

# 基于突触巩固机制的前馈小世界神经网络设计

李文静<sup>1,2,3,4,5</sup> 李治港<sup>1,2,3,4,5</sup> 乔俊飞<sup>1,2,3,4,5</sup>

**摘要** 小世界神经网络具有较快的收敛速度和优越的容错性,近年来得到广泛关注.然而,在网络构造过程中,随机重连可能造成重要信息丢失,进而导致网络精度下降.针对该问题,基于 Watts-Strogatz (WS) 型小世界神经网络,提出了一种基于突触巩固机制的前馈小世界神经网络 (Feedforward small-world neural network based on synaptic consolidation, FSWNN-SC). 首先,使用网络正则化方法对规则前馈神经网络进行预训练,基于突触巩固机制,断开网络不重要的权值连接,保留重要的连接权值;其次,设计重连规则构造小世界神经网络,在保证网络小世界属性的同时实现网络稀疏化,并使用梯度下降算法训练网络;最后,通过 4 个 UCI 基准数据集和 2 个真实数据集进行模型性能测试,并使用 Wilcoxon 符号秩检验对对比模型进行显著性差异检验.实验结果表明:所提出的 FSWNN-SC 模型在获得紧凑的网络结构的同时,其精度显著优于规则前馈神经网络及其他 WS 型小世界神经网络.

**关键词** 小世界神经网络, 突触巩固机制, 网络正则化, 重连规则, Wilcoxon 符号秩检验

**引用格式** 李文静, 李治港, 乔俊飞. 基于突触巩固机制的前馈小世界神经网络设计. 自动化学报, 2023, 49(10): 2145–2158

**DOI** 10.16383/j.aas.c220638

## Structure Design for Feedforward Small-world Neural Network Based on Synaptic Consolidation Mechanism

LI Wen-Jing<sup>1,2,3,4,5</sup> LI Zhi-Gang<sup>1,2,3,4,5</sup> QIAO Jun-Fei<sup>1,2,3,4,5</sup>

**Abstract** Because of faster convergence speed and superior fault tolerance, small-world neural network has attracted wide attention in recent years. However, in the construction process, it may cause the loss of important information due to random reconnection, which may lead to the decline of network accuracy. To solve this problem, derived from the Watts-Strogatz (WS) small-world neural network, a feedforward small-world neural network based on synaptic consolidation (FSWNN-SC) mechanism is proposed in this study. Firstly, the regular feedforward neural network is pre-trained by using the network regularization method. Based on the synaptic consolidation mechanism, the unimportant connection weights of the network are disconnected and the important connection weights are retained. Secondly, the rewiring rules are designed to construct a small-world neural network, which can realize the sparseness of the network while ensuring the small-world properties of the network. The gradient descent algorithm is used to train the network. Finally, four UCI benchmark experiments and two practical experiments are carried out to evaluate the model performance, and the Wilcoxon signed-ranks test is performed to test the significant differences between comparative models. Experimental results show that the FSWNN-SC model proposed in this study not only obtains a compact network structure, but also has significantly better accuracy than regular feedforward neural networks and other WS small-world neural networks.

**Key words** Small-world neural networks, synaptic consolidation mechanism, network regularization, rewiring rule, Wilcoxon signed-ranks test

**Citation** Li Wen-Jing, Li Zhi-Gang, Qiao Jun-Fei. Structure design for feedforward small-world neural network based on synaptic consolidation mechanism. *Acta Automatica Sinica*, 2023, 49(10): 2145–2158

收稿日期 2022-08-11 录用日期 2022-11-12

Manuscript received August 11, 2022; accepted November 12, 2022

国家重点研发计划 (2021ZD0112301), 国家自然科学基金 (62173008, 62021003, 61890930-5) 资助

Supported by National Key Research and Development Program of China (2021ZD0112301) and National Natural Science Foundation of China (62173008, 62021003, 61890930-5)

本文责任编辑 何晖光

Recommended by Associate Editor HE Hui-Guang

1. 北京工业大学信息学部 北京 100124 2. 北京人工智能研究院 北京 100124 3. 智能感知与自主控制教育部工程研究中心 北京 100124 4. 计算智能与智能系统北京市重点实验室 北京 100124 5. 智慧环保北京实验室 北京 100124

神经网络是受生物神经网络启发而设计出来的一种数学计算模型,具有良好的容错能力、万能的逼近特性以及优越的自适应和自学习功能<sup>[1]</sup>,同时具备高速并行处理信息的结构<sup>[2]</sup>,可以解决复

1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124 2. Beijing Artificial Intelligence Institute, Beijing 100124 3. Engineering Research Center of Intelligence Perception and Autonomous Control, Ministry of Education, Beijing 100124 4. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124 5. Beijing Laboratory for Intelligent Environmental Protection, Beijing 100124

杂的工程问题. 这些优点使得人工神经网络成为当今最成功的人工智能模型之一<sup>[3]</sup>, 已广泛地应用于众多领域, 如非线性系统建模<sup>[4-5]</sup>、数据挖掘<sup>[6-7]</sup>、计算机视觉<sup>[8-9]</sup>和自然语言处理<sup>[10-11]</sup>等.

1998年, Watts等<sup>[12]</sup>发现许多生物网络、社会网络介于随机网络和规则网络之间, 其拓扑结构呈现出稀疏特性, 且具有较大的聚类系数和较短的平均路径长度, 并将其定义为小世界网络. 随着人们对生物神经网络研究的深入, 借助弥散张量成像、磁共振脑功能成像等技术, 研究者发现生物神经网络在结构和功能上均呈现出小世界属性<sup>[13-15]</sup>. 研究表明, 生物神经网络较短的平均路径长度可以加快信息在网络中的传播速度, 较大的聚类系数能够增加网络的容错能力<sup>[16]</sup>. 尽管人工神经网络通过构建大量节点(神经元)之间的相互连接模拟人脑的信息处理方式, 但是其拓扑结构及功能与生物神经网络仍相去甚远.

近年来, 已有很多学者尝试将小世界属性引入人工神经网络设计中, 旨在提高人工神经网络的信息处理及容错能力. 前馈神经网络(Feedforward neural network, FNN)由于结构简单灵活<sup>[17]</sup>, 且可以以任意精度逼近任一连续函数<sup>[18-19]</sup>, 在模式识别、函数逼近等方面得到广泛应用<sup>[20-21]</sup>. 目前, 基于FNN的小世界神经网络设计得到越来越多的关注. Simard等<sup>[22]</sup>以多层感知器为基础模型, 建立了一种多层前馈小世界神经网络, 发现小世界神经网络比规则网络和随机网络有更快的学习速度和更高的精度. Li等<sup>[23]</sup>探究了重连概率对小世界神经网络性能的影响, 发现重连概率在0.1附近时, 网络收敛速度最快. 可见, 将小世界属性融入人工神经网络结构设计能够提升网络的收敛速度和泛化能力. 基于此, 小世界神经网络已广泛地应用到智能控制<sup>[24-25]</sup>、风力预测<sup>[26]</sup>、医疗诊断<sup>[27-28]</sup>、污水处理<sup>[29-30]</sup>等多个领域, 取得了良好的效果.

小世界神经网络的拓扑结构直接影响着网络的性能. 目前, 构造小世界神经网络的方法主要包括Watts-Strogatz (WS)<sup>[12]</sup>和Newman-Watts (NW)<sup>[31]</sup>两种方式. WS型前馈小世界神经网络通过在FNN上以一定概率随机断开相邻层连接再进行随机跨层重连实现小世界网络构建, 而NW型前馈小世界神经网络则是通过在FNN上直接随机跨层加边进行小世界网络构造. 从网络的构造方式可以看出, WS型小世界神经网络相比NW型小世界神经网络拓扑结构更加稀疏, 因此得到更多学者的关注. 在WS方式构造的基础上, 不少学者通过改进其断开或重连策略实现构造方式的优化. 例如, 李小虎等<sup>[32]</sup>

对随机断开连接加以限制, 即禁止断开最后一个隐含层与输出层之间的连接, 以防止孤立神经元的产生. 王爽心等<sup>[33]</sup>提出基于层连优化的小世界神经网络的改进算法, 引入了同层节点重连的策略, 改善了小世界神经网络聚类系数偏低的问题, 并且发现输入层和输出层直接相连会造成网络性能下降. 此外, 近年来研究学者在WS构造方式的基础上对网络进行稀疏化, 以进一步提高网络的泛化性能. Guo等<sup>[34]</sup>提出了一种基于E-信息熵的剪枝算法用于稀疏化WS型前馈小世界神经网络, 在一定程度上改善了因网络结构过大而出现拟合的问题. Li等<sup>[30]</sup>利用节点的Katz中心性衡量网络中节点的重要性, 删除不重要的节点使网络结构更加紧凑, 同时提高了网络的泛化性能. 尽管以上研究通过优化网络结构提升了网络性能, 但是在WS型构造方式中, 网络的随机跨层重连是在随机断开相邻层间连接的基础上实现的, 而连接断开的随机性可能会导致网络重要信息丢失, 在一定程度上影响网络的建模精度.

