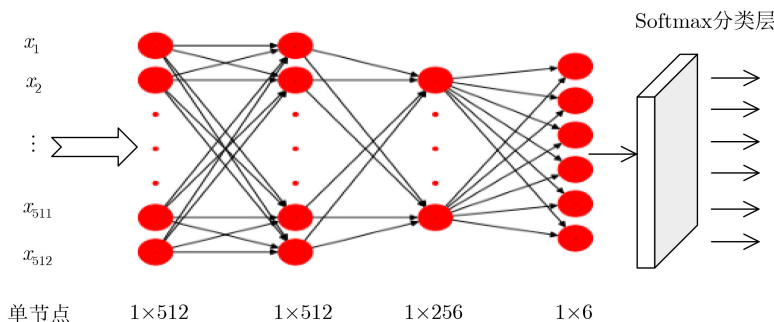


图7 MLP

表2列出了WSOINN-GCN,WSOINN-MLP在不同节点标注率、有无Dropout策略下,模型对拓扑图中剩余无标注节点的自动标注精度。可见:

(1)随着节点标注率增加,两类模型的精度都会增加。WSOINN-GCN性能要优于WSOINN-MLP,当节点标注率为0.3时,前者精度可达93%,而后者仅为86%。这是因为GCN会考虑单节点的1阶邻域信息,有更强的泛化能力,而MLP训练过程一直是单节点前向传播,其无法考虑节点之间的连接性。

(2)对于WSOINN-GCN而言,当节点标注率较低时(≤ 0.4),含有Dropout比未含有Dropout有更好的性能表现;当节点标注率较高时,Dropout会降低GCN自动标注精度。这是因为节点标注率低时,虽标注的节点剪代表性强,但图中高阶信息未完全挖掘,GCN易对已标注的节点过拟合,此时使用Dropout会增加模型的泛化能力,而当节点标注率增加时,已标注的节点有足够的全局剪代表性,加之GCN会获取节点1阶邻域信息,此时使用Dropout正则化方法往往会矫枉过正,降低模型性能。

(3)对于WSOINN-MLP而言,无论节点标注率多少,使用Dropout策略能一直提高其精度。因为MLP不能考虑节点间的关联信息,易出现过拟合现象,Dropout在一定程度上能弥补这一缺陷。

表2 不同节点标注率情况下剩余节点自动标注精度(%)

节点标注率(rate)	WSOINN- GCN		WSOINN- MLP	
	有Dropout	无Dropout	有Dropout	无Dropout
0.1	83	80	77	76
0.2	89	88	82	81
0.3	93	93	86	84
0.4	95	92	91	90
0.5	93	94	91	90
0.6	97	99	97	97
0.7	97	99	99	97

3.4 金相图分类精度

表3列举不同节点标注率下所有金相图自动分类精度;表4列举了节点标注率为0.3时,不同类别的金相图的准确率和召回率。由表4可见:(1)节点标注率越高,金相图自动分类准确率呈上升趋势,但是对于当节点标注率达到0.6时,WSOINN-GCN和WSOINN-MLP的性能均无法再提高;(2)高碳片状马氏体召回率虽高,精确率却低至74%,贝氏体精度虽高,但召回率低,存在交叉误判,原因可能是VGG16卷积层虽能暴力提取特征,但是对这种平均像素强度相近的灰度金相图,VGG16无法进一步获取到深度区分特征。

3.5 新模型的竞争力

表5列举了选择30%比例标注,不同方法的所有图片分类精度、在相同环境下所需的训练时间及

表3 不同节点标注率情况下所有金相图分类精度(%)

节点标注率(rate)	WSOINN-GCN		WSOINN- MLP	
	有Dropout	无Dropout	有Dropout	无Dropout
0.1	81	81	74	79
0.2	88	87	81	80
0.3	91	91	82	81
0.4	92	91	89	85
0.5	92	92	88	92
0.6	90	94	90	92
0.7	90	94	89	92

表4 节点标注率为0.3时,不同类别金相图的精确率与召回率(%)

金相图类别	精确率	召回率
下贝氏体	90	92
低碳板条马氏体	93	93
珠光体	99	92
贝氏体	93	73
铁素体	100	100
高碳片状马氏体	74	97

表5 选择30%标注, 不同方法的所有图片自动分类效果

	VGG-ICAM	SOINN	MLP	WSOINN-GCN	WSOINN-MLP
人工标注量(张)	730	730	730	88	88
训练时间(min)	240	1	1	2.5	2.5
分类精度(%)	74	39	51	91	82

人工所需标注图像数量。针对金相图数据集, 参照文献[11], 搭建了深度卷积网络VGG-ICAM, 标注30%的图像数据样本用于训练, 预测所有样本; SOINN先用VGG16卷积模块提取特征, 同VGG-ICAM标注30%样本用于拓扑学习, 最后基于欧氏距离判断所有图像数据; MLP先用VGG16提取特征, 然后搭建如图7所示的分类器, 标注30%样本用于训练, 预测所有样本; WSOINN-GCN, WSOINN-MLP用WSOINN学习所有样本后, 按照胜利次数选择30%节点标注, 用GCN, MLP自动标注剩余所有节点, 最后基于欧氏距离分类所有图像数据。

