

随机配置网络研究进展*

张成龙¹, 丁世飞^{1,2}, 郭丽丽^{1,2}, 张健^{1,2}

¹(中国矿业大学 计算机科学与技术学院, 江苏 徐州 221116)

²(矿山数字化教育部工程研究中心 (中国矿业大学), 江苏 徐州 221116)

通信作者: 丁世飞, E-mail: dingsf@cumt.edu.cn



摘 要: 随机配置网络 (stochastic configuration network, SCN) 是一种新兴的增量式神经网络模型, 与其他随机化神经网络方法不同, 它能够通过监督机制进行隐含层节点参数配置, 保证了模型的快速收敛性能. 因其具有学习效率高、人为干预程度低和泛化能力强等优点, 自 2017 年提出以来, SCN 吸引了大量国内外学者的研究兴趣, 得到了快速地推广和发展. 从 SCN 的基础理论、典型算法变体、应用领域以及未来研究方向等方面切入, 全面地概述 SCN 研究进展. 首先, 从理论的角度分析 SCN 的算法原理、通用逼近性能及其优点; 其次, 重点研究深度 SCN、二维 SCN、鲁棒 SCN、集成 SCN、分布式并行 SCN、正则化 SCN 等典型变体; 随后介绍 SCN 在硬件实现、计算机视觉、医学数据分析、故障检测与诊断、系统建模预测等不同领域的应用进展; 最后指出 SCN 在卷积神经网络架构、半监督学习、无监督学习、多视图学习、模糊神经网络、循环神经网络等研究方向的发展潜力.

关键词: 随机配置网络; 神经网络; 深度学习; 随机化学习; 研究进展

中图法分类号: TP18

中文引用格式: 张成龙, 丁世飞, 郭丽丽, 张健. 随机配置网络研究进展. 软件学报. <http://www.jos.org.cn/1000-9825/6804.htm>

英文引用格式: Zhang CL, Ding SF, Guo LL, Zhang J. Research Progress on Stochastic Configuration Network. Ruan Jian Xue Bao/Journal of Software (in Chinese). <http://www.jos.org.cn/1000-9825/6804.htm>

Research Progress on Stochastic Configuration Network

ZHANG Cheng-Long¹, DING Shi-Fei^{1,2}, GUO Li-Li^{1,2}, ZHANG Jian^{1,2}

¹(School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China)

²(Mine Digitization Engineering Research Center of Ministry of Education (China University of Mining and Technology), Xuzhou 221116, China)

Abstract: Stochastic configuration network (SCN), as an emerging incremental neural network model, is different from other randomized neural network methods. It can configure the parameters of hidden layer nodes through supervision mechanisms, thereby ensuring the fast convergence performance of SCN. Due to the advantages of high learning efficiency, low human intervention, and strong generalization ability, SCN has attracted a large number of national and international scholars and developed rapidly since it was proposed in 2017. In this study, SCN research is summarized from the aspects of basic theories, typical algorithm variants, application fields, and future research directions of SCN. Firstly, the algorithm principles, universal approximation capacity, and advantages of SCN are analyzed theoretically. Secondly, typical variants of SCN are studied, such as DeepSCN, 2DSCN, Robust SCN, Ensemble SCN, Distributed SCN, Parallel SCN, and Regularized SCN. Then, the applications of SCN in different fields, including hardware implementation, computer vision, medical data analysis, fault detection and diagnosis, and system modeling and prediction are introduced. Finally, the development potential of SCN in convolutional neural network architectures, semi-supervised learning, unsupervised learning, multi-view learning, fuzzy neural network, and recurrent neural network is pointed out.

Key words: stochastic configuration network (SCN); neural network; deep learning; randomized learning; survey

* 基金项目: 国家自然科学基金 (61976216, 62276265, 61672522)

收稿时间: 2021-12-20; 修改时间: 2022-07-17; 采用时间: 2022-09-16; jos 在线出版时间: 2022-12-30

随着《新一代人工智能发展规划》的实施,人工智能在智能制造、生物计算以及目标检测等研究领域高速发展,如何快速有效地分析处理各行业的大规模复杂数据,已经成为当前人工智能发展亟需关注解决的难题^[1,2].神经网络(neural network, NN)^[3-6]因其具有较强的特征学习能力和非线性逼近特性得到了快速地发展,深度学习模型(deep learning)通过构建多层网络结构,能够从高维大规模数据中挖掘多级抽象的特征,在特征表示、图像分类和模式识别领域取得了突破性的成功^[7,8].但是,深度置信网络(deep belief network, DBN)^[9,10]、深度玻尔兹曼机(deep Boltzmann machine, DBM)^[11]、卷积神经网络(convolutional neural network, CNN)^[12]等模型通常需要采用反向传播(back propagation, BP)算法逐层求解梯度的方式更新权重,网络结构复杂、训练速度慢且存在大量的网络参数.

随机化学习(randomized learning)方法因其建模的快速有效性,在机器学习中展现出了巨大的发展潜力^[13-15].其中,神经网络的随机学习技术起源于20世纪80年代后期,在20世纪90年代早期得到进一步发展,Pao等人提出了一种随机向量函数链接网络(random vector functional link network, RVFL)^[16,17],Schmidt等人提出了具有随机权重的前馈神经网络(feedforward neural network with random weights, FNNRW)^[18],其基本思想和训练过程均为:随机分配隐含层节点的输入权重和偏置,利用最小二乘法或其正则化方法计算隐含层节点的输出权重,简化了模型的学习过程,提高了其学习效率.随后,在隐含层节点输入权重和偏置满足一致性分布条件下,Igelnik等人证明了RVFL对于连续函数的通用逼近性^[19].但是后续研究表明,此类随机化神经网络的通用逼近性取决于隐含层节点数量和随机参数范围,如果参数设置不当,模型无法以很高的概率逼近目标函数^[20-22].为确保随机化神经网络的通用逼近性,提高模型的泛化能力,Wang等人于2017年提出了随机配置网络(stochastic configuration network, SCN),SCN采取增量学习方式,并引入监督机制,能够从一个较小的网络结构开始,根据训练样本数据增量式地生成隐含层节点,构建网络结构^[23].与传统的RVFL等随机化神经网络不同,SCN引入了监督机制,利用不等式约束对随机参数进行赋值,并自适应地选择随机参数的取值范围,通过最小二乘法计算节点输出权重,保证了学习模型的通用逼近性,在较少人为干预下,提高了网络的学习精度和学习效率^[23].

本文首次综述了近年来SCN的主要进展,主要从算法理论、重要算法变体、算法应用和未来研究方向等方面对SCN代表性模型进行了详细的描述.本文第1节介绍了SCN算法理论,说明了其通用逼近性,并分析了其建模优、劣势.第2节重点研究了SCN发展历程及其典型变体;随后归纳了SCN在不同领域的应用进展.最后指出了SCN的潜在未来研究方向.希望本文有助于国内外研究人员全面了解SCN的最新研究进展和应用,为他们未来的研究工作提供一定帮助.

1 随机配置网络

1.1 算法原理

SCN是一种增量式快速建模方法,其网络结构由一个隐含层节点,利用“数据依赖”的监督机制随机配置隐含层节点输入权重和偏置,逐渐增加隐含层节点个数,并使用最小二乘方法计算输出权重,完成网络构建.其中,SCN网络结构与传统的单隐层前馈神经网络(single-hidden layer feedforward neural network, SFLN)相似,如图1所示,包括输入层、隐含层和输出层^[23].

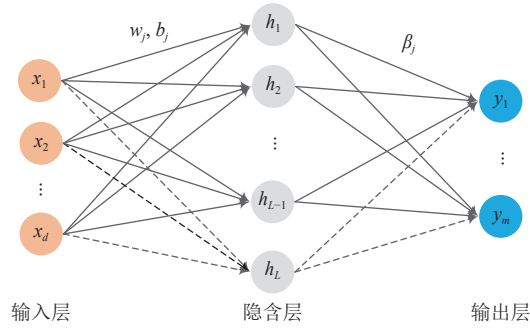
SCN构建过程可以描述如下.

给定一组训练数据集 $\{X, Y\}$, $X = \{x_1, x_2, \dots, x_N\}$ 表示特征数据, $x_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,d}\} \in \mathbb{R}^d$; $Y = \{y_1, y_2, \dots, y_N\}$ 表示相应的标签数据, $y_i = \{y_{i,1}, y_{i,2}, \dots, y_{i,m}\} \in \mathbb{R}^m$; $i = 1, 2, \dots, N$.

Step 1. 给定目标函数 $f: \mathbb{R}^d \rightarrow \mathbb{R}^m$, 假设SCN隐含层的 $L-1$ 个节点已经生成, 则当前网络输出结果可由公式(1)计算得出:

$$f_{L-1}(X) = \sum_{j=1}^{L-1} \beta_j g_j(w_j^T X + b_j), L = 1, 2, \dots, L_{\max}, f_0 = 0 \quad (1)$$

其中, β_j 表示隐含层节点 j 的输出权重; $g(\cdot)$ 表示激活函数; w_j 和 b_j 分别表示隐含层第 j 个节点的输入权重和偏置; $j = 1, 2, \dots, L_{\max}$.