针对以上问题, 本文提出一种基于突触巩固机制<sup>[35-36]</sup>的前馈小世界神经网络(Feedforward small-world neural network based on synaptic consolidation, FSWNN-SC). 首先, 使用正则化方法对FNN进行预训练, 基于突触巩固机制选择性断开网络连接; 其次, 设计小世界网络重连规则, 同时实现网络的稀疏化, 并采用梯度下降学习算法训练网络; 最后, 通过4个UCI基准数据集和2个真实数据集进行模型性能测试, 并使用Wilcoxon符号秩检验<sup>[37]</sup>对实验结果进行显著性分析.

## 1 背景知识

### 1.1 前馈神经网络

前馈神经网络, 又称多层感知器, 采用级联方式连接实现信息的前向传导, 其结构由输入层、隐含层和输出层组成, 如图1(以4层为例)所示. 假设FNN共包含 $L$ 层, 使用 $\mathbf{X}$ 表示输入数据,  $y^l$  ( $1 \leq l \leq L$ )表示第 $l$ 层输出, FNN各层功能及表示详述如下.

1) 输入层. 输入层将输入数据导入FNN, 该层神经元激活函数常采用线性的, 以得到数据的原始特征. 若输入数据为 $n$ 维, 即 $\mathbf{X} = [x_1, x_2, \dots, x_n]^T$ , 则输入层包含 $n$ 个神经元, 其中第 $i$ 个神经元的输出为

$$y_i^1 = x_i, 1 \leq i \leq n \quad (1)$$

2) 隐含层. FNN包含至少一层隐含层, 通过激活函数实现该层输入数据的非线性映射, 激活函数

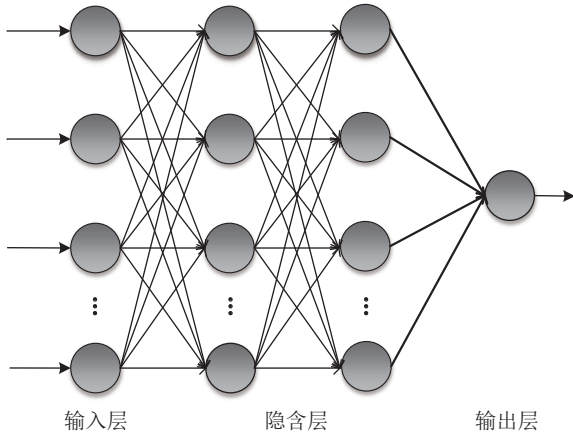


图 1 前馈神经网络结构示意图

Fig.1 The architecture of feedforward neural network

采用 sigmoid 函数, 因其导数便于计算而广泛使用. FNN 第  $l$  ( $1 < l < L$ ) 层 (即第  $l-1$  个隐含层) 的第  $j$  个神经元的输出为

$$y_j^l = f \left( \sum_{i=1}^{n_{l-1}} w_{ij}^{l-1} y_i^{l-1} \right) \quad (2)$$

其中,  $w_{ij}^{l-1}$  是第  $l-1$  层第  $i$  个神经元与第  $l$  层的第  $j$  个神经元之间的连接权值,  $n_{l-1}$  表示第  $l-1$  层的神经元个数,  $f(\cdot)$  为激活函数.

3) 输出层. 输出层是隐含层输出的集成, 该层激活函数常采用线性的, 进而实现对隐含层的线性加权, 输出层神经元个数视系统问题而定. 以含有一个输出神经元的 FNN 为例, 其输出为

$$y^L = \sum_{i=1}^{n_{L-1}} w_i^{L-1} y_i^{L-1} \quad (3)$$

其中,  $w_i^{L-1}$  是最后一个隐含层第  $i$  个神经元与输出神经元之间的连接权值. 若输出含有多个神经元, 可按照式 (3) 做相应扩展.

## 1.2 小世界属性

1998 年, Watts 等<sup>[12]</sup> 发现生物、技术和社交等网络的连接方式介于规则网络的“规则”和随机网络的“无序”之间, 具有较大的聚类系数和较短的特征路径长度等特征, 将其定义为小世界网络. 图论是描述网络特征的重要工具, 网络可以看作一个无向图, 由节点和边组成, 则网络的平均聚类系数  $AC$  和平均最短路径长度  $AL$  可由式 (4) 和式 (5) 计算得到

$$AC = \frac{1}{N} \sum_{i=1}^N \frac{2E_i}{D_i(D_i - 1)} \quad (4)$$

$$AL = \frac{1}{N(N-1)} \sum_{i \neq j \in G} l_{ij} \quad (5)$$

其中,  $N$  是网络中的节点数;  $D_i$  是节点  $i$  的度, 即节点  $i$  的相邻节点数, 则  $D_i(D_i - 1)/2$  表示这些相邻节点理论最多连接数量; 而  $E_i$  是这些相邻节点的实际连接数量;  $l_{ij}$  表示节点  $i$  到节点  $j$  的最短距离. 当网络的聚类系数远大于随机网络的聚类系数, 且特征路径长度接近于随机网络时, 即  $AC \gg AC_{\text{rand}}$  且  $AL \approx AL_{\text{rand}}$  时, 该网络具有小世界属性. 因此, 定义小世界网络属性指标为

$$\eta = \frac{AC}{\frac{AC_{\text{rand}}}{AL_{\text{rand}}}} \quad (6)$$

当网络满足  $\eta > 1$  时, 该网络为小世界网络<sup>[38]</sup>.

对于前述前馈神经网络, 其连接遵循一定的规则, 即相邻层之间为全连接结构, 同一层内神经元之间无连接且不存在跨层连接. 因此, 对任一神经元, 其近邻神经元之间不存在边, 由式 (4) 可得网络聚类系数为 0. 通过在前馈神经网络规则连接的基础上进行随机跨层重连, 可生成小世界神经网络<sup>[30, 32-34]</sup>. 相对于规则前馈神经网络, 小世界神经网络由于存在跨层连接, 其聚类系数始终大于 0, 且由于位于不同层的两个神经元之间的最短路径减小使得其特征路径长度减小. 相对于完全随机网络, 小世界神经网络的拓扑结构含有随机网络的“无序”特性, 使其特征路径长度接近于完全随机网络的特征路径长度, 同时小世界神经网络的拓扑结构仍然保留一定的“规则”连接, 这使得小世界神经网络的聚类系数远大于完全随机网络的聚类系数, 因此具备小世界属性.

## 1.3 突触巩固机制

在生物神经网络中, 突触是神经细胞之间的连接, 也是细胞间信息传递的“桥梁”. 研究表明, 突触具有可塑性, 可以自主调节其连接强度, 突触的形态和功能可发生较为持久的改变. 当学习特定任务时, 有些突触兴奋性会增强, 有些会被抑制, 这就是突触巩固机制<sup>[39]</sup>. 突触巩固的实现依赖突触的长时程增强 (Long-term potentiation, LTP)<sup>[40]</sup> 和长时程抑制 (Long-term depression, LTD)<sup>[41]</sup>. 如果两个神经元之间存在持久的信息传递, LTP 将加强它们之间的突触连接; 如果两个神经元之间的信息传递弱且呈间歇性, LTD 会抑制其间突触连接. 因此, 突触巩固机制可以选择性地增强和削弱特定的突触. 如图 2 所示, 图 2(a) 是神经元与突触模型, 图 2(b)

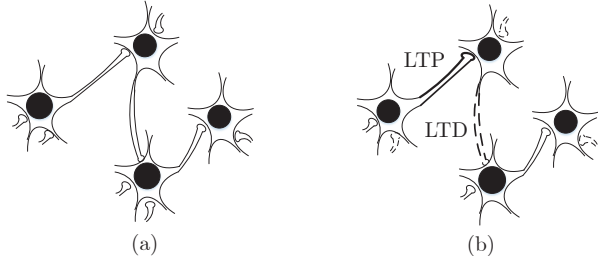


图 2 突触巩固

Fig. 2 Synaptic consolidation

展示了突触巩固机制作用下突触的变化. 其中, 粗线表示重要的突触在突触巩固的作用下会增强其连接强度 (LTP), 而虚线表示不重要的突触在突触巩固作用下会消失 (LTD).

研究者通过模拟生物神经系统的结构特征和突触的可塑性机制, 设计人工神经网络结构及学习规则. 突触巩固机制在人工神经网络结构设计中也得到了广泛应用, 如模型压缩<sup>[42]</sup>和增量学习<sup>[43]</sup>等.

## 2 基于突触巩固机制的小世界神经网络设计

受到突触巩固机制的启发, 本文提出一种前馈小世界神经网络 (FSWNN-SC). 首先, 通过前馈神经网络正则化对其进行预训练, 保留网络中重要的权值连接、断开网络中不重要的权值连接以模拟生物神经网络的突触巩固; 其次, 制定小世界网络重连规则, 对断开的连接进行随机跨层重连, 实现小世界神经网络的构造; 最后使用梯度下降法对网络进行训练, 实现网络权值更新.

### 2.1 FNN 预训练

对于规则 FNN, 本文首先使用网络正则化方法对其进行预训练, 旨在通过网络稀疏化保留网络重要权值连接、断开不重要的权值连接. 具体步骤如下:

1) 初始化. 初始化网络结构包括层数和神经元个数等参数, 网络的连接权值设置为  $[-1, 1]$  区间内的随机数. 设置惩罚系数  $\lambda$ 、网络学习率  $\mu$ 、预训练迭代步数  $S$  和重连概率  $P$ .