由表5可见: 按照选择30%标注, WSOINN-GCN, WSOINN-MLP所需的人工标注量仅为其他方案的12%; 在训练时间上, 相较于VGG-ICAM减少了99%, 相较于SOINN, MLP虽稍微增加了训练时间, 但精度大幅度提升, 具有明显的优势。

表6列举了所有图片分类精度达到90%以上, VGG-ICAM, SOINN, MLP, WSOINN-GCN, WSOINN-MLP所需的人工标注数量及训练时间。可见, WSOINN-GCN所需的人工标注量仅为VGG-ICAM的5.6%, SOINN, MLP的5.2%, 且训练时间相较于VGG-ICAM大幅缩减。

表6 分类精度达到90%, 不同方法所需的标注量及训练时间

	VGG-ICAM	SOINN	MLP	WSOINN-GCN	WSOINN-MLP
人工标注量(张)	1580	1702	1702	88	115
训练时间(min)	420	1.5	1.5	2.5	2.5

4 结论

(1) 针对深度学习中图像数据标注困难的问题, 融合拓扑学习与图卷积理论, 本文提出一种新的基于自组织增量-图卷积神经网络(WSOINN-GCN)的半监督学习方法。通过引入连接权重来改进自组织增量神经网络(WSOINN), 从而提取数据图结构, 并按照节点重要性指标节点胜利次数选择部分节点进行标注, 进而搭建图卷积网络(GCN)挖掘图中节点的潜在联系, 融合Dropout正则化手段与Adam算法对GCN进行网络参数寻优, 自动标注剩余节点信息, 并基于欧氏距离来自动分类金相图, 结果表明该模型具有可行性。

(2) 针对从某国家重点实验室扫描电子显微镜拍摄到的钢铁材料微观组织图片样本, 比较了不同节点标注率、有无Dropout对模型的影响, 结果表明: 随着节点标注率增加, WSOINN-GCN与WSOINN-MLP模型的精度都会增加, 且前者性能要优于后者, 当节点标注率为0.3时, 前者精度可达93%, 而后者仅为86%; 对于WSOINN-GCN而言, 当节点标注率较低时(≤ 0.4), 含有Dropout比未含有Dropout有更好的性能表现; 对于WSOINN-MLP而言, 无论节点标注率多少, 使用Dropout策略能一直提高其精度。

(3) 与现有的人工标注或其他监督学习算法相比, 本文所提出的WSOINN-GCN模型有效解决了实际应用时金相图片训练集数据人工标注困难的问题, 为金相图片数据标注、分类等提供了新的解决思路。在金相图片标注量仅为VGG等传统深度学习网络的12%时, 新模型比传统模型精度高, 分类准确度高达91%; 在达到相同的分类精度90%时, 人工标注量仅为传统模型的5.6%, 同时保证了效率优势。WSOINN-GCN具有自动提取数据图结构、实施半监督学习、动态调整网络结构等特性, 在图片数据标注、分类等领域有理论研究价值和广阔的应用前景。

参考文献

- [1] ZHAO H, WYNNE B P, and PALMIERE E J. A phase quantification method based on EBSD data for a continuously cooled microalloyed steel[J]. *Materials Characterization*, 2017, 123: 339-348. doi: [10.1016/j.matchar.2016.11.024](https://doi.org/10.1016/j.matchar.2016.11.024).
- [2] YANG Youwen, HE Chongxian, E Dianyu, et al. Mg bone implant: Features, developments and perspectives[J]. *Materials & Design*, 2020, 185: 108259. doi: [10.1016/j.matdes.2019.108259](https://doi.org/10.1016/j.matdes.2019.108259).
- [3] TERASAKI H, MIYAHARA Y, HAYASHI K, et al. Digital identification scheme for steel microstructures in low-carbon