图 1 随机配置网络结构^[23]

Step 2. 计算当前网络残差向量如公式 (2) 所示:

$$e_{L-1} = f - f_{L-1}(X) = [e_{L-1,1}(X), e_{L-1,2}(X), \dots, e_{L-1,m}(X)]^T \in \mathbb{R}^{N \times m} \quad (2)$$

Step 3. 如 $\|e_{L-1}\|^2$ 未达到预设误差 ε 或未达到最大节点数 $L_{\max}$, 增加第 L 个隐含层节点, 并根据公式 (4) 监督机制确定其输入权重和偏置:

$$h_L = [g_L(w_L^T x_1 + b_L), g_L(w_L^T x_2 + b_L), \dots, g_L(w_L^T x_N + b_L)]^T \in \mathbb{R}^N \quad (3)$$

$$\xi_{L,q} = \frac{(e_{L-1,q}^T \cdot h_L)^2}{h_L^T h_L} - (1 - r - \mu_L) \|e_{L-1,q}\|^2 \quad (4)$$

其中, $q = 1, 2, \dots, m$; h_L 表示隐含层节点 L 的输出; w_L 和 b_L 分别表示节点 L 的候选参数; $T_{\max}$ 表示最大候选池大小; $r \in (0, 1)$; $\{\mu_L\}$ 表示非负实数序列, 其中 $\mu_L \leq 1 - r$, $\lim_{L \rightarrow +\infty} \mu_L = 0$; 满足 $\xi_L = \sum_{q=1}^m \xi_{L,q} \geq 0$ 最大值的候选节点参数作为第 L 个节点参数。

Step 4. 根据公式 (5) 确定隐含层节点输出权重:

$$\beta = \arg \min_{\beta} \|H\beta - Y\|^2 = H^+ Y \quad (5)$$

其中, H^+ 表示 H 的摩尔-彭若斯 (Moore-Penrose) 广义逆, $H = [h_1, h_2, \dots, h_L]$ 。

Step 5. 计算网络输出结果 f :

$$f = H\beta \quad (6)$$

1.2 通用逼近性证明

定理 1^[23,24]. 假设 $\Gamma = \{g_1, g_2, g_3, \dots\}$ 代表一组实值函数, $\text{span}(\Gamma)$ 表示 Γ 张成的函数空间, $\text{span}(\Gamma)$ 稠密于 L_2 空间且 $\forall g \in \Gamma$, $0 < \|g\| < b_g, b_g \in \mathbb{R}^+$ 。

对于 $L = 1, 2, 3, \dots$, 定义 $\delta_L = \sum_{q=1}^m \delta_{L,q}, \delta_{L,q} = (1 - r - \mu_L) \|e_{L-1,q}\|^2$, 其中, $0 < r < 1; \mu_L \leq 1 - r, \lim_{L \rightarrow +\infty} \mu_L = 0$ 。

若随机基函数 g_L 满足下列不等式约束:

$$\langle e_{L-1,q}, g_L \rangle^2 \geq b_g^2 \delta_{L,q}, q = 1, 2, \dots, m \quad (7)$$

隐含层节点输出权重:

$$\beta = [\beta_1, \beta_2, \dots, \beta_L] = \arg \min_{\beta} \|f - \sum_{j=1}^L \beta_j g_j\| \quad (8)$$

则 $\lim_{L \rightarrow +\infty} \|f - f_L\| = 0$ 。

1.3 SCN 优、劣势分析

1.3.1 SCN 优势

随机配置 (stochastic configuration, SC) 算法与传统的神经网络学习方法 (BP 算法等)、无不等式约束的随机

化权重方法 (random weights, RW) 相比, 其主要优势如下.

1) 传统的随机化神经网络的无限逼近性取决于隐含层节点数量和随机参数范围, 如果参数设置不当, 模型无法以很高的概率逼近目标函数, SCN 采取增量式建模方法, 根据预设参数范围, 利用监督机制配置隐含层节点的输入权重和偏置, 保证了良好的通用逼近性.

2) 与深度神经网络模型相比, 其通过伪逆运算计算隐含层节点的输出权重, 无需通过反向传播迭代更新, 提高了学习速率的同时, 有效避免陷入局部最优、梯度消失或爆炸等问题, 适用于实时性需求高的应用场景, 且易于实现并行、集成操作.

3) SCN 参数设置具有较少的人为干预, 其随机参数通过监督机制依赖数据产生, 当模型未达到预设误差时, 其能自适应增加隐含层节点数或扩展为 DeepSCN^[24], 避免重新调整网络结构, 消耗训练时间, 具有较优的泛化性能.

1.3.2 SCN 缺点与不足

尽管 SCN 存在上述优势, 但是仍存在以下缺点与不足.

1) SCN 的不等式约束及其监督机制提高了随机化神经网络算法的通用逼近性能和学习精度, 但是增量式构建方式, 一定程度上增加了其时间复杂度, 影响了学习效率.

2) SCN 采用随机化参数配置方式, 相比于基于反向传播的深度神经网络模型等, 节点参数不能得到足够优化, 在较大规模图像数据分析方面, 虽能节省训练时间, 学习精度仍存在差异.

3) 不等式约束以及监督机制降低了人为干预程度, 但是 SCN 仍存在部分超参数需要人为设置, 缺少理论支撑.

2 随机配置网络变体

2.1 深度 SCN

近 10 年来, 深度学习模型因其强大的学习表征能力, 在计算机视觉、模式识别、自然语言处理等领域取得了巨大的成功^[7-12]. 但是, 深度神经网络模型结构复杂、训练速度慢且存在大量的网络参数, 易出现梯度爆炸等问题. 深度 SCN (deep stochastic configuration network, DeepSCN)^[24]提出于 2018 年, 网络结构如图 2 所示, 其配置过程与 SCN 相似, 网络结构由一个隐含层和一个隐含层节点开始, 逐渐增加隐含层节点个数和隐含层层数, 通过监督机制配置隐含层节点参数, 并将所有隐含层都直接连接到输出层, 利用最小二乘法计算输出权重.

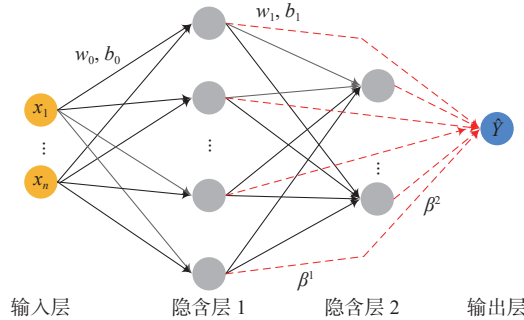


图 2 深度随机配置网络结构^[24]

DeepSCN 构建过程可以描述如下^[24].

给定一组训练数据集 $\{X, Y\}$, $X = \{x_1, x_2, \dots, x_N\}$ 表示特征数据, $x_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,d}\} \in \mathbb{R}^d$; $Y = \{y_1, y_2, \dots, y_N\}$ 表示相应的标签数据, $y_i = \{y_{i,1}, y_{i,2}, \dots, y_{i,m}\} \in \mathbb{R}^m$; $i = 1, 2, \dots, N$.

对于目标函数 $f: \mathbb{R}^d \rightarrow \mathbb{R}^m$, 假设隐含层 n 的第 $L_n - 1$ 个节点已经配置完成, 则当前残差向量为:

$$e_{L_n-1}^{(n)} = e_{L_n-1}^{(n)}(X) = f - f_{L_n-1}^{(n)}(X) = [e_{L_n-1,1}^{(n)}(X), e_{L_n-1,2}^{(n)}(X), \dots, e_{L_n-1,m}^{(n)}(X)]^T \quad (9)$$

其中, $e_{L_n-1,q}^{(n)}(X) = [e_{L_n-1,q}^{(n)}(x_1), e_{L_n-1,q}^{(n)}(x_2), \dots, e_{L_n-1,q}^{(n)}(x_N)] \in \mathbb{R}^N$, $q = 1, 2, \dots, m$.

如 $\|e_{L_n-1}^{(n)}\|^2$ 未达到预设误差 ε 或未达到最大节点数 $L_{\max}^{(n)}$, 增加第 L_n 个节点, 利用公式 (11) 计算当前节点输出 $h_{L_n}^{(n)}$:

$$X^{(n-1)} = \Phi(X^{(n-2)}, W^{(n-2)}, B^{(n-2)}) = [\phi_{n-1,1}(X^{(n-2)}), \phi_{n-1,2}(X^{(n-2)}), \dots, \phi_{n-1,L_{n-1}}(X^{(n-2)})], n \geq 2 \quad (10)$$

$$h_{L_n}^{(n)} = h_{L_n}^{(n)}(X) = [\phi_{n,L_n}(x_1^{(n-1)}), \phi_{n,L_n}(x_2^{(n-1)}), \dots, \phi_{n,L_n}(x_N^{(n-1)})]^T \quad (11)$$

其中, $\phi(\cdot)$ 表示 DeepSCN 的激活函数; $X^{(0)} = X = [x_1, x_2, \dots, x_N]$; $W^{(n-2)} = [w_1^{(n-2)}, w_2^{(n-2)}, \dots, w_{L_{n-1}}^{(n-2)}]$; $B^{(n-2)} = [b_1^{(n-2)}, b_2^{(n-2)}, \dots, b_{L_{n-1}}^{(n-2)}]$; $w_{L_{n-1}}^{(n-1)}$ 和 $b_{L_{n-1}}^{(n-1)}$ 分别表示隐含层 n 第 L_n 个节点的输入权重和偏置。

根据公式 (12) 监督机制确定其输入权重和偏置:

$$\xi_{L_n,q}^{(n)} = \frac{\langle e_{L_{n-1},q}^{(n)}, h_{L_n}^{(n)} \rangle^2}{\langle h_{L_n}^{(n)}, h_{L_n}^{(n)} \rangle^2} - (1-r) \langle e_{L_{n-1},q}^{(n)}, e_{L_{n-1},q}^{(n)} \rangle^2 \quad (12)$$

其中, $q = 1, 2, \dots, m$; $r \in (0, 1)$; 满足 $\xi_{L_n,q}^{(n)} = \sum_{q=1}^m \xi_{L_n,q}^{(n)} \geq 0$ 最大值的候选节点参数作为第 L_n 个节点参数。