2) 定义损失函数. 即

$$L_1(\mathbf{W}) = E(\mathbf{W}) + \lambda g(\mathbf{W}) \quad (7)$$

该损失函数由两项组成, 其中第 1 项  $E(\mathbf{W})$  为误差项, 通过式 (8) 计算得到

$$E(\mathbf{W}) = \frac{1}{2} \sum_{m=1}^M (d_m - y_m^L)^2 \quad (8)$$

其中,  $M$  为训练样本数,  $d_m$  与  $y_m^L$  为样本  $m$  的期望

与实际输出. 第 2 项  $\lambda g(\mathbf{W})$  是正则化项, 其中  $\lambda$  是惩罚系数,  $g(\mathbf{W})$  为  $L_q$  ( $0 \leq q \leq 1$ ) 范数, 即

$$g(\mathbf{W}) = \|\mathbf{W}\|_q \quad (9)$$

此处分别列出  $q = 0, 1/2, 1$  时的计算式, 即

$$g(\mathbf{W}) = \|\mathbf{W}\|_0 = \sum_{l=1}^{L-1} \sum_{i=1}^{n_l} \sum_{j=1}^{n_{l+1}} (1 - e^{-\beta |w_{ij}^l|}) \quad (10)$$

$$g(\mathbf{W}) = \|\mathbf{W}\|_{\frac{1}{2}} = \sum_{l=1}^{L-1} \sum_{i=1}^{n_l} \sum_{j=1}^{n_{l+1}} |w_{ij}^l|^{\frac{1}{2}} \quad (11)$$

$$g(\mathbf{W}) = \|\mathbf{W}\|_1 = \sum_{l=1}^{L-1} \sum_{i=1}^{n_l} \sum_{j=1}^{n_{l+1}} |w_{ij}^l| \quad (12)$$

其中,  $L$  为网络的总层数,  $w_{ij}^l$  代表第  $l$  层的第  $i$  个神经元与第  $l+1$  层的第  $j$  个神经元的连接权值. 由于其在原点处不可导, 将其在原点处进行平滑<sup>[44]</sup>, 即

$$g(\mathbf{W}) = \begin{cases} \|\mathbf{W}\|_q, & |w| \geq a \\ -\frac{1}{8a^3}w^4 + \frac{3}{4a}w^2 + \frac{3}{8}a, & |w| < a \end{cases} \quad (13)$$

其中,  $a$  为接近于 0 的常数, 本文中取值  $a = 0.005$ .

3) 使用梯度下降法更新网络权值. 更新式为

$$\mathbf{W}(t+1) = \mathbf{W}(t) + \Delta \mathbf{W}(t) \quad (14)$$

其中,  $t$  为迭代步数,  $\Delta \mathbf{W}(t)$  是权值变化矩阵, 矩阵中第  $l$  层的第  $i$  个神经元与第  $l+1$  层的第  $j$  个神经元的连接权值变化量  $\Delta w_{ij}^l$  可由式 (15) 计算得到

$$\Delta w_{ij}^l(t) = -\mu \left( \frac{\partial E(\mathbf{W})}{\partial w_{ij}^l(t)} + \lambda \frac{\partial g(\mathbf{W})}{\partial w_{ij}^l(t)} \right) = -\mu \left( \delta_j^{l+1}(t) y_i^l(t) + \lambda \frac{\partial g(\mathbf{W})}{\partial w_{ij}^l(t)} \right) \quad (15)$$

其中,  $\mu$  是学习率,  $\lambda$  是惩罚系数.  $\delta_j^{l+1}$  按照式 (16) 计算得到

$$\delta_j^{l+1} = \begin{cases} \mathbf{W}_j^{l+1} \delta^{l+2} (y_j^{l+1})', & l \in [1, L-2] \\ \sum_{m=1}^M (d_m - y_m^L), & l = L-1 \end{cases} \quad (16)$$

其中,  $\mathbf{W}_j^{l+1}$  是第  $l+1$  层的第  $j$  个神经元与第  $l+2$  层神经元的权值向量.  $(y_j^{l+1})'$  是第  $l+1$  层第  $j$  个神经元的输出的导数, 通过式 (17) 计算得到

$$(y_j^{l+1})' = y_j^{l+1} (1 - y_j^{l+1}) \quad (17)$$

当迭代次数达到预设步数  $S$  时, 训练停止.

4) 网络稀疏化. 根据设定重连概率  $P$ , 计算断开连接数  $N_c$ . 将完成预训练后的网络权值降序排列, 设置前  $N_c$  个连接权值为 0.

通过网络预训练, 规则 FNN 将按照重连概率  $P$  断开网络不重要的权值连接, 其他权值连接被保留.

## 2.2 小世界神经网络构造方法

本文对 WS 重连规则进行改进, 设计小世界神经网络构造方法.

标记经过预训练后断开连接的神经元, 假设  $l$  层的第  $i$  个神经元与  $l+1$  层的第  $j$  个神经元之间的权值连接被断开, 起始及终止神经元分别记为  $v_i^l$  和  $v_j^{l+1}$ .

制定重连规则如下: 神经元  $v_i^l$  向后跨层寻找可重连的神经元. 若存在, 则对其进行随机重连, 并在  $[-1, 1]$  范围内对连接权值随机赋值; 若不存在, 则从神经元  $v_j^{l+1}$  向前跨层寻找可连接的神经元进行随机重连, 并在  $[-1, 1]$  范围内对连接权值随机赋值. 遍历所有标记神经元直至完成重连, 若出现孤立神经元, 则将其删除.

同时设定约束条件: 输入层与输出层的神经元不能直接相连.

本文所提出的小世界神经网络构造方法可以用图 3 表示, 规则网络先经过预训练断开网络中不重要的连接 (图 3(a)), 然后按照重连规则进行跨层重连 (图 3(b)), 最后删除孤立神经元 (图 3(c)).

## 2.3 小世界神经网络结构描述

在规则前馈神经网络的基础上, 本文通过以上重连规则引入跨层连接构造小世界神经网络. 两者输入层的数学描述相同, 如式 (1) 所示. 而由于引入了跨层连接, 小世界神经网络隐含层及输出层的数学描述与规则前馈神经网络有所不同.

对于隐含层, 以小世界神经网络第  $l$  ( $1 < l < L$ ) 层的第  $j$  个神经元为例, 其输出通过式 (18) 计算得到

$$y_j^l = f \left( \sum_{s=1}^{l-1} \sum_{i=1}^{n_s} w_{ij}^{sl} y_i^s \right) \quad (18)$$

其中,  $n_s$  表示第  $s$  层神经元的个数,  $f(\cdot)$  为激活函数,  $w_{ij}^{sl}$  为第  $s$  层第  $i$  个神经元与第  $l$  层的第  $j$  个神经元之间的连接权值. 若不存在连接, 则  $w_{ij}^{sl}$  为 0. 由式 (18) 可以看出, 第  $l$  层隐含层的输入需要同时考虑来自前面所有层的输出.

对于输出层, 其输入需要同时考虑来自所有隐含层的输出, 其输出通过式 (19) 计算得到

$$y^L = \sum_{s=1}^{L-1} \sum_{i=2}^{n_s} w_i^{sL} y_i^s \quad (19)$$

其中,  $w_i^{sL}$  表示第  $s$  层的第  $i$  个神经元与输出层神经元的连接权值.

## 2.4 小世界神经网络参数学习

本文使用梯度下降算法更新小世界神经网络的连接权值. 首先, 基于误差函数定义损失函数, 即

$$L_2(\mathbf{W}) = E(\mathbf{W}) = \frac{1}{2} \sum_{m=1}^M (d_m - y_m^L)^2 \quad (20)$$

其中,  $M$  是训练样本数,  $d_m$  与  $y_m^L$  分别表示样本  $m$  的期望输出与实际输出.

按照式 (21) 更新连接权值, 即

$$\mathbf{W}(t+1) = \mathbf{W}(t) + \Delta \mathbf{W}(t) \quad (21)$$

其中,  $t$  为迭代步数,  $\Delta \mathbf{W}$  为权值变化矩阵. 矩阵中第  $l$  层的第  $i$  个神经元与第  $s$  层的第  $j$  个神经元的

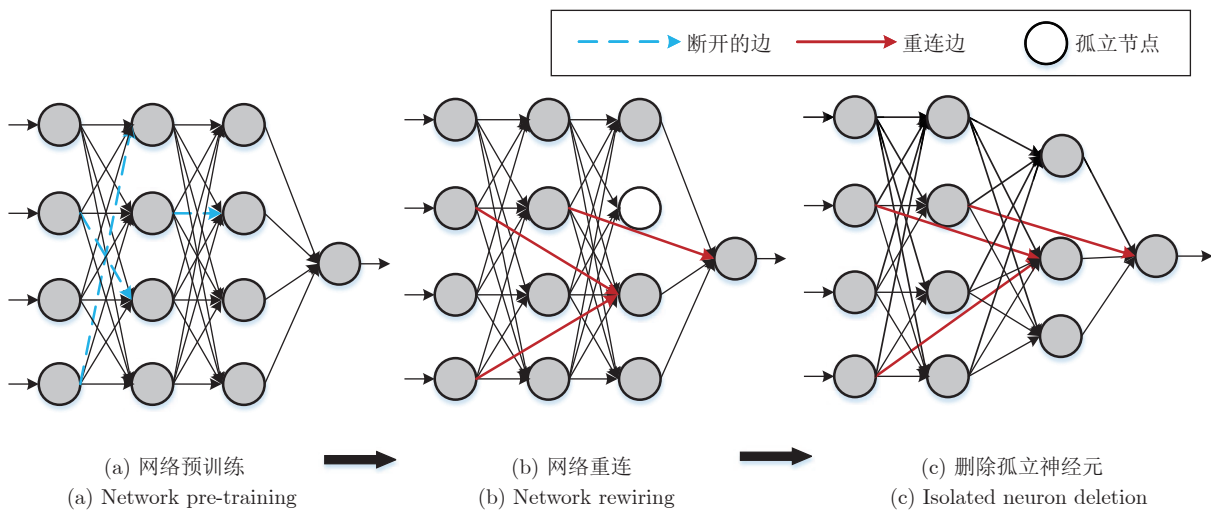


图 3 基于突触巩固小世界神经网络构造流程

Fig.3 Construction process of small-world neural network based on synaptic consolidation