- steel[J]. *Materials Characterization*, 2017, 129: 305–312. doi: [10.1016/j.matchar.2017.05.021](https://doi.org/10.1016/j.matchar.2017.05.021).
- [4] PANCHAL J H, KALIDINDI S R, and MCDOWELL D L. Key computational modeling issues in Integrated Computational Materials Engineering[J]. *Computer-Aided Design*, 2013, 45(1): 4–25. doi: [10.1016/j.cad.2012.06.006](https://doi.org/10.1016/j.cad.2012.06.006).
- [5] LU Xiaochong, ZHANG Xu, SHI Mingxing, *et al.* Dislocation mechanism based size-dependent crystal plasticity modeling and simulation of gradient nano-grained copper[J]. *International Journal of Plasticity*, 2019, 113: 52–73. doi: [10.1016/j.ijplas.2018.09.007](https://doi.org/10.1016/j.ijplas.2018.09.007).
- [6] WANG Hongda, RIVENSON Y, JIN Yiyin, *et al.* Deep learning enables cross-modality super-resolution in fluorescence microscopy[J]. *Nature Methods*, 2019, 16(1): 103–110. doi: [10.1038/s41592-018-0239-0](https://doi.org/10.1038/s41592-018-0239-0).
- [7] DECOST B L, LEI Bo, FRANCIS T, *et al.* High throughput quantitative metallography for complex microstructures using deep learning: A case study in ultrahigh carbon steel[J]. *Microscopy and Microanalysis*, 2019, 25(1): 21–29. doi: [10.1017/S1431927618015635](https://doi.org/10.1017/S1431927618015635).
- [8] PAULY J, BRITZ D, and MÜCKLICH F. Advanced microstructure classification using data mining methods[C]. In *TMP - 5th International Conference on ThermoMechanical Processing*, Milan, Italy, 2016: 12–25.
- [9] CHOWDHURY A, KAUTZ E, YENER B, *et al.* Image driven machine learning methods for microstructure recognition[J]. *Computational Materials Science*, 2016, 123: 176–187. doi: [10.1016/j.commatsci.2016.05.034](https://doi.org/10.1016/j.commatsci.2016.05.034).
- [10] AZIMI S M, BRITZ D, ENGSTLER M, *et al.* Advanced steel microstructural classification by deep learning methods[J]. *Scientific Reports*, 2018, 8(1): 2128. doi: [10.1038/s41598-018-20037-5](https://doi.org/10.1038/s41598-018-20037-5).
- [11] 李维刚, 谌竟成, 范丽霞, 等. 基于卷积神经网络的钢铁材料微观组织自动辨识[J]. *钢铁研究学报*, 2020, 32(1): 33–43. doi: [10.13228/j.boyuan.issn1001-0963.20190147](https://doi.org/10.13228/j.boyuan.issn1001-0963.20190147).
LI Weigang, SHEN Jingcheng, FAN Lixia, *et al.* Automatic identification of microstructure of iron and steel material based on convolutional neural network[J]. *Journal of Iron and Steel Research*, 2020, 32(1): 33–43. doi: [10.13228/j.boyuan.issn1001-0963.20190147](https://doi.org/10.13228/j.boyuan.issn1001-0963.20190147).
- [12] LI Bentian and PI Dechang. Network representation learning: A systematic literature review[J]. *Neural Computing and Applications*, 2020, 32(21): 16647–16679. doi: [10.1007/s00521-020-04908-5](https://doi.org/10.1007/s00521-020-04908-5).
- [13] 康世泽, 吉立新, 张建朋. 一种基于图注意力网络的异质信息网络表示学习框架[J]. *电子与信息学报*, 2021, 43(4): 915–922. doi: [10.11999/JEIT200034](https://doi.org/10.11999/JEIT200034).
KANG Shize, JI Lixin, and ZHANG Jianpeng. Heterogeneous information network representation learning framework based on graph attention network[J]. *Journal of Electronics & Information Technology*, 2021, 43(4): 915–922. doi: [10.11999/JEIT200034](https://doi.org/10.11999/JEIT200034).
- [14] 李世宝, 张益维, 刘建航, 等. 基于知识图谱共同邻居排序采样的推荐模型[J]. *电子与信息学报*, 待发表. doi: [10.11999/JEIT200735](https://doi.org/10.11999/JEIT200735).
LI Shibao, ZHANG Yiwei, LIU Jianhang, *et al.* Recommendation model based on public neighbor sorting and sampling of knowledge graph[J]. *Journal of Electronics & Information Technology*, To be published. doi: [10.11999/JEIT200735](https://doi.org/10.11999/JEIT200735).
- [15] ORTEGA A, FROSSARD P, KOVAČEVIĆ J, *et al.* Graph signal processing: Overview, challenges, and applications[J]. *Proceedings of the IEEE*, 2018, 106(5): 808–828. doi: [10.1109/JPROC.2018.2820126](https://doi.org/10.1109/JPROC.2018.2820126).
- [16] HAMILTON W L, YING R, and LESKOVEC J. Inductive representation learning on large graphs[J]. arXiv preprint arXiv: 1706.02216, 2017.
- [17] KIPF T N and WELING M. Semi-supervised classification with graph convolutional networks[J]. arXiv preprint arXiv: 1609.02907, 2016.
- [18] CHEN Jie, MA Tengfei, and XIAO Cao. FastGCN: Fast learning with graph convolutional networks via importance sampling[J]. arXiv preprint arXiv: 1801.10247, 2018.
- [19] SHEN Furao and HASEGAWA O. An incremental network for on-line unsupervised classification and topology learning[J]. *Neural Networks*, 2006, 19(1): 90–106. doi: [10.1016/j.neunet.2005.04.006](https://doi.org/10.1016/j.neunet.2005.04.006).
- 李维刚: 1977年生, 教授, 博士生导师, 研究方向为人工智能与机器学习算法.
谌竟成: 1997年生, 硕士生, 研究方向为图像处理.
谢璐: 1996年生, 博士生, 研究方向为语义分割.
赵云涛: 1983年生, 副教授, 研究方向为3维点云数据处理.

责任编辑: 余蓉