确定隐含层节点输出权重:

$$\beta = \arg \min_{\beta} \|H\beta - Y\|^2 = H^+ Y \quad (13)$$

其中, H^+ 表示 H 的摩尔-彭若斯 (Moore-Penrose) 广义逆, $H = [H_{L_1}^{(1)}, H_{L_2}^{(2)}, \dots, H_{L_n}^{(n)}]$, $H_{L_n}^{(n)} = [h_1^{(n)}, h_2^{(n)}, \dots, h_{L_n}^{(n)}]$ 。

为说明 DeepSCN 的优越性, 文献 [24] 通过图 3 所示的合成手写数字数据集预测手写数字的旋转角度, 包含其中训练集和测试集各包含 5 000 张不同旋转角度的灰度图像, 每一幅图像大小为 (28×28) 。以可接受误差界限内的预测值的百分比 (percentage of predictions within an acceptable error margin, *PPA*) (公式 (14)) 和预测旋转角度和实际旋转角度的均方根误差 (root mean square error, *RMSE*) (公式 (15)), 实验结果见表 1, 其中, SCN^1 、 SCN^2 各隐含层节点数量分别为 1 000、2 000, DeepSCN¹ 和 DeepSCN² 均采用 4 个隐含层, DeepSCN¹ 和 DeepSCN² 每个隐含层节点数量分别设置为 250、500。DeepSCN 相比于 SCN 具有较大的 *PPA* 和较小的 *RMSE* 值, 说明了 DeepSCN 模型的有效性。

$$PPA = \frac{N_p}{N} \quad (14)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - f_i)^2} \quad (15)$$

其中, N 表示样本数量; N_p 表示可接受误差界限内 (阈值为 10) 的预测结果的数量; y_i 表示样本 i 的实际旋转角度; f_i 表示样本 i 的预测旋转角度。

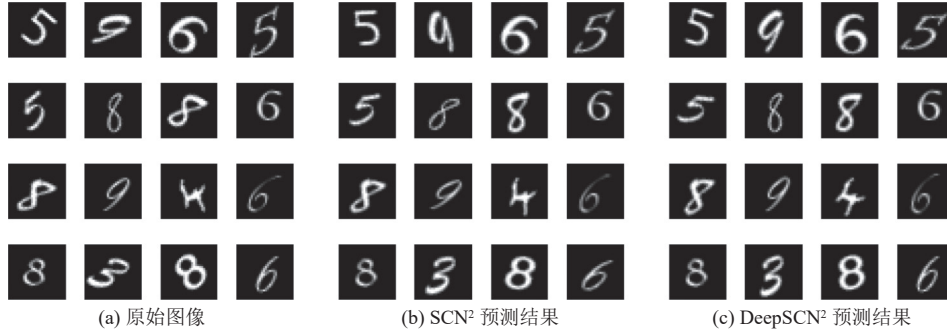


图 3 手写数字数据集样本及 SCN^2 、DeepSCN² 预测结果^[24]

2.2 二维 SCN

为保证图像、视频等数据信息的完整性, 研究具有矩阵输入或多维输入的全连接神经网络 (fully connected neural network, FCNN) 理论和算法具有重要意义。Gao 等人首次提出了矩阵神经网络 (matrix neural network, MatNet)^[25], 将传统的 BP 算法扩展为 2D 输入版本, 并验证了其处理图像数据的优势^[25,26]。但是, MatNet 仍存在局部最优和收敛速度慢等梯度下降方法潜在的缺陷, Lu 等人提出了具有矩阵输入的 RVFL (2DRVFL), 通过构建 2D

随机学习模型提高了图像数据分析的高效性^[27]. 为克服 RVFL 的缺点, Li 等人将 SCN 扩展为二维 SCN (two dimensional stochastic configuration network, 2DSCN), 利用 SCN 的监督机制配置隐含层节点参数, 并应用于解决图像数据建模问题^[28].

表 1 DeepSCN 和 SCN 的性能对比

算法	PPA (%)		RMSE	
	训练结果	测试结果	训练结果	测试结果
DeepSCN ¹ $L_1=L_2=L_3=L_4=250$	82.21	65.21	7.0265	14.8231
SCN ¹ $L=1000$	78.14	60.35	7.5623	15.9176
DeepSCN ² $L_1=L_2=L_3=L_4=500$	84.73	66.36	5.2324	12.0971
SCN ² $L=2000$	80.12	60.00	5.6797	14.7775

2DSCN 构建过程可以描述如下^[28].

给定一组训练数据集 $\{X, Y\}$, $X = \{x_1, x_2, \dots, x_N\}$ 表示特征数据, $x_i \in \mathbb{R}^{d_1 \times d_2}$; $Y = \{y_1, y_2, \dots, y_N\}$ 表示相应的标签数据, $y_i = \{y_{i,1}, y_{i,2}, \dots, y_{i,m}\} \in \mathbb{R}^m$; $i = 1, 2, \dots, N$.

对于目标函数 $f: \mathbb{R}^{d_1 \times d_2} \rightarrow \mathbb{R}^m$, 假设 2DSCN 的 $L-1$ 个隐含层节点已经配置完成, 并计算当前残差向量 (公式 (17)):

$$f_{L-1}(x) = \sum_{j=1}^{L-1} \beta_j g_j(u_j^T x v_j + b_j), \quad L = 1, 2, \dots, L_{\max}, \quad f_0 = 0 \quad (16)$$

$$e_{L-1} = f - f_{L-1} = [e_{L-1,1}, e_{L-1,2}, \dots, e_{L-1,m}]^T \quad (17)$$

其中, β_j 表示隐含层节点 j 的输出权重; $g_j(\cdot)$ 表示激活函数; u_j 和 v_j 表示隐含层节点 j 的输入权重; b_j 表示其偏置; $j = 1, 2, \dots, L_{\max}$.

2DSCN 隐含层节点 L 的输出 $h_L(X)$ 定义为:

$$h_L(X) = [g_L(u_L x_1 v_L + b_L), \dots, g_L(u_L x_N v_L + b_L)]^T, \quad u_L \in \mathbb{R}^{d_1}, \quad v_L \in \mathbb{R}^{d_2} \quad (18)$$

其中, $g_L(\cdot)$ 表示节点 L 的激活函数; u_L 和 v_L 表示其输入权重; b_L 表示偏置.

定义一组时间变量 $\xi_{L,q}$, 其中 $q = 1, 2, \dots, m$:

$$\xi_{L,q} = \frac{[e_{L-1,q}^T(X) \cdot h_L(X)]^2}{h_L^T(X) \cdot h_L(X)} - (1-r)e_{L-1,q}^T(X)e_{L-1,q}(X) \quad (19)$$

其中, $r \in (0, 1)$; 满足 $\xi_L = \sum_{q=1}^m \xi_{L,q} \geq 0$ 最大值的候选节点参数作为第 L 个节点参数.

根据最小二乘法, 计算隐含层节点输出权重:

$$\beta = \arg \min_{\beta} \|H\beta - Y\|^2 = H^+ Y \quad (20)$$

其中, $\beta = [\beta_1, \beta_2, \dots, \beta_L]$, $H = [h_1, h_2, \dots, h_L]$.

为说明 2DSCN 的有效性, 表 2 展示了 2DSCN 和经典深度 CNN 模型在 CIFAR-10 数据集上的训练和测试结果. 其中, 训练数据随机选择“鹿”“狗”“青蛙”“猫”这 4 类样本各 500 幅组成, 测试数据为测试集中的所有 4 类样本 4 000 幅组成, 具体实验环境和实验参数见文献 [28].

在同样训练样本和测试样本下, 2DSCN 在仅使用单隐含层, 且未使用卷积、池化操作的条件下, 其训练精度和测试精度能够与大多数深度 CNN 模型相当的同时, 消耗最少的训练时间和测试时间, 说明了其在图像分析领域巨大的潜力^[28]. 但是, 实验仅选择“鹿”“狗”“青蛙”“猫”这 4 类样本, 淡化了样本数量对于训练时间的影响, 虽然 2DSCN 训练时间和测试时间存在较大优势, 样本数量的减少未能完全体现经典深度 CNN 模型的分类精度.

表 2 2DSCN 和经典深度 CNN 模型性能对比^[28]

模型	精确度 (%)		计算时间 (s)	
	训练结果	测试结果	训练结果	测试结果
VGG16 ^[29]	90.81	89.08	31 721	978
VGG19 ^[29]	91.17	90.03	38 025	1 267
AlexNet ^[12]	88.29	86.00	2 460	76
ResNet-18 ^[30]	88.88	86.25	6 088	175
ResNet-50 ^[30]	89.00	87.33	18 523	448
ResNet-101 ^[30]	90.27	88.90	30 669	720
GoogLeNet ^[31]	90.47	89.20	7 287	236
DenseNet-201 ^[32]	91.73	90.65	68 955	1 124
Inception-v3 ^[33]	84.88	83.98	25 888	736
Inception-ResNet-v2 ^[34]	85.54	84.70	106 795	1 424
2DSCN ^[28]	88.20	85.63	313	16

2.3 鲁棒 SCN

在实际工程应用中, 传感器易受到设备故障、人为干扰、工作环境等因素影响, 采集数据存在不同程度的噪声和异常值^[35-37], 进而降低学习模型的泛化性. 为有效解决不确定数据回归问题, Wang 等人首次提出了一种鲁棒 SCN (robust stochastic configuration network, Robust SCN), 简称 RSC-KDE, 利用核密度估计 (kernel density estimation, KDE) 方法计算训练样本惩罚权值, 并通过加权最小二乘方法计算输出权重, 以减少噪声数据或异常值对学习模型的负面影响, 其中训练样本惩罚权值和输出权重采用交替优化 (alternating optimization, AO) 策略进行更新^[38], Xie 等人将 RSC-KDE 用于高炉炼铁中铁水质量多输出预测^[39]. 此外, Li 等人^[40]提出了 RSC-MCC 用来处理异常值和噪声数据提高 SCN 的鲁棒性, 通过最大熵准则 (maximum correntropy criterion, MCC) 方法训练输出权值. Dai 等分别使用基于 Huber 损失函数的 M 估计 (M-estimation with Huber loss function)、基于四分位距的 M 估计 (M-estimation with interquartile range)、非参数核密度估计 (nonparametric kernel density estimation, NKDE) 计算 Robust SCN 输出权重, 并提出 Huber-RSCN、IQR-RSCN、NKDE-RSCN 这 3 种模型^[41].