连接权值变化量  $\Delta w_{ij}^{ls}$  可通过式 (22) 计算得到

$$\Delta w_{ij}^{ls}(t) = -\mu \frac{\partial E(\mathbf{W})}{\partial w_{ij}^{ls}(t)} = -\mu \delta_j^s(t) y_i^l(t) \quad (22)$$

其中,  $\mu$  是学习率,  $y_i^l$  为第  $l$  ( $1 \leq l < s \leq L$ ) 层的第  $i$  个神经元的输出,  $\delta_j^s$  可由式 (23) 计算得到

$$\delta_j^s = \begin{cases} \sum_{d=s+1}^L \mathbf{W}_j^{sd} \delta^d (y_j^s)', & s \in [2, L-1] \\ \sum_{m=1}^M (d_m - y_m^L), & s = L \end{cases} \quad (23)$$

其中,  $\delta^d$  是第  $d$  层神经元的误差;  $\mathbf{W}_j^{sd}$  为第  $s$  层的第  $j$  个神经元与第  $d$  层神经元的连接权值向量;  $(y_j^s)'$  是第  $s$  层第  $j$  个神经元的输出的导数, 通过式 (24) 计算得到.

$$(y_j^s)' = y_j^s(1 - y_j^s) \quad (24)$$

当迭代步数达到最大迭代步数 ( $iter_{\max}$ ) 或训练误差达到均方根误差 (Root mean squared error, RMSE) 的期望值  $RMSE_d$  时, 网络训练结束.

### 2.5 FSWNN-SC 算法流程

本文在规则前馈神经网络的基础上, 通过网络正则化断开不重要的权值连接, 设计网络重连规则构造小世界神经网络, 使用梯度下降算法更新网络权值, 算法流程图如图 4 所示.

## 3 实验仿真

通过实验仿真, 本文首先研究预训练关键参数 (重连概率  $P$  及预训练次数  $S$ ) 对实验结果的影响, 为后续实验的参数选取提供实验依据; 其次, 从建

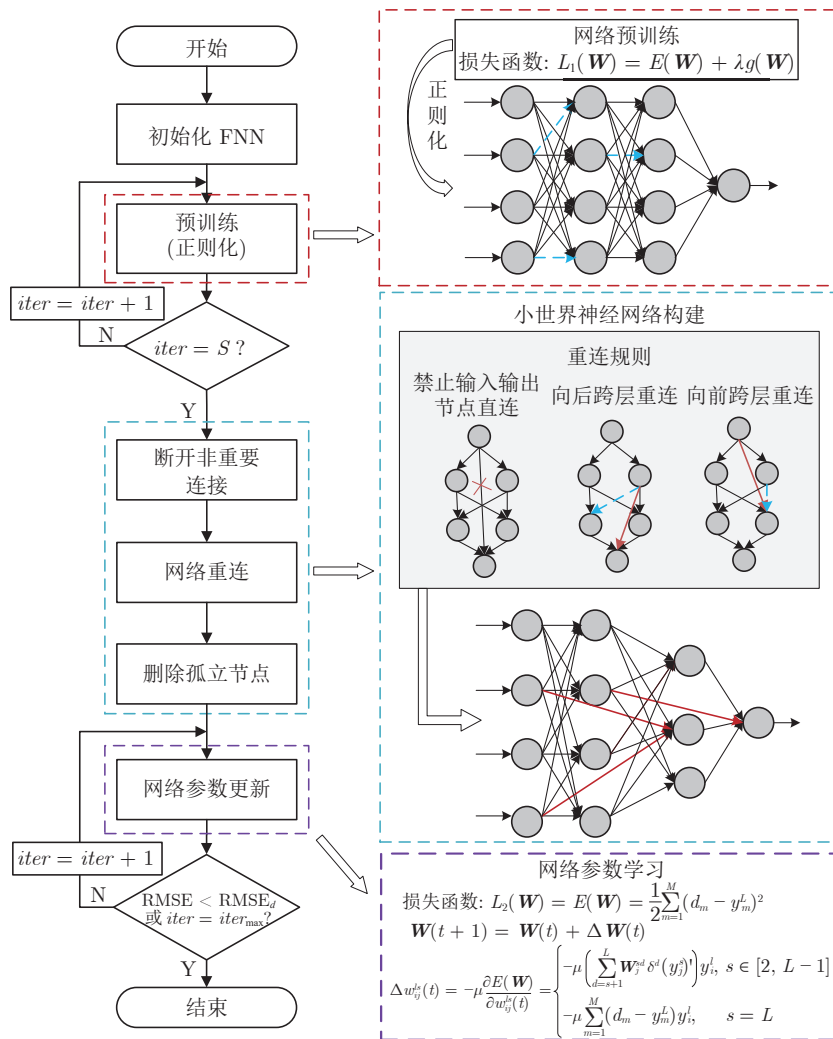


图 4 FSWNN-SC 算法流程图  
Fig.4 The flowchart of FSWNN-SC

模精度及训练时间等方面评价 FSWNN-SC 模型的性能及有效性; 最后, 使用 Wilcoxon 符号秩检验方法进一步验证 FSWNN-SC 模型相比于其他模型的显著性优势.

### 3.1 实验数据

本文选取了 4 个 UCI 数据集<sup>[45]</sup>, 包括 2 个分类数据集(数据集 1 和数据集 2) 和 2 个回归数据集(数据集 3 和数据集 4), 以及 2 个真实数据集(数据集 5 和数据集 6) 进行实验. 数据集信息如下:

1) 数据集 1: HTRU2. 该分类数据集有 17898 组样本, 8 个特征变量, 1 个输出变量(取值  $-1, 1$  分别代表所属两种类别). 抽取 3278 组样本, 正负样本比例为  $1 : 1.2$ , 2295 组样本用于网络训练, 983 组样本用于网络性能测试.

2) 数据集 2: Banknote Authentication. 该分类数据集有 1372 组样本, 4 个特征变量, 1 个输出变量(取值  $-1, 1$  分别代表所属两种类别), 正负样本比例为  $1 : 1.2492$ . 将数据集划分为训练集和测试集, 960 组样本用于网络训练, 412 组样本用于网络性能测试.

3) 数据集 3: Boston Housing. 该回归数据集有 506 组样本, 13 个特征变量, 1 个输出变量. 将数据集划分为训练集和测试集, 354 组样本用于网络训练, 152 组样本用于网络性能测试.

4) 数据集 4: Concrete Compressive Strength. 该数据集有 1030 组样本, 8 个特征变量, 1 个输出变量. 721 组样本用于网络训练, 309 组样本用于网络性能测试.

5) 数据集 5: 风速预测. 风速时间序列从中国某风电场采集, 采样间隔为 2011 年 4 月 6 日凌晨 0:00 至 2011 年 4 月 12 日凌晨 0:00, 采样间隔为 10 分钟. 用  $y(t)$  表示  $t$  时刻的风速值, 使用  $y(t-5)$  至  $y(t)$  作为网络输入, 预测下一时刻的风速值, 即  $y(t+1)$  网络输出. 697 组样本用于网络训练, 299 组样本用于网络性能测试.

6) 数据集 6: 出水生化需氧量 (Biochemical

oxygen demand, BOD) 浓度预测. 该数据来自北京市某污水处理厂, 共包含 365 组数据, 输入变量包括: a) 出水总氮; b) 出水氨氮; c) 进水总氮; d) 进水 BOD; e) 进水氨氮; f) 出水磷酸盐; g) 混合液悬浮固体 (Mixed liquid suspended solids, MLSS); h) 溶解氧 (Dissolved oxygen, DO); i) 进水磷酸盐; j) 进水化学需氧量 (Chemical oxygen demand, COD). 输出变量为出水 BOD 浓度. 265 组样本用于网络训练, 100 组样本用于网络性能测试.

为了消除特征向量之间量级不同导致的影响, 对 6 个实验的数据集都进行归一化处理, 特征变量归一化至  $[-1, 1]$ , 输出变量归一化至  $[0, 1]$ .

### 3.2 实验设置

本文选取基于  $L_{1/2}$  范数(即  $q = 1/2$ ) 的正则化方法进行网络预训练, 进而构建 FSWNN-SC 网络. 实验设置超参数, 包括网络结构、预训练惩罚系数  $\lambda$ 、学习率  $\mu$ 、最大迭代步数  $iter_{max}$  和期望均方根误差  $RMSE_d$  等, 详见表 1.

本文使用分类精度 (Accuracy, Acc) 和标准均方根误差 (Normalized root mean squared error, NRMSE) 分别评价模型对分类和回归实验的建模精度, 具体定义为

$$Acc = \frac{1}{M} \sum_{k=1}^K TP_k \quad (25)$$

$$NRMSE = \sqrt{\frac{\sum_{m=1}^M (d_m - y_m)^2}{\sum_{m=1}^M (d_m - \bar{d})^2}} \quad (26)$$

其中,  $K$  是分类问题中的类别总数,  $TP_k$  是第  $k$  个类别下真阳性样本总量;  $M$  是样本总量,  $d_m$  和  $y_m$  分别为第  $m$  个样本的期望输出与实际输出,  $\bar{d}$  是网络期望输出的均值. 可见,  $Acc$  反映了网络分类准确性,  $Acc$  越大, 网络性能越好;  $NRMSE$  反映了网络实际输出与期望输出之间的误差大小,  $NRMSE$  越

表 1 实验超参数设置  
Table 1 Setting of the hyperparameters in experiments

数据集	网络结构	$\lambda$	$\mu$	$iter_{max}$	$RMSE_d$
数据集 1	8-15-15-1	$1.0 \times 10^{-3}$	0.0003	6000	0.001
数据集 2	4-15-15-1	$1.0 \times 10^{-3}$	0.0008	6000	0.001
数据集 3	13-20-20-1	$1.0 \times 10^{-6}$	0.0008	10000	0.001
数据集 4	8-20-20-1	$1.0 \times 10^{-6}$	0.0008	10000	0.001
数据集 5	6-20-20-1	$1.0 \times 10^{-6}$	0.0005	10000	0.001
数据集 6	10-20-20-1	$1.0 \times 10^{-6}$	0.0008	10000	0.001

小, 网络性能越好.

此外, 由于预训练后可能会产生孤立节点, 本文提出的 FSWNN-SC 在删除孤立节点后, 会产生一定的稀疏化作用. 为了评价模型的稀疏化程度, 定义稀疏度指标为

$$SP = \frac{N_F}{N_I} \quad (27)$$

其中,  $N_I$  是网络初始状态下非零权值的数量,  $N_F$  是训练结束后网络中非零权值的数量.  $SP$  越小表示稀疏化程度越高; 反之,  $SP$  越大, 稀疏化程度越低.

本文实验所使用的软件为 MATLAB R2021b, 操作系统为 Windows11, CPU 为 AMD 5800H 16 GB.

### 3.3 预训练参数影响分析

#### 3.3.1 重连概率 $P$ 对小世界属性的影响

如第 2.1 节所述, 在预训练过程中, 重连概率  $P$  决定了预训练后需要断开并进行重连的连接边数  $N_c$ , 直接影响网络的小世界属性. 本文通过计算不同  $P$  取值情况下的小世界网络属性  $\eta$  值, 来研究重连概率  $P$  对网络小世界属性的影响.

在本研究中, 重连概率  $P$  定义为重连边数  $N_c$  与最大可重连边数  $N_{\max}$  的比值, 即

$$P = \frac{N_c}{N_{\max}} \quad (28)$$

其中, 最大重连边数  $N_{\max}$  取值为可断开边数  $N_d$  和可跨层重连边数  $N_r$  的最小值, 即

$$N_{\max} = \min\{N_d, N_r\} \quad (29)$$

其中,

$$N_d = \sum_{l=1}^{L-2} n_l \times n_{l+1} - \sum_{l=1}^{L-2} \max\{n_l, n_{l+1}\} \quad (30)$$

$$N_r = \sum_{l=1}^{L-2} \sum_{i=2}^{L-l} n_l \times n_{l+i} \quad (31)$$

其中,  $n_l$  代表第  $l$  层的神经元个数. 当重连概率  $P = 1$  时, 规则前馈神经网络变为随机网络, 神经网络的小世界属性根据式 (6) 计算得出.

本文在  $[0, 1]$  范围内, 对  $P$  进行不同取值, 计算相应重连概率下的小世界属性  $\eta$  值, 在各数据集上进行验证, 画出  $P$ - $\eta$  曲线, 如图 5 所示. 由图 5 可见, 在各实验中, 随着重连概率  $P$  的增大,  $\eta$  值均呈现先增大后减小的趋势, 一般在  $P = 0.5$  左右,  $\eta$  达到峰值. 基于以上实验结果, 为了保障所建立模型具有较好的小世界属性, 本文中设置重连概率  $P = 0.5$ .

#### 3.3.2 预训练次数 $S$ 对模型精度的影响

本文在  $[1, 5000]$  范围内对预训练次数  $S$  进行取值, 计算不同预训练次数下建模精度指标, 以研究预训练次数对网络性能的影响, 绘制曲线如图 6 所示.

在分类问题 (数据集 1 和数据集 2) 中, 随着预训练次数的增加, 测试集的 Acc 逐渐减小; 在回归问题 (数据集 3~6) 中, 随着预训练次数的增加, 测试集的 NRMSE 逐渐增大. 可见, 预训练次数的增加将会导致网络性能变差. 因此, 为了保证网络性能, 本文设置预训练次数  $S$  为较小值, 取值  $S = 20$ .

### 3.4 FSWNN-SC 网络性能测试

本文使用 6 个数据集, 分别从训练过程和测试结果对提出的 FSWNN-SC 网络的性能进行评价. FSWNN-SC 网络训练过程 RMSE 曲线如图 7 所示. 可见, 在训练过程中, 当迭代步数达到设定的预训练次数时, 网络按照重连规则结构发生变化以构造小世界神经网络, 训练 RMSE 发生一次跳变, 之

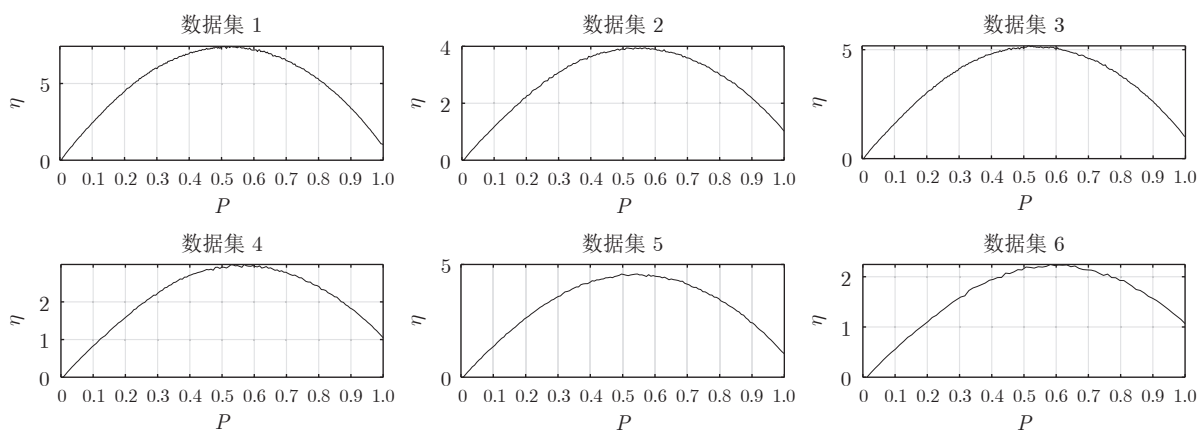


图 5 网络小世界属性  $\eta$  与重连概率  $P$  的关系曲线 ( $P$ - $\eta$  曲线)

Fig.5 The curves for the relationship between the small-world property  $\eta$  and the rewiring probability  $P$  ( $P$ - $\eta$  curves)