综上所述, 给定一组训练数据集 $\{X, Y\}$, $X = \{x_1, x_2, \dots, x_N\}$ 表示特征数据, $x_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,d}\} \in \mathbb{R}^d$; $Y = \{y_1, y_2, \dots, y_N\}$ 表示相应的标签数据, $y_i = \{y_{i,1}, y_{i,2}, \dots, y_{i,m}\} \in \mathbb{R}^m$; $i = 1, 2, \dots, N$. Robust SCN 模型可通过求解加权最小二乘 (weighted least squares, WLS) 问题建立:

$$\min_{\beta, \theta} \sum_{i=1}^N \theta_i \left\| \sum_{j=1}^L \beta_j \phi(w_j, b_j, x_i) - y_i \right\|^2 \quad (21)$$

其中, $\theta_i \geq 0$ ($i = 1, 2, \dots, N$) 表示样本 i 的惩罚权值; $\phi(\cdot)$ 表示激活函数, L 表示隐含层节点数量, w_j 和 b_j 分别表示节点 j 的权重和偏置, β_j 表示输出权重; $j = 1, 2, \dots, L$.

根据公式 (21), 加权的 $e_{L-1}(X)$ 和 $h_L(X)$ 可以被分别定义为 $e'_{L-1}(X) = \Theta e_{L-1}(X)$, $h'_L(X) = \Theta h_L(X)$, $\Theta = \text{diag}\{\sqrt{\theta_1}, \sqrt{\theta_2}, \dots, \sqrt{\theta_N}\}$, 令 $\xi'_L = \sum_{q=1}^m \xi'_{L,q}$, 则表示为:

$$\xi'_{L,q} = \frac{[e'_{L-1,q}(X) \cdot h'_L(X)]^2}{h'^T_L(X) \cdot h'_L(X)} - (1 - r - \mu_L) e'^T_{L-1,q}(X) e'_{L-1,q}(X) \quad (22)$$

其中, $q = 1, 2, \dots, m$; h'_L 表示隐含层第 L 个节点输出; $r \in (0, 1)$ 表示正则化参数; $\{\mu_L\}$ 表示非负实数序列, 其中 $\mu_L \leq 1 - r$, $\lim_{L \rightarrow +\infty} \mu_L = 0$; 满足 $\xi'_L = \sum_{q=1}^m \xi'_{L,q} \geq 0$ 最大值的候选节点参数作为第 L 个节点参数.

根据惩罚权值, 隐含层节点输出权重计算可得:

$$\beta = \arg \min_{\beta} (H\beta - Y)^T \Lambda (H\beta - Y) = (H^T \Lambda H)^+ H^T \Lambda Y \quad (23)$$

其中, $\beta = [\beta_1, \beta_2, \dots, \beta_L]$, $H = [h_1, h_2, \dots, h_L]$, $\Lambda = \Theta^2 = \text{diag}\{\theta_1, \theta_2, \dots, \theta_N\}$.

除上述描述的 Robust SCN 模型, Wu 等人提出了 Bayesian SCN (Bayesian stochastic configuration networks) 用于处理不确定数据, 通过引入贝叶斯推理 (Bayesian inference, BI) 方法计算 SCN 的输出权值, 并能够根据给定数据自动推断超参数, 为处理噪声数据提供了新的解决方法^[42]. Lu 等人基于贝叶斯框架, 分别提出了一种混合高斯和拉普拉斯分布的 MoGL-SCN 和基于贝叶斯学习的稀疏 SCN (Bayesian-learning-based sparse stochastic configuration network, BSSCN), 并采用期望最大化 (expectation-maximization, EM) 算法估计 MoGLSCN 和 BSSCN 的参数, 提高了 SCN 的鲁棒性^[43,44], 上述模型通过贝叶斯稀疏学习框架提高了神经网络模型的鲁棒性^[45].

2.4 集成 SCN

集成学习能够通过训练多个基学习器模型, 有效结合不同算法或多个模型个体的能力, 提高算法的学习精度, 降低模型过拟合的风险^[46]. 其中, 最常用的集成学习算法包括 Bagging^[47], Boosting^[48]和 Stacking^[49].

Bagging 是一种并行集成方法, 通过 Bootstrap 方法从原始数据集进行采样, 并基于采样子数据集并行训练多个基模型, 使得基模型间具有一定的独立性和差异性, 从而提高集成模型的泛化性能, 且并行训练的方式能节省大量的时间成本. Lu 等人将 Bootstrap 从原始数据集进行采样, 提高了区间预测的鲁棒性^[43,44,50]. 李康等人基于 Bootstrap 采样策略, 提出了一种 Bagging-SCN 集成模型^[51]. Boosting 是一种串行集成方法, 其中 AdaBoost 是 Boosting 算法的典型代表, 能够根据样本训练误差调整基模型的权重并加权提高模型的性能, 对于异常值和噪声数据具有较强的敏感性, 可以有效克服模型的过拟合问题. Qu 等人提出了一种 AdaBoost-SCN 模型, 对多个 SCN 基模型进行迭代训练, 自适应调节基模型权重, 组合加权得到最终模型^[52]. 除此之外, Qu 等人还将 Bootstrap 和 AdaBoost 用于 SCN 节点配置过程中, 提出了 Bootstrap-SCN、AdaBoost-SCN 和 AdaBoost-Bootstrap-SCN 用于分类任务^[53]. Pratama 等人基于 Stacking 方法提出了一种新的深度堆叠随机配置网络 (deep stacked stochastic configuration network, DSSCN), 简化了数据流学习算法, 能够根据数据流自动构建深度堆叠网络结构, 可被认为是一种集成深度 SCN 算法^[54].

除上述基于典型集成算法构建的集成 SCN 模型外, Wang 等人提出了一种基于异构特征的广义神经集成框架用于大规模数据分析^[55]. 该模型以 SCN 为基础学习模型, 采用负相关学习策略 (negative correlation learning, NCL) 计算集成模型输出权重, NCL 能够在集成模型方差项和偏差项不变的情况下, 降低基模型之间的协方差, 被广泛应用于神经网络集成模型构建^[56,57].

其中, 具体 SCNE-NCL 算法描述如下.

将原始的数据集 $\{X, Y\}$ 根据先验知识或其他特征选择方法构建异构特征集, 每个 SCN 基模型对应的样本数据集定义为 $D_m = \{(x_n^{(m)}, y_n), n = 1, 2, \dots, N\}$, $x_n^{(m)} \in \mathbb{R}^{d_m}$, $y_n \in \mathbb{R}$, $m = 1, 2, \dots, M$. 其模型建立过程分为 SCN 基模型配置和基于 NCL 算法的模型集成 2 部分组成.

在 SCNE-NCL 中, SCN 基模型的节点输入权重和偏置由 SC 算法配置完成, 集成模型的输出权重 $\{\beta_1, \beta_2, \dots, \beta_M\}$ 可由 NCL 得到, 即:

$$\{\beta_1, \beta_2, \dots, \beta_M\} = \arg \min_{\beta_m} \{e_m\}, m = 1, 2, \dots, M \quad (24)$$

具有 NCL 惩罚项的 SCN 基模型 f_m 的代价函数 e_m 可表示为公式 (25) 矩阵形式:

$$e_m = \frac{1}{2} \left(\|H_m \beta_m - Y\|^2 - \lambda \|H_m \beta_m - \frac{1}{M} H B\|^2 \right) \quad (25)$$

其中, H_m 表示基模型 m 的隐含层输出矩阵; B 表示 SCNE-NCL 的输出权重; $0 < \lambda < 1$ 表示正则化参数; $H = [H_1, H_2, \dots, H_M]_{N \times L}$, L 表示 SCNE-NCL 中的隐含层节点总数.

Wang 等人^[55]针对 SCNE-NCL 的输出权重 B 的计算过程分别给出了伪逆计算方法和 Block Jacobi 和 Gauss-Seidel 迭代求解方法, 其中伪逆计算方法如下.

通过对代价函数 e_m 进行求导:

$$\frac{\partial e_m}{\partial \beta_m} = c_1 H_m^T H_m \beta_m + c_2 H_m^T \tilde{H}_m B - H_m^T Y = 0, m = 1, 2, \dots, M \quad (26)$$

可得:

$$\tilde{H}_m = [H_1, \dots, H_{m-1}, 0_{N \times L_m}, H_{m+1}, \dots, H_M]_{N \times L} \quad (27)$$

$$c_1 = 1 - \frac{\lambda(M-1)^2}{M^2}, c_2 = \frac{\lambda(M-1)^2}{M^2} \quad (28)$$

根据公式 (26) 能够得到一个巨大的线性方程组 $\mathbb{H}B = H^T y$, 其中:

$$\mathbb{H} = \begin{bmatrix} c_1 H_1^T H_1 & c_2 H_1^T H_2 & \dots & c_2 H_1^T H_M \\ c_2 H_2^T H_1 & c_1 H_2^T H_2 & \dots & c_2 H_2^T H_M \\ \vdots & \vdots & \ddots & \vdots \\ c_2 H_M^T H_1 & c_2 H_M^T H_2 & \dots & c_1 H_M^T H_M \end{bmatrix}_{L \times L} \quad (29)$$

通过伪逆运算可得 $B = \mathbb{H}^+ H^T Y$, 其中 $\mathbb{H}^+$ 表示矩阵 $\mathbb{H}$ 的伪逆.