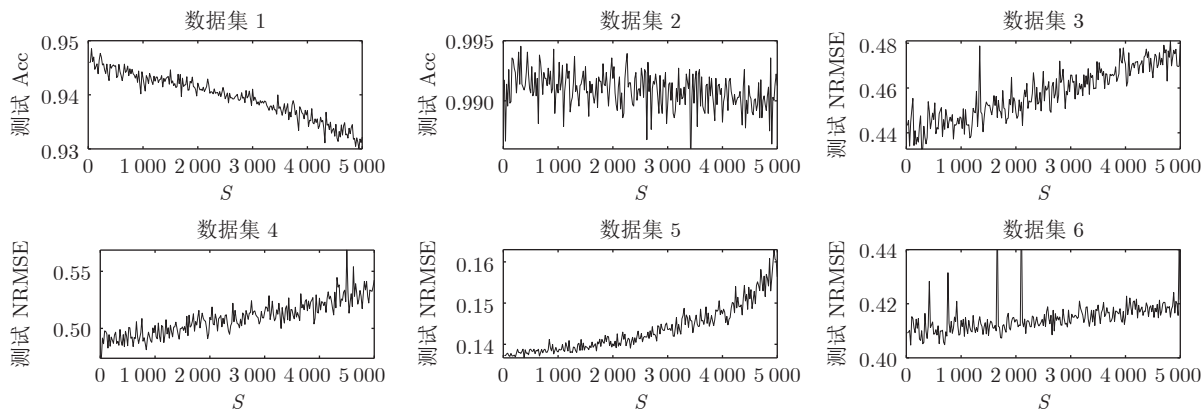


图 6 预训练次数对网络性能的影响

Fig.6 Influence of pre-training epochs on network performance

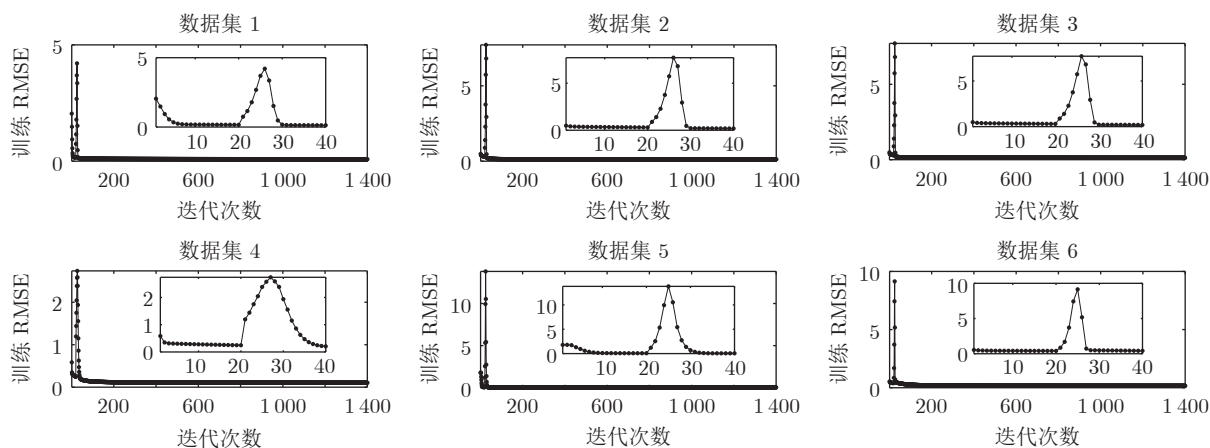


图 7 训练过程 RMSE 曲线

Fig.7 The RMSE curves in the training process

后快速下降直至网络收敛。

FSWNN-SC 网络对测试样本的分类和预测效果如图 8 所示。可见, 对于分类问题, FSWNN-SC 网络能够达到较高的分类精度; 对于回归问题, FSWNN-SC 网络能够较好地输出进行拟合, 实验结果验证了 FSWNN-SC 网络的有效性。

为了进一步研究 FSWNN-SC 网络的性能, 本文将其与多层前馈神经网络 (FNN)、基于 WS 重连规则构造的前馈小世界神经网络 (FSWNN-WS)<sup>[32]</sup>、基于层连优化的新型小世界神经网络 (FSWNN-TO)<sup>[33]</sup>、基于 Katz 中心性剪枝的前馈小世界神经网络 (PFSWNN-Katz)<sup>[30]</sup>、基于正则化 ( $L_{1/2}$  范数) 的删减型小世界神经网络 (PFSWNN-SL) 等模型性能进行对比。所有网络设置相同的初始结构, 小世界神经网络的构造过程重连概率均设置为  $P = 0.5$ , 网络学习率、期望 RMSE 和最大迭代次数均按照表 1 进行设置。实验独立运行 20 次, 通过计算各模型的测试 Acc 和测试 NRMSE 等性能指标评价模

型精度, 记录训练时间对模型的学习速度进行评价, 同时计算网络稀疏度, 实验结果见表 2 和表 3。

从表中可以看出, 相比于其他网络, FSWNN-SC 网络在分类问题中具有更高的分类精度, 在回归问题中具有更小的测试 NRMSE, 说明 FSWNN-SC 能够获得更好的建模精度。FSWNN-SC、PFSWNN-SL 和 PFSWNN-Katz 模型具有网络稀疏性, 其中 FSWNN-SC 模型通过在网络重连过程中删除产生的孤立节点实现稀疏化, 而 PFSWNN-SL 和 PFSWNN-Katz 是在学习过程中分别通过网络正则化和合并节点实现网络稀疏化, 实验验证网络稀疏化能够获得相对于 FSWNN-WS 和 FNN 更紧凑的结构, 然而其训练时间也相应增加。虽然 FSWNN-SC 的稀疏化程度要低于 PFSWNN-SL 和 PFSWNN-Katz, 但是其训练时间相对于 PFSWNN-SL 和 PFSWNN-Katz 缩短, 更接近于 FSWNN-WS 和 FNN 的训练时间。相比于 FSWNN-TO, FSWNN-SC 的训练时间明显短, 因为 FSWNN-SC 没有

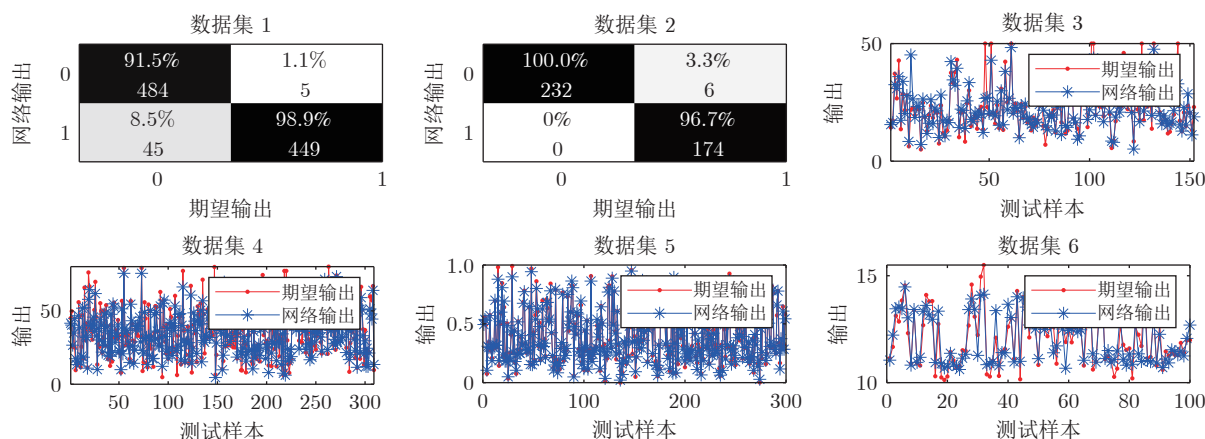


图 8 测试集样本拟合与分类效果

Fig. 8 Test set sample fitting and classification effects

表 2 分类实验结果对比

Table 2 Comparison results in classification experiments

分类实验	网络	网络结构	稀疏度 $SP$	测试 Acc		训练时间 (s)	
				均值	标准差	均值	标准差
数据集 1	FSWNN-SC	8-15-12-1	0.8861	<b>0.9472</b>	0.0034	3.9631	0.1936
	PFSWNN-SL	8-15-11-1	0.7511	0.9403	0.0026	5.6645	0.2085
	PFSWNN-Katz	8-12-10-1	<b>0.6056</b>	0.9396	0.0126	4.0555	0.2764
	FSWNN-TO	8-15-15-1	—	0.9392	0.0066	5.4922	0.0147
	FSWNN-WS	8-15-15-1	—	0.9374	0.0073	3.9371	0.1255
	FNN	8-15-15-1	—	0.9195	0.0093	<b>3.7201</b>	0.0609
数据集 2	FSWNN-SC	4-15-12-1	0.8950	<b>0.9883</b>	0.0049	2.7552	0.4252
	PFSWNN-SL	4-15-10-1	0.6608	0.9788	0.0081	4.6556	0.2525
	PFSWNN-Katz	4-10-11-1	<b>0.5463</b>	0.9823	0.0054	2.8007	0.1837
	FSWNN-TO	4-15-15-1	—	0.9840	0.0040	3.6596	0.0614
	FSWNN-WS	4-15-15-1	—	0.9782	0.0071	2.3605	0.0419
	FNN	4-15-15-1	—	0.9756	0.0132	<b>2.3402</b>	0.0347

引入同层节点连接, 在权值更新时 FSWNN-SC 的误差反传“路径长度”比 FSWNN-TO 短. 同时, 通过对比 FNN 与其他 5 个具有小世界属性的网络, 验证了将小世界属性加入神经网络模型的设计中, 能够提高模型精度. 综上所述, 通过对比实验分析可见, 本文提出的 FSWNN-SC 网络在获得紧凑网络结构的同时, 具有更高的建模精度. 分析其原因, FSWNN-SC 网络优越的建模精度依赖于其预训练环节, 该环节模仿生物神经网络的突触巩固过程, 通过网络正则化方法保留了重要的权值连接并断开了不重要的权值连接, 使该阶段网络所学到的知识得以保留, 可以有效避免灾难性遗忘问题.

### 3.5 Wilcoxon 符号秩检验

为了进一步验证本文所提出模型是否具有显著

优势, 对第 3.4 节的实验结果进行 Wilcoxon 符号秩检验. 它是一种非参数检验方法, 可以评估多个事件下两个模型是否有显著性差异<sup>[46]</sup>, 具体流程如下.

**步骤 1.** 提出原假设  $H_0$ : 两个模型之间没有显著性差异.

**步骤 2.** 对于多个事件, 通过式 (32) 和式 (33) 分别计算正负秩

$$R^+ = \sum_{d_i > 0} \text{rank}(d_i) + \frac{1}{2} \sum_{d_i = 0} \text{rank}(d_i) \quad (32)$$

$$R^- = \sum_{d_i < 0} \text{rank}(d_i) + \frac{1}{2} \sum_{d_i = 0} \text{rank}(d_i) \quad (33)$$

其中,  $d_i$  是第  $i$  个事件下两个模型的差异, 本文通过精度性能指标的差值计算得到,  $\text{rank}(\cdot)$  表示降序排名的名次. 基于此得到 Wilcoxon 统计量, 为正负

表 3 回归实验结果对比  
Table 3 Comparison results in regression experiments

回归实验	网络	网络结构	稀疏度 $SP$	测试 NRMSE		训练时间 (s)	
				均值	标准差	均值	标准差
数据集 3	FSWNN-SC	13-20-13-1	0.7941	<b>0.4331</b>	0.0199	2.9838	0.0978
	PFSWNN-SL	13-20-14-1	<b>0.7265</b>	0.4546	0.0187	6.9352	0.2077
	PFSWNN-Katz	13-15-16-1	0.7563	0.4551	0.0200	4.6810	0.1358
	FSWNN-TO	13-20-20-1	—	0.4476	0.0193	4.3250	0.0267
	FSWNN-WS	13-20-20-1	—	0.4582	0.0232	<b>2.9583</b>	0.0609
	FNN	13-20-20-1	—	0.5728	0.0235	3.1481	0.1228
数据集 4	FSWNN-SC	8-20-16-1	0.8865	<b>0.4814</b>	0.0308	4.7431	0.1883
	PFSWNN-SL	8-20-17-1	<b>0.7706</b>	0.5104	0.0275	8.4518	0.3075
	PFSWNN-Katz	8-17-18-1	0.8064	0.5159	0.0234	5.6207	0.5053
	FSWNN-TO	8-20-20-1	—	0.4944	0.0147	5.8352	0.0231
	FSWNN-WS	8-20-20-1	—	0.5142	0.0222	4.6306	0.1288
	FNN	8-20-20-1	—	0.6691	0.0058	<b>4.4024</b>	0.0585
数据集 5	FSWNN-SC	6-20-14-1	0.7952	<b>0.1351</b>	0.0017	5.0063	0.2048
	PFSWNN-SL	6-20-14-1	0.6698	0.1405	0.0080	8.3014	0.3069
	PFSWNN-Katz	6-17-14-1	<b>0.6647</b>	0.1371	0.0031	5.2003	0.4510
	FSWNN-TO	6-20-20-1	—	0.1374	0.0032	5.5165	0.1494
	FSWNN-WS	6-20-20-1	—	0.1378	0.0026	<b>4.8520</b>	0.2943
	FNN	6-20-20-1	—	0.1544	0.0084	5.0213	0.4910
数据集 6	FSWNN-SC	10-20-16-1	0.8663	<b>0.4055</b>	0.0101	2.7706	0.1334
	PFSWNN-SL	10-20-15-1	<b>0.7298</b>	0.4168	0.0112	6.2909	0.0112
	PFSWNN-Katz	10-15-18-1	0.7649	0.4139	0.0093	3.5227	0.4455
	FSWNN-TO	10-20-20-1	—	0.4124	0.0143	3.2057	0.0388
	FSWNN-WS	10-20-20-1	—	0.4144	0.0102	2.7778	0.0161
	FNN	10-20-20-1	—	0.4309	0.0134	<b>2.7206</b>	0.0132

秩两者中的较小值, 即

$$T = \min(R^+, R^-) \quad (34)$$

步骤 3. 通过 Wilcoxon 统计量计算得到  $Z$  的临界值, 计算式为

$$Z = \frac{T - \frac{1}{4}N(N+1)}{\sqrt{\frac{1}{24}N(N+1)(2N+1)}} \quad (35)$$

其中,  $N$  是事件数, 通过查正态分布  $Z$  值表得到对应的  $P_w$ , 若  $P_w > 0.05$ , 则接受原假设  $H_0$ , 反之, 则拒绝原假设, 即两个模型间存在显著性差异, 且  $P_w$  值越小差异越显著.

针对第 3.4 节中的实验结果进行 Wilcoxon 符号秩检验, 其中事件数为实验独立运行的次数 20, 分别对比 FSWNN-SC 模型与其他 5 种模型的性能, 检验结果如表 4 所示 (\*表示存在显著性差异). 由表 4 可见, 本文提出的 FSWNN-SC 模型在精度上显著优于其他 5 种模型. 此外, 本文实验中, FSWNN-

SC 在不同场景、不同任务 (分类、回归) 下表现出的显著优越性, 反映了该模型对不同应用环境的适应性及可塑性, 同时在实际应用中的准确建模反映了其具有一定的抗扰性.

## 4 结束语

针对 WS 构造小世界神经网络方式中随机断开规则网络中的权值连接, 可能会造成重要信息丢失, 进而导致网络精度下降的问题, 本文提出了 FSWNN-SC 模型. 该模型具有以下特点:

- 1) FSWNN-SC 模型使用正则化方法对 FNN 进行预训练, 断开对网络不重要的权值连接, 以减少重要信息丢失的概率, 并制定了重连规则构造小世界神经网络, 保证构造的小世界神经网络的性能;
- 2) FSWNN-SC 模型通过在网络重连过程中删除产生的孤立节点实现网络稀疏化, 能够获得紧凑的网络结构;

表 4 Wilcoxon 符号秩检验结果  
Table 4 Results of Wilcoxon signed-rank test

实验	模型	$R^+$	$R^-$	$Z$	$P_w$
数据集 1	FSWNN-SC vs. PFSWNN-SL	206	4	-3.7706	0.0002*
	FSWNN-SC vs. PFSWNN-Katz	179	31	-2.7626	0.0058*
	FSWNN-SC vs. FSWNN-TO	203	7	-3.6586	0.0002*
	FSWNN-SC vs. FSWNN-WS	198.5	11.5	-3.4906	0.0004*
	FSWNN-SC vs. FNN	210	0	-3.9199	0*
数据集 2	FSWNN-SC vs. PFSWNN-SL	203.5	6.5	-3.6773	0.0002*
	FSWNN-SC vs. PFSWNN-Katz	177	33	-2.6880	0.0074*
	FSWNN-SC vs. FSWNN-TO	176.5	33.5	-2.6693	0.0076*
	FSWNN-SC vs. FSWNN-WS	199.5	10.5	-3.5279	0.0004*
	FSWNN-SC vs. FNN	206.5	3.5	-3.7893	0.0004*
数据集 3	FSWNN-SC vs. PFSWNN-SL	187	23	-3.0613	0.0022*
	FSWNN-SC vs. PFSWNN-Katz	207	3	-3.8079	0.0002*
	FSWNN-SC vs. FSWNN-TO	190	20	-3.1733	0.0016*
	FSWNN-SC vs. FSWNN-WS	209	1	-3.8826	0.0002*
	FSWNN-SC vs. FNN	210	0	-3.9199	0*
数据集 4	FSWNN-SC vs. PFSWNN-SL	184	26	-2.9493	0.0032*
	FSWNN-SC vs. PFSWNN-Katz	210	0	-3.9199	0.0000*
	FSWNN-SC vs. FSWNN-TO	159	51	-2.0160	0.0434*
	FSWNN-SC vs. FSWNN-WS	208	2	-3.8453	0.0002*
	FSWNN-SC vs. FNN	210	0	-3.9199	0*
数据集 5	FSWNN-SC vs. PFSWNN-SL	187	23	-3.0613	0.0022*
	FSWNN-SC vs. PFSWNN-Katz	169	41	-2.3893	0.0168*
	FSWNN-SC vs. FSWNN-TO	177	33	-2.6880	0.0074*
	FSWNN-SC vs. FSWNN-WS	190	20	-3.1733	0.0016*
	FSWNN-SC vs. FNN	210	0	-3.9199	0*
数据集 6	FSWNN-SC vs. PFSWNN-SL	171	39	-2.4640	0.0138*
	FSWNN-SC vs. PFSWNN-Katz	160	50	-2.0533	0.0434*
	FSWNN-SC vs. FSWNN-TO	177	33	-2.6880	0.0074*
	FSWNN-SC vs. FSWNN-WS	172	38	-2.5013	0.0124*
	FSWNN-SC vs. FNN	210	0	-3.9199	0*

3) 实验结果表明, 相比于 PFSWNN-SL、PFSWNN-Katz、FSWNN-TO、FSWNN-WS 以及 FNN, FSWNN-SC 在获得紧凑结构的同时, 在模型精度方面具有显著优势。

### References

- Tran V P, Santoso F, Garrat M A, Anavatti S G. Neural network-based self-learning of an adaptive strictly negative imaginary tracking controller for a quadrotor transporting a cable-suspended payload with minimum swing. *IEEE Transactions on Industrial Electronics*, 2021, **68**(10): 10258-10268
- Zhang G H, Li B, Wu J X, Wang R, Lan Y Z, Sun L, et al. A low-cost and high-speed hardware implementation of spiking neural network. *Neurocomputing*, 2020, **382**: 106-115
- Lv H, Wen M, Lu R A, Li J. An adversarial attack based on incremental learning techniques for unmanned in 6G scenes. *IEEE Transactions on Vehicular Technology*, 2021, **70**(6): 5254-5264
- Li W J, Li M, Zhang J K, Qiao J F. Design of a self-organizing reciprocal modular neural network for nonlinear system modeling. *Neurocomputing*, 2020, **411**: 327-339
- Qiao Jun-Fei, Ding Hai-Xu, Li Wen-Jing. Structure design for recurrent fuzzy neural network based on wavelet transform fuzzy markov chain. *Acta Automatica Sinica*, 2020, **46**(11): 2367-2378 (乔俊飞, 丁海旭, 李文静. 基于 WTFMC 算法的递归模糊神经网络结构设计. *自动化学报*, 2020, **46**(11): 2367-2378)
- Feng Yong, Chen Yi-Gang, Qiang Bao-Hua. Social and comment text CNN model based automobile recommendation. *Acta Automatica Sinica*, 2019, **45**(3): 518-529 (冯永, 陈以刚, 强保华. 融合社交因素和评论文本卷积网络模型的汽车推荐研究. *自动化学报*, 2019, **45**(3): 518-529)
- Wang S, Cao J, Yu P S. Deep learning for spatio-temporal data mining: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 2022, **34**(8): 3681-3700
- Chen Qing-Jiang, Zhang Xue. Single image dehazing based on multiple convolutional neural networks. *Acta Automatica Sinica*, 2021, **47**(7): 1739-1748