此外, Huang 等人分析了影响随机化模型泛化能力的因素, 提出了一个具有可选择性的神经网络集成框架, 能够根据基模型泛化能力确定较优的基模型组合, 并以 SCN 为例进行了验证^[58].

2.5 分布式并行 SCN

分布式技术和并行技术作为 2 种不同的体系架构, 均能通过并行运行多个算法, 提高算法运行速度, 进行大规模数据处理^[59]. 与集中式 SCN 不同, 假设分布式 SCN (distributed stochastic configuration network, Distributed SCN) 通信网络中包含 V 个 agent, 每个 agent 对应的样本数据集定义为 $D_i = \{(x_{n,i}, y_{n,i}), n = 1, 2, \dots, N_i\}$, $x_{n,i} \in \mathbb{R}^d$, $y_{n,i} \in \mathbb{R}^m$, $i = 1, 2, \dots, V$, 则在分布式学习问题中全局权重 β 可以被定义为:

$$\beta = \arg \min_{\beta} \frac{1}{2} \sum_{i=1}^V \sum_{n=1}^{N_i} \|e_{n,i}\|^2 \text{ s.t. } h(x_{n,i})\beta = Y_{n,i}^T - e_{n,i}^T \quad (30)$$

则隐含层输出矩阵 $H = [H_1, H_2, \dots, H_V]^T$, $H_i = [h_{1,i}, h_{2,i}, \dots, h_{N_i,i}]^T$; $Y = [Y_1, Y_2, \dots, Y_V]^T$, $Y_i = [y_{1,i}, y_{2,i}, \dots, y_{N_i,i}]^T$.

将约束条件 (30) 代入目标函数, 得到优化问题:

$$\arg \min_{\beta} \frac{1}{2} \sum_{i=1}^V \|H_i \beta - Y_i\|^2 \quad (31)$$

为降低模型过拟合风险, 通常使用 L_2 正则化惩罚项修正目标函数, Ai 等人在基于合作范式的 Distributed SCN 中定义目标函数为公式 (32), 并使用并行交替方向乘子法 (parallel alternating direction method of multipliers, PADMM) 迭代优化模型参数^[60], Ge 等人使用离散时间零梯度和 (zero-gradient-sum, ZGS) 算法计算模型参数^[61]:

$$\arg \min_{\beta} \frac{1}{2} \sum_{i=1}^V \|H_i \beta_i - Y_i\|^2 + \frac{C}{2V} \sum_{i=1}^V \|\beta_i\|^2, C \in [0, 1] \quad (32)$$

Zhao 等人在目标函数公式 (31) 中加入 Elastic-Net 正则化, 修正目标函数为公式 (33), 并使用 ADMM 算法求解模型^[62]:

$$\arg \min_{z, \beta} \frac{1}{2} \sum_{i=1}^V \|H_i \beta_i - Y_i\|^2 + CP(z), C \in [0, 1] \quad (33)$$

$$P(z) = \alpha \|z\| + \frac{1-\alpha}{2} \|z\|^2, \alpha \in [0, 1] \quad (34)$$

Zhou 等人使用 sparse ADMM 优化公式 (35)^[63]:

$$\arg \min_{\beta} \frac{1}{2} \sum_{i=1}^V \|H_i \beta_i - Y_i\|^2 + \frac{C}{2V} \sum_{i=1}^V \|\beta_i\|, C \in [0, 1] \quad (35)$$

除上述 Distributed SCN, Dai 等人提出一种模型与数据混合并行的随机配置网络 (hybrid parallel stochastic configuration networks, HPSCN), 通过构建不同方式的 2 个 SCN 模型建立样本数据的动态分块方法, 加快了候选节点参数的建立, 降低了模型计算量^[64]. Zhang 等人提出了一种基于数据融合的并行 SCN (parallel stochastic configuration network, Parallel SCN), 通过训练不同目标函数的 SCN 模型, 并对输出结果进行数据融合, 提高了模型的鲁棒性^[65].

2.6 正则化 SCN

尽管 SCN 的监督机制保障了其无限逼近性能, 但是随着隐含层节点数量的增加, 容易面临过拟合的风险, 采

取一定的正则化策略, 有利于提高模型的泛化性能.

2.6.1 L_1 和 L_2 正则化

传统的 SCN 根据公式 (5) 计算隐含层节点的输出权重, 基于 L_1 和 L_2 正则化的 SCN 输出权重 β 由误差项和正则化惩罚项共同决定^[66,67], 分别通过公式 (36) 和公式 (37) 计算:

$$\beta = \arg \min_{\beta} \|H\beta - Y\|^2 + l_1 \|\beta\| \quad (36)$$

$$\beta = \arg \min_{\beta} \|H\beta - Y\|^2 + l_2 \|\beta\|^2 \quad (37)$$

其中, l_1 和 l_2 分别表示 L_1 和 L_2 的正则化系数, 基于 L_1 正则化容易得到稀疏解. 相比于 L_1 正则化, L_2 正则化能够得到较为平滑的解比较平滑, 有利于降低模型的复杂度.

其中, 王前进等人将 L_1 正则化引入到 SCN, 提出一种简约随机配置网络 (parsimonious stochastic configuration network, Parsimonious SCN), 并利用交替方向乘子 (alternating direction method of multipliers, ADMM) 算法更新输出权重 β ^[66]. 赵立杰等人^[67]和王前进等人^[68]将 L_2 正则化引入到 SCN, 并给出了权重 β 计算方式:

$$\beta = (H^T H + l_2 I)^{-1} H^T Y \quad (38)$$

与此同时, 在 Distributed SCN 的构造过程中, 分布式模型的最终权重计算中同样引入了正则化范数以提高模型的泛化性^[60-63]. 除此之外, 还可采用基于 L_1 和 L_2 正则化混合的模式计算输出权重 β , 均衡稀疏性和平滑性:

$$\beta = \arg \min_{\beta} \|H\beta - Y\|^2 + l_1 \|\beta\| + l_2 \|\beta\|^2 \quad (39)$$

2.6.2 Dropout 正则化

除 L_1 和 L_2 正则化方法外, Dropout 技术被广泛用于深度模型, 通过在输出层增加误差反馈, 动态调整神经网络参数, 提高模型泛化能力. Li 等人提出了一种基于 Dropout 的 SCN, 每一个节点服从概率为 p 的伯努利 (Bernoulli) 分布, 确保节点之间的随机性和独立性, 其输出权重 β 的目标函数被修正为^[69]:

$$\beta = \arg \min_{\beta} \|H\hat{\beta} - Y\|^2 + p(1-p)\|\Gamma\hat{\beta}\|^2 \quad (40)$$

其中, p 表示 Dropout 概率; $\hat{\beta} = p\beta$; $\Gamma = [\text{diag}(H^T H)]^{1/2}$ 表示不同节点之间的标准偏差, $\text{diag}(H^T H)$ 表示由 $H^T H$ 的对角元素组成的对角矩阵.

与标准 L_2 正则化方法不同, Dropout-SCN 能够对输出权重 β 进行约束, 可通过计算标准最小二乘法得到最终权重:

$$\beta = [H^T H + p(1-p)\text{diag}(H^T H)]^{-1} H^T Y \quad (41)$$

2.6.3 流形正则化

流形学习 (manifold learning) 方法能够利用非线性多元系统变量之间的几何关系, 保持流形中相邻数据点的局部关系, 挖掘数据关键信息, 降低数据冗余. 潘承燕等人提出了一种流形正则化随机配置网络 (manifold regularized stochastic configuration network, MRSCN)^[70].

相比 SCN 训练过程, MRSCN 能够挖掘增量式节点的结构信息, 体现样本的局部几何结构, 其目标函数定义为:

$$\min_{\beta} \sum_{i=1}^N \frac{1}{2} \left\| \sum_{j=1}^L \beta_j g(w_j, b_j, x_i) - y_i \right\|^2 + \frac{\lambda_M}{2} \text{tr}[(h_L \beta_L)^T L_p h_L \beta_L] \quad (42)$$

其中, λ_M 表示流形正则化参数, L_p 表示训练样本的图拉普拉斯矩阵 (graph laplacians). $\frac{\lambda_M}{2} \text{tr}[(h_L \beta_L)^T L_p h_L \beta_L]$ 表示流形正则化参数项, 用于学习样本局部几何特征.

$\xi_{L,q}$ 可表示为:

$$\xi_{L,q} = \frac{(e_{L,q}^T \cdot h_L)^2}{(h_L^T h_L + \lambda_M h_L^T L_p h_L) / (h_L^T h_L + 2\lambda_M h_L^T L_p h_L)} - (1-r-\mu_L) e_{L-1,q}^T e_{L-1,q} \quad (43)$$

其输出权重 β 的目标函数为:

$$\beta = \arg \min_{\beta} \frac{1}{2} \|H\beta - Y\|^2 + \frac{\lambda_M}{2} \text{tr}[(H\beta)^T L_p(H\beta)] \quad (44)$$

通过计算标准最小二乘法得到最终权重:

$$\beta = (H^T H + \lambda_M H^T L_p H)^{-1} H^T Y \quad (45)$$

2.7 其他变体

除上述几类典型 SCN 变体外, 国内外学者针对不同的应用场景设计了诸多 SCN 变体. 因篇幅有限, 本节简要总结了其他几种重要变体.

2.7.1 基于 SCN 的深度神经网络模型

在图像分析领域, 自动编码器 (auto encoder, AE) 能够从输入数据中提取高级抽象特征信息, Pan 等人为了提高模型精度, 保障运行效率, 将 SCN 和 AE 进行结合, 提出了一种 SCN-AE 模型. 该模型 AE 部分采用随机化编码器权重和偏置分配模式, 分别利用公式 (46) 和公式 (47) 更新解码器输出权重 β_{de} 和输出结果 H_{de} , 并将输出结果 H_{de} 作为 SCN 的输入, 提高了图像分类的精度^[71].

$$\beta_{de} = (CI + H_{en}^T)^{-1} H_{en}^T X \quad (46)$$

$$H_{de} = \phi(\beta_{de}^T X + b_{de}) \quad (47)$$

其中, C 表示正则化参数; H_{en} 表示编码器的输出; X 表示输入图像特征数据; b_{de} 表示解码器随机偏置.

CNN 通过卷积、激活、池化等特征学习过程, 极大提高了图像数据分类的准确性, Li 等人^[72]和 Zhang 等人^[73]利用 SCN 的通用逼近性能, 将 SCN 作为自适应卷积神经网络 (adaptive convolutional neural network, ACNN) 的分类器, 分别采用交替优化 (alternating optimization, AO) 技术和动态迁移学习机制对 ACNN 模型进行更新, 提高了模型的泛化能力和识别精度. Liu 等人将超分辨率感知卷积神经网络 (super-resolution perception convolutional neural network, SRPCNN) 和 SCN 进行结合, 提出了一种光伏发电预测模型, 能够补全缺失数据, 提高预测精度^[74].

2.7.2 SCN 优化模型

SCN 引入了监督机制, 利用不等式约束对随机参数进行赋值, 并自适应地选择随机参数的取值范围, 通过最小二乘法计算输出节点权重, 保证了学习模型的通用逼近性, 在较少人为干预下, 提高了网络的学习精度和学习效率^[23].

为进一步优化 SCN 模型, 提高学习效率和学习精度, Sheng 等人通过自适应选择截断阈值的截断奇异值分解 (truncated singular value decomposition, TSVD) 方法计算 SCN 输出权重, 提高了模型的精度和泛化性能^[75]. Dai 等人^[76]和 Tian 等人^[77]分别提出了不同形式的变增量 SCN 模型 (stochastic configuration networks with block increments, Block SCN) 和 (stochastic configuration network with variable increments, vi-SCN) 替代原有的点增量配置模式, Cao 等人提出了一种双向 SCN (bidirectional stochastic configuration network, Bidirectional SCN)^[78], 提高了模型的运算效率. Wang 等人提出了一种基于驱动量的 SCN (driving amount based stochastic configuration network, DASCN) 算法, 提高了 SCN 的泛化性和结构性^[79].

参数优化方面, Li 等人提出了一种基于遗传算法的 (genetic algorithm, GA) 的 SCN 模型, 通过 GA 优化随机参数的控制范围的矩阵^[80]. Zhang 等人分别使用天牛须搜索 (beetle antennae search, BAS) 和混沌麻雀搜索算法 (chaotic sparrow search algorithm, CSSA) 2 种不同形式的生物启发式优化算法搜索 SCN 随机参数控制因子^[65,81]. Zhu 等人提出了 2 个新的不等式约束分配隐含层节点随机参数, 提高了 SCN 的学习效率^[82]. Niu 等人使用不同重尾分布 (heavy-tailed distributions) 方法更新隐含层节点权重和偏置, 验证了基于重尾分布的 SCN 模型的最优鲁棒性^[83].

除此之外, 一些混合模型充分利用了传统模型和 SCN 的优势. 赵立杰等人利用偏最小二乘方法 (partial least squares, PLS) 替代经典的最小二乘方法, 通过正交投影增强了模型的泛化性能^[84]. Wang 等人利用粗糙集 (rough set) 对数据约简的优势提出了 RS-SCN, 相比 SCN, 进一步提高了预测精度^[85]. Wang 等人提出了一种基于广义可加模型 (generalized additive model) 和 SCN 的混合模型, 验证了混合模型的有效性^[86]. Felicetti 等人针对 DeepSCN 超参数优化问题, 引入了蒙特卡洛树搜索 (Monte-Carlo tree search, MCTS) 和随机搜索策略^[87].

2.8 小 结

自 SCN 诞生以来,广泛吸引了国内外学者的研究兴趣和关注,本节总结了深度 SCN、二维 SCN、鲁棒 SCN、集成 SCN、分布式并行 SCN、正则化 SCN 等代表性变体算法及其特点,如表 3 所示.此外,并对具有里程碑意义的变体进行总结,如图 4 所示.

表 3 SCN 主要变体分析

所属类别	代表算法	主要特点
深度SCN	DeepSCN ^[24]	增量式深度神经网络模型
二维SCN	2DSCN ^[28]	矩阵输入,主要用于处理图像数据
鲁棒SCN	RSC-KDE ^[38] 、RSC-MCC ^[40] 、Huber-RSCN ^[41] 、IQR-RSCN ^[41] 、NKDE-RSCN ^[41] 、Bayesian SCN ^[42] 、MoGL-SCN ^[43] 、MoGLSCN ^[44] 、BSSCN ^[44]	主要用于处理含有一定比例噪声或异常值的数据
集成SCN	Bagging-SCN ^[51] 、Bootstrap-SCN ^[53] 、AdaBoost-SCN ^[53] 和AdaBoost-Bootstrap-SCN ^[53] 、DSSCN ^[54] 、SCNE-NCL ^[55] 、K-RNNE ^[58]	主要用于模型集成,提高算法泛化能力
分布式并行SCN	PADMM-SCN ^[60] 、ZGS-SCN-ET ^[61] 、ADMM-SCN-ENet ^[62] 、SADMM-SCN ^[63] 、HPSCN ^[64] 、PSCN ^[65]	主要用于分布式数据处理,克服能耗、通信带宽、计算能力限制
正则化SCN	Parsimonious SCN ^[66] 、L ₂ -RSCN ^[67,68] 、Dropout-SCN ^[69] 、MRSCN ^[70]	主要用于解决模型因节点增多而导致的过拟合问题
其他变体	SCN-AE ^[71] 、ACNN-SCN ^[72,73] 、SCN-TSVD ^[75] 、Block SCN ^[76] 、vi-SCN ^[77] 、Bidirectional SCN ^[78] 、DASCN ^[79] 、GA-SCN ^[80] 、CSSA-SCN ^[81] 、PLS-SCN ^[84] 、RS-SCN ^[85] 、GAM-SCN ^[86] 、Optimized DeepSCN ^[87]	不同角度的SCN优化变体模型

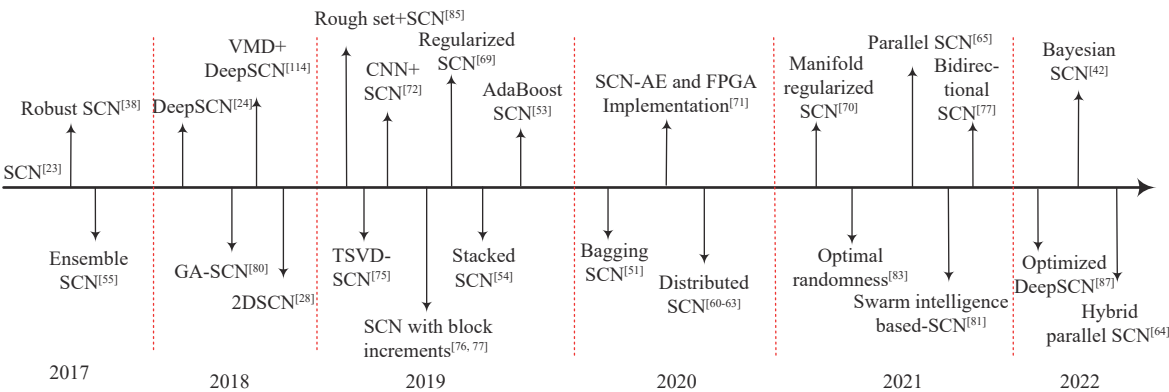


图 4 随机配置网络发展过程

3 应用领域

SCN 作为一种新型的随机化神经网络模型,其基于监督机制的增量式节点配置模式,使得 SCN 训练过程具有计算参数量小、学习效率高、人为干预程度低和泛化能力强等优点.因此,SCN 自提出以来,已经逐渐应用于硬件实现、计算机视觉、医学数据分析、故障检测与诊断以及各类系统建模预测等领域,并取得了巨大的成功.

3.1 硬件实现

现场可编程门阵列 (field-programmable gate array, FPGA) 具有成本低、精度高、处理速度快和可重构性等优点,适用于工业领域等实时性强的应用场景^[88,89].近年来,随着 FPGA 技术的快速发展,促进了神经网络等人工智能算法的硬件实现,有利于降低网络负载,节省计算资源^[90,91].相比于深度神经网络模型,SCN 等随机化神经网络

模型具有模型复杂度低、参数量少、运算效率高等优势,易于基于硬件系统实现,保障系统的正常运行. Pan 等人基于 FPGA 框架实现了 SCN-AE 算法并应用于机器人抓取识别,与基于 Matlab 软件实现的 SCN-AE 结果相似,取得了较高的识别精度^[71].与此同时, Gao 等人以提高 SCN 预测精度和 FPGA 硬件资源利用率为目标,提出了一种新的非线性激活函数,基于 FPGA 硬件实现了 SCN 的回归预测,显著提高了预测性能^[92].