- (陈清江, 张雪. 基于并联卷积神经网络的图像去雾. *自动化学报*, 2021, **47**(7): 1739–1748)
- 9 Jiao Y, Yao H, Xu C. SAN: Selective alignment network for cross-domain pedestrian detection. *IEEE Transactions on Image Processing*, 2021, **30**: 2155–2167
  - 10 Otter D W, Medina J R, Kalita J K. A survey of the use of deep learning for natural language processing. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, **32**(2): 604–624
  - 11 Xi Xue-Feng, Zhou Guo-Dong. A survey on deep learning for natural language processing. *Acta Automatica Sinica*, 2016, **42**(10): 1445–1465  
(奚雪峰, 周国栋. 面向自然语言处理的深度学习研究. *自动化学报*, 2016, **42**(10): 1445–1465)
  - 12 Watts D J, Strogatz S H. Collective dynamics of small world networks. *Nature*, 1998, **393**(4): 440–442
  - 13 Bassett D S, Bullmore E. Small-world brain networks. *Neuroscientist*, 2006, **12**(6): 512–523
  - 14 Strogatz S H. Exploring complex networks. *Nature*, 2001, **410**: 268–276
  - 15 Pessoa L. Understanding brain networks and brain organization. *Physics of Life Reviews*, 2014, **11**(3): 400–435
  - 16 Latora V, Marchiori M. Efficient behavior of small-world networks. *Physical Review Letters*, 2001, **87**(19): Article No. 198701
  - 17 Li H, Zhang L. A bilevel learning model and algorithm for self-organizing feed-forward neural networks for pattern classification. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, **32**(11): 4901–4915
  - 18 Guliyev N J, Ismailov V E. On the approximation by single hidden layer feedforward neural networks with fixed weights. *Neural Networks*, 2017, **98**: 296–304
  - 19 Huang G B, Chen L, Siew C K. Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Transactions on Neural Networks*, 2006, **17**(4): 879–892
  - 20 Qiao J F, Li F, Yang C L, Li W J, Gu K. A self-organizing RBF neural network based on distance concentration immune algorithm. *IEEE/CAA Journal of Automatica Sinica*, 2022, **7**(1): 276–291
  - 21 Yu Q, Song S, Ma C, Wei J, Chen S, Tan K C. Temporal encoding and multispike learning framework for efficient recognition of visual patterns. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, **33**(8): 3387–3399
  - 22 Simard D, Nadeau L, Kroger H. Fastest learning in small-world neural networks. *Physics Letters A*, 2004, **336**(1): 8–15
  - 23 Li X H, Li X L, Zhang J H, Zhang Y L, Li M L. A new multilayer feedforward small-world neural network with its performances on function approximation. In: Proceedings of the IEEE International Conference on Computer Science and Automation Engineering (CSAE). Shanghai, China: IEEE, 2011. 353–357
  - 24 Li X, Xu F, Zhang J, Wang S. A multilayer feed forward small-world neural network controller and its application on electrohydraulic actuation system. *Journal of Applied Mathematics*, 2013, **21**: 1–8
  - 25 Dong Z K, Duan S K, Hu X F, Li H. A novel memristive multilayer feedforward small-world neural network with its applications in PID control. *The Scientific World Journal*, 2014, **14**: 1–12
  - 26 Wang S X, Zhao X, Wang H, Li M. Small-world neural network and its performance for wind power forecasting. *CSEE Journal of Power and Energy Systems*, 2020, **6**(2): 362–373
  - 27 Erkamaz O, Ozer M. Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes. *Chaos, Solitons and Fractals*, 2016, **83**: 178–185
  - 28 Erkamaz O, Ozer M, Perc M. Performance of small-world feedforward neural networks for the diagnosis of diabetes. *Applied Mathematics and Computation*, 2017, **311**: 22–28
  - 29 Zhang R C, Hu X L. Effluent quality prediction of wastewater treatment system based on small-world ANN. *Journal of Computers*, 2012, **7**(9): 2136–2143
  - 30 Li W J, Chu M H, Qiao J F. A pruning feedforward small-world neural network based on Katz centrality for nonlinear system modeling. *Neural Networks*, 2020, **130**: 269–285
  - 31 Newman M E J, Watts D J. Renormalization group analysis of the small-world network model. *Physics Letters A*, 1999, **263**(4): 341–346
  - 32 Li Xiao-Hu, Du Hai-Feng, Zhang Jin-Hua, Wang Sun-An. Multilayer feedforward small-world neural networks and its function approximation. *Control Theory and Applications*, 2010, **27**(7): 836–842  
(李小虎, 杜海峰, 张进华, 王孙安. 多层前向小世界神经网络及其函数逼近. *控制理论与应用*, 2010, **27**(7): 836–842)
  - 33 Wang Shuang-Xin, Yang Cheng-Hui. Novel small-world neural network based on topology optimization. *Control and Decision*, 2014, **29**(1): 77–82  
(王爽心, 杨成慧. 基于层连优化的新型小世界神经网络. *控制与决策*, 2014, **29**(1): 77–82)
  - 34 Guo D, Yang L. Research on trim of multilayer feedforward small world network based on E-exponential information entropy. In: Proceedings of the 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC). Hangzhou, China: IEEE, 2017. 155–158
  - 35 Grutzendler J, Kasthuri N, Gan W B. Long-term dendritic spine stability in the adult cortex. *Nature*, 2002, **420**: 812–816
  - 36 Zuo Y, Lin A, Chang P, Gan W B. Development of long-term dendritic spine stability in diverse regions of cerebral cortex. *Neuron*, 2005, **46**: 181–189
  - 37 Demiar J, Schuurmans D. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 2006, **7**(1): 1–30
  - 38 Humphries M D, Gurney K. Network “Small-world-ness”: A quantitative method for determining canonical network equivalence. *Plos One*, 2008, **3**(4): Article No. e0002051
  - 39 Ziegler L, Zenke F, Kastner D B, Gerstner W. Synaptic consolidation: From synapses to behavioral modeling. *Journal of Neuroscience*, 2015, **35**(3): 1319–1334
  - 40 Bliss T V P, Lomo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of Physiology*, 1973, **232**(2): 331–356
  - 41 Dudek S M, Bear M F. Homosynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. *Proceedings of the National Academy of Sciences*, 1992, **89**(10): 4363–4367
  - 42 Rathi N, Panda P, Roy K. STDP-based pruning of connections and weight quantization in spiking neural networks for energy-efficient recognition. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2019, **38**(4): 668–677
  - 43 Peng J, Tang B, Jiang H, Li Z, Lin T, Li H F. Overcoming long-term catastrophic forgetting through adversarial neural pruning and synaptic consolidation. *IEEE Transactions on Neural Networks and Learning Systems*, 2022, **33**(9): 4243–4256



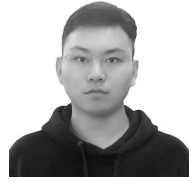
- 44 Wang J, Xu C, Yang X, Zurada J M. A novel pruning algorithm for smoothing feedforward neural networks based on group lasso method. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(5): 2012–2024
- 45 Bache K, Lichman M. UCI machine learning repository [Online], available: <https://archive.ics.uci.edu/ml>, December 20, 2021
- 46 Papantoni-Kazakos P. Small-sample efficiencies of rank tests. *IEEE Transactions on Information Theory*, 1975, **21**(2): 150–157



**李文静** 北京工业大学信息学部副教授. 主要研究方向为神经网络计算, 污水处理过程智能建模. 本文通信作者. E-mail: wenjing.li@bjut.edu.cn

**(LI Wen-Jing** Associate professor at the Faculty of Information Technology, Beijing University of Technology.

Her research interest covers neural network computation and intelligent modelling in wastewater treatment process. Corresponding author of this paper.)



**李治港** 北京工业大学信息学部硕士研究生. 主要研究方向为神经网络结构设计与优化, 污水处理过程特征建模.

E-mail: lzg551602@emails.bjut.edu.cn

**(LI Zhi-Gang** Master student at

the Faculty of Information Technology, Beijing University of Technology. His research interest covers structure design and optimization of neural networks, and feature modelling in wastewater treatment process.)



**乔俊飞** 北京工业大学信息学部教授. 主要研究方向为污水处理过程智能控制, 神经网络结构设计与优化.

E-mail: adqiao@bjut.edu.cn

**(QIAO Jun-Fei** Professor at the Faculty of Information Technology, Beijing University of Technology.

His research interest covers intelligent control of wastewater treatment process and structure design and optimization of neural networks.)