3.2 计算机视觉

计算设备性能的不断提高,使得计算机视觉得到了快速的发展,并应用于医疗、金融、军事等诸多行业.目前深度学习在机器视觉领域取得了巨大的成功,但是存在模型参数量大、计算复杂度高、训练速度缓慢等问题.近年来,SCN 及其相关变体模型逐渐应用于图像分类、人脸识别、目标检测等图像处理问题,为计算机视觉和图像处理提供了新的思路和方法. Wang 等人将 SCN 用于合成孔径雷达 (synthetic aperture radar, SAR) 图像目标识别^[93]. Cao 等人提出了面向回归问题的 Bidirectional SCN 并用于根据人脸图像年龄预测问题^[78]. Niu 等人探讨了 SCN 的最优鲁棒性,并用于解决手写数字识别问题^[83].为提高 SCN 的特征学习能力,有效获取图像信息, Wang 等人将 SCN 扩展为 DeepSCN,探讨了 SCN 的深度可扩展性,并用于手写数字旋转角度预测^[24]. Niu 等人使用 UAV (unmanned aerial vehicle) 无人机采集图像,并通过 DeepSCN 估计石榴园作物实际蒸散量^[94,95].随后, Li 等人提出了具有矩阵输入的 2DSCN 新模式,提高了 SCN 人脸识别的精确度和手写数字旋转角度预测精度^[28].除此之外,SCN 与传统神经网络模型的结合,能够进一步提高特征学习能力, Pan 等人将 SCN 和 AE 进行结合,提出了一种 SCN-AE 算法并应用于机器人抓取识别^[71]. Li 等人^[72]和 Zhang 等人^[73]将 SCN 和 CNN 框架进行结合,分别用于青梅等级评定和输电线路玻璃绝缘子自爆状态检测,探索了 SCN 和 CNN 模型的可结合性.

3.3 医学数据分析

基于 SCN 的生物医学数据分析,能够为人类疾病预防和诊断、患者康复、健康监测等过程提供关键的隐性信息.脑电图 (electroencephalogram, EEG) 数据记录了大脑活动时无线电波的变化,通常被用作情绪识别和心理负荷评估等应用领域^[96]. Pang 等人利用 SCN 对飞行员的脑电图数据进行心理负荷分类,有利于实现飞行员心理负荷的实时监测,为心理负荷分类提供了一种有效的建模方法^[97]. Wang 等人建立了一种结合集成学习和特征选择的思想,建立了食管鳞状细胞癌 (esophageal squamous cell carcinoma, ESCC) 术后风险预测的集成学习模型,利用 SCN、逻辑回归 (logistic regression, LR)、支持向量机 (support vector machine, SVM)、决策树 (decision tree, DT) 和随机森林 (random forest, RF) 等 5 种模型对 ESCC 术后风险因素的重要性进行排序,并进行术后风险评估^[98].

3.4 故障检测与诊断

为保障工业过程设备系统安全可靠运行,如何针对传感器、物联网采集的系统运行数据进行实时性分析处理已经成为亟需解决的问题^[99,100]. SCN 快速的建模效率为系统故障实时检测和诊断提供了全新的解决方案.王前进等人^[66,68,79]分别将 Parsimonious SCN、Regularized SCN 和 DASCN 用于矿井井下供给风量实时监测,保证矿井工作主通风机平稳运行,以提供充足的井下供给风量.赵立杰等人提出一种基于 L_2 范数正则化的 SCN,并用于矿业领域球磨机负荷工况识别,有利于提高设备的矿石粉碎效率,降低运行成本^[67].潘承燕等人提出一种 MRSCN,用于识别化工过程故障识别^[70].机载电子吊舱作为搭载多功能机载电子设备的主要平台, Pang 等人将 SCN 用于机载电子吊舱温度识别,有利于实时监测电子设备运行环境,保障吊舱内设备可靠运行^[101].与此同时,滚动轴承、齿轮等旋转机械作为机电装备的典型零部件被广泛应用于设备故障诊断与识别仿真建模分析, Zhang 等人将 Parallel SCN 用于滚动轴承剩余使用寿命预测^[65]. Huang 等人提出一种基于变分模态分解 VMD 和 SCN 的齿轮箱振动信号故障诊断方法,取得了更快、更准确的结果^[102]. Liu 等人使用移动时间窗对振动信号进行特征提取,通过 SCN 有效地完成了齿轮故障分类^[103].玻璃绝缘子的自爆状态直接影响输电线路的安全性和可靠性, Zhang 等人利用 SCN 和 CNN 组合框架进行输电线路玻璃绝缘子自爆状态检测^[73].光纤预警系统 (optical fiber pre-warning system, OFPS) 能够实时监测运输管道,确定管道泄漏破坏位置, Li 等人利用基于 Dropout 正则化的 SCN 进行噪声光纤振动信号识别^[69], Qu 等人分别使用基于 AdaBoost-SCN 等不同形式的集成 SCN 提高了识别的准确率^[52,53]. Sheng 等人利用 TSVD 替代了 SVD,提高了 SCN 模型泛化性^[75], Tian 等人提出了 vi-SCN,使用变增量方式进行 SCN 节点

配置,提高了光纤振动信号识别效率^[77]. Chen 等人利用最小冗余、最大关联方法 (minimum redundancy, maximum relevance) 分析了不同时段、频段特征的重要性,并提高了入侵信号的平均识别准确率^[104]. 除此之外, Ding 等人提出了一种基于 SCN 和案例推理 (case-based reasoning, CBR) 的垃圾焚烧过程故障检测模型^[105].

3.5 其他系统建模预测

SCN 除被应用于硬件实现、计算机视觉、医学数据分析、故障检测与诊断外,还被广泛用于各类系统建模预测任务. 在水产养殖领域,为实时监测水产养殖过程中海水氨氮浓度,王魏等人提出了基于 SCN 的海水养殖氨氮浓度软测量模型^[106]. 随后,为提高海水氨氮浓度预测精度, Wang 等人相继提出了基于 Rough set 和 SCN 的混合模型和基于 GAM 和 SCN 的混合模型^[85,86],李康等人分别使用了 Bagging 集成技术和 GA 优化,提出了 Bagging-SCN 和 GA-SCN^[51,80]. 针对原油含碳残余量预测等问题, Lu 等人分别使用 DeepSCN 等 SCN 相关变体构建区间预测模型,克服了点预测的不确定性,提高了预测精度^[43,44,50,107]. 针对非线性系统建模问题,陶金梅等人结合了带遗忘因子的递推最小二乘法 (forgetting factor recursive least square, FFRLS) 和 SCN 进行智能建模^[108]. Lian 等人利用 SCN 预测非线性网络控制系统的数据包无序的非线性序列^[109]. Wang 等人通过 SCN 实现了非线性和高时延的电厂主汽温预测^[110]. 针对云数据中心工作量预测问题, Bi 等人通过 Savitzky-Golay 滤波和小波分解进行特征提取,并使用 SCN 完成了预测任务^[111,112]. 在电力领域, Huang 等人通过 SCN 估计电力数据网络的流量状态^[113], Liu 等人针对光伏发电预测数据缺失问题,使用 SRPCNN 恢复缺失数据,并利用 SCN 预测光伏发电量^[74]. Cai 等人将 VMD 和 DeepSCN 进行结合实现了电能质量扰动的检测与分类^[114]. 赵允文等人提出一种基于相空间重构 (phase space reconstruction, PSR) 和 SCN 的电力负荷短期预测模型,节省了预测时间,提高了预测精度^[115]. Hao 等人提出了一种基于 SCN 和 LM (Levenberg-Marquardt) 算法生成非对称多电平逆变器的选择性谐波消去切换角度,保证了求解效率和切换角度的精度^[116]. 在其他应用领域,针对数据流学习问题, Pratama 等人提出了一种深度堆叠 SCN (deep stacked stochastic configuration network, DSSCN) 用于非平稳数据流学习^[54]. Cao 等人将 Bidirectional SCN 用于空气污染预测问题^[78]. Wang 等人利用 SCN 进行化学溶液组成成分估计^[117]. Han 等人基于随机采样的 SCN 预测抽油机井的淹没深度,有利于抽油机高效运行^[118]. Zhao 等人结合了小波包变换 (wavelet packet transform, WPT) 和 SCN 用于混凝土超声波缺陷检测^[119]. Tian 等人提出了一种 SVM 和 SCN 集成模型,用于短期风速预测^[120].

4 总结与展望

SCN 作为一种具有监督机制的随机化学习模型,克服了传统随机化模型方法的参数选择问题的弊端,提高了模型的泛化性能. 自 SCN 诞生以来,其被广泛地应用于学术研究和工业领域应用,本文对 SCN 理论和应用的最新研究发展进行了全面的综述. 首先,详细介绍了 SCN 的基本理论,包括 SCN 的增量式节点配置方法和输出层权值的伪逆计算方法,阐述了 SCN 的通用逼近性能,分析了 SCN 的优点和不足. 其次,本文列出了 SCN 的里程碑路线图,帮助研究人员了解 SCN 的发展,并详细介绍了深度 SCN、二维 SCN、鲁棒 SCN、集成 SCN、分布式并行 SCN、正则化 SCN 等变体模型,并进行了部分实验数据分析. 此外,总结了 SCN 的当前应用领域,包括硬件实现、计算机视觉、医学数据分析、故障检测与诊断、系统建模与预测等,体现了 SCN 在运算效率和计算精度上的显著优势.

本文认为, SCN 还可在以下几个研究方向进行扩展.

1) 卷积 DeepSCN. DeepSCN、2DSCN 的高效性,展现了 SCN 在图像处理问题中的发展潜能. CNN 作为计算机视觉领域的经典模型,通过卷积、激活、池化等技术极大提高了图像数据分类的准确性,但是同时存在计算复杂度高、参数量大、结构复杂等问题. 将 2D、3D 输入的 DeepSCN 与 CNN 框架进行结合应用于彩色图像、视频数据分析将进一步提高 SCN 在机器视觉领域的性能.

2) 半监督 SCN、无监督 SCN. 半监督学习 (semi-supervised learning, SSL) 和无监督学习 (unsupervised learning) 能够分别通过学习少量的标记数据和未标记数据信息执行模式识别任务. 相比于基于深度神经网络的半监督学习模型,如何将 SCN 独特的监督机制及其通用逼近特性和流形正则化 (manifold regularization) 等图正则化

框架进行结合, 设计一种半监督 SCN 或无监督 SCN, 保证学习精度的同时, 提高其学习效率具有进一步研究的价值。

3) 多视图 SCN. 多视图学习 (multi-view learning) 能够通过充分学习不同视图的数据信息, 更加全面详细的描述事物对象。目前, SCN 已经被广泛应用于诸多应用领域, 但是尚未在多视图学习中得到应用。通过借助后验一致性、边缘一致性等方法, 平衡不同视图特征学习和模型之间的关系, 有利于借助 SCN 的数据依赖性及其无限逼近性, 提高 SCN 模型分类准确性。

4) 模糊 SCN. 模糊神经网络 (fuzzy neural network) 结合了神经网络和模糊系统的优势, 能够通过神经网络训练算法更新模糊规则中与模糊集对应的隶属函数的参数。但是, 面对较大规模, 较高维度数据处理时, 模糊规则参数更新过程消耗大量的训练时间。SCN 采用增量式建模方法, 其监督机制保障了模型的通用逼近性, 借助模糊系统和 SCN 分别学习特征信息并进行知识融合, 构造一种层次融合的模糊深度 SCN, 有利于进一步提高 SCN 模型的收敛性。

5) 循环 SCN. 循环神经网络 (recurrent neural network, RNN) 作为一类具有特殊循环单元的神经网络模型, 能够有效挖掘数据中的时序信息, 被广泛用于处理视频、语音、文本等时间序列数据。传统的 RNN 模型大多采用 BP 算法进行参数更新, 当输入序列较长时, 易产生梯度爆炸或梯度消失问题。因此使用 SC 算法配置 RNN, 用于解决短时记忆的时间序列问题, 具有重要的研究价值。

除上述描述外, 如何进一步扩展 SCN 结构, 提高 DeepSCN 的鲁棒性, 发挥 SCN 人为干预程度低、运算效率高的优势, 将 SCN 应用于迁移学习 (transfer learning)、强化学习 (reinforcement learning, RL)、边缘计算 (edge computing)、类脑智能 (brain-like intelligence) 等研究领域, 都是具有挑战性的研究方向。希望本综述能帮助研究人员理解 SCN 研究, 并进一步完善 SCN 的理论和方法研究。

References:

- [1] Ding SF, Xu X, Wang YR. Optimized density peaks clustering algorithm based on dissimilarity measure. Ruan Jian Xue Bao/Journal of Software, 2020, 31(11): 3321–3333 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5813.htm> [doi: 10.13328/j.cnki.jos.005813]
- [2] Zhang J, Ding SF, Ding L, Zhang CL. Deep generative neural networks based on real-valued RBM with auxiliary hidden units. Ruan Jian Xue Bao/Journal of Software, 2021, 32(12): 3802–3813 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6126.htm> [doi: 10.13328/j.cnki.jos.006126]
- [3] Cybenko G. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems, 1989, 2(4): 303–314. [doi: 10.1007/BF02551274]
- [4] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. Neural Networks, 1989, 2(5): 359–366. [doi: 10.1016/0893-6080(89)90020-8]
- [5] Park J, Sandberg IW. Universal approximation using radial-basis-function networks. Neural Computation, 1991, 3(2): 246–257. [doi: 10.1162/neco.1991.3.2.246]
- [6] Chen TP, Chen H. Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems. IEEE Trans. on Neural Networks, 1995, 6(4): 911–917. [doi: 10.1109/72.392253]
- [7] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521(7553): 436–444. [doi: 10.1038/nature14539]
- [8] Deng L, Yu D. Deep learning: Methods and applications. Foundations and Trends® in Signal Processing, 2014, 7(3–4): 197–387. [doi: 10.1561/20000000039]
- [9] Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. Neural Computation, 2006, 18(7): 1527–1554. [doi: 10.1162/neco.2006.18.7.1527]
- [10] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. Science, 2006, 313(5786): 504–507. [doi: 10.1126/science.1127647]
- [11] Salakhutdinov R, Hinton G. An efficient learning procedure for deep Boltzmann machines. Neural Computation, 2012, 24(8): 1967–2006. [doi: 10.1162/NECO_a_00311]
- [12] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Communications of the ACM, 2017, 60(6): 84–90. [doi: 10.1145/3065386]
- [13] Lukoševičius M, Jaeger H. Reservoir computing approaches to recurrent neural network training. Computer Science Review, 2009, 3(3):

- 127–149. [doi: [10.1016/j.cosrev.2009.03.005](https://doi.org/10.1016/j.cosrev.2009.03.005)]
- [14] Scardapane S, Wang DH. Randomness in neural networks: An overview. *WIREs Data Mining and Knowledge Discovery*, 2017, 7(2): e1200. [doi: [10.1002/widm.1200](https://doi.org/10.1002/widm.1200)]
- [15] Wang DH. Editorial: Randomized algorithms for training neural networks. *Information Sciences*, 2016, 364–365: 126–128. [doi: [10.1016/j.ins.2016.05.021](https://doi.org/10.1016/j.ins.2016.05.021)]
- [16] Pao YH, Takefuji Y. Functional-link net computing: Theory, system architecture, and functionalities. *Computer*, 1992, 25(5): 76–79. [doi: [10.1109/2.144401](https://doi.org/10.1109/2.144401)]
- [17] Pao YH, Park GH, Sobajic DJ. Learning and generalization characteristics of the random vector functional-link net. *Neurocomputing*, 1994, 6(2): 163–180. [doi: [10.1016/0925-2312\(94\)90053-1](https://doi.org/10.1016/0925-2312(94)90053-1)]
- [18] Schmidt WF, Kraaijveld MA, Duin RPW. Feedforward neural networks with random weights. In: *Proc. of the 11th IAPR Int'l Conf. on Pattern Recognition*. The Hague: IEEE, 1992. 1–4. [doi: [10.1109/ICPR.1992.201708](https://doi.org/10.1109/ICPR.1992.201708)]
- [19] Igel'nik B, Pao YH. Stochastic choice of basis functions in adaptive function approximation and the functional-link net. *IEEE Trans. on Neural Networks*, 1995, 6(6): 1320–1329. [doi: [10.1109/72.471375](https://doi.org/10.1109/72.471375)]
- [20] Tyukin IY, Prokhorov DV. Feasibility of random basis function approximators for modeling and control. In: *Proc. of the 2009 IEEE Control Applications, (CCA) & Intelligent Control, (ISIC)*. Saint Petersburg: IEEE, 2009. 1391–1396. [doi: [10.1109/CCA.2009.5281061](https://doi.org/10.1109/CCA.2009.5281061)]
- [21] Gorban AN, Tyukin IY, Prokhorov DV, Sofeikov KI. Approximation with random bases: Pro et Contra. *Information Sciences*, 2016, 364–365: 129–145. [doi: [10.1016/j.ins.2015.09.021](https://doi.org/10.1016/j.ins.2015.09.021)]
- [22] Li M, Wang DH. Insights into randomized algorithms for neural networks: Practical issues and common pitfalls. *Information Sciences*, 2017, 382–383: 170–178. [doi: [10.1016/j.ins.2016.12.007](https://doi.org/10.1016/j.ins.2016.12.007)]
- [23] Wang DH, Li M. Stochastic configuration networks: Fundamentals and algorithms. *IEEE Trans. on Cybernetics*, 2017, 47(10): 3466–3479. [doi: [10.1109/TCYB.2017.2734043](https://doi.org/10.1109/TCYB.2017.2734043)]
- [24] Wang DH, Li M. Deep stochastic configuration networks with universal approximation property. In: *Proc. of the 2018 Int'l Joint Conf. on Neural Networks*. Rio de Janeiro: IEEE, 2018. 1–8. [doi: [10.1109/IJCNN.2018.8489695](https://doi.org/10.1109/IJCNN.2018.8489695)]
- [25] Gao JB, Guo Y, Wang ZY. Matrix neural networks. In: *Proc. of the 14th Int'l Symp. on Advances in Neural Networks*. Sapporo: Springer, 2017. 313–320. [doi: [10.1007/978-3-319-59072-1_37](https://doi.org/10.1007/978-3-319-59072-1_37)]
- [26] Liu SM, Sun YF, Hu YL, Gao JB, Ju FJ, Yin BC. Matrix variate RBM model with Gaussian distributions. In: *Proc. of the 2017 Int'l Joint Conf. on Neural Networks*. Anchorage: IEEE, 2017. 808–815. [doi: [10.1109/IJCNN.2017.7965935](https://doi.org/10.1109/IJCNN.2017.7965935)]
- [27] Lu J, Zhao JW, Cao FL. Extended feed forward neural networks with random weights for face recognition. *Neurocomputing*, 2014, 136: 96–102. [doi: [10.1016/j.neucom.2014.01.022](https://doi.org/10.1016/j.neucom.2014.01.022)]
- [28] Li M, Wang DH. 2-D stochastic configuration networks for image data analytics. *IEEE Trans. on Cybernetics*, 2021, 51(1): 359–372. [doi: [10.1109/TCYB.2019.2925883](https://doi.org/10.1109/TCYB.2019.2925883)]
- [29] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv:1409.1556*, 2015.
- [30] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: *Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 770–778. [doi: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90)]
- [31] Szegedy C, Liu W, Jia YQ, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: *Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015. 1–9. [doi: [10.1109/cvpr.2015.7298594](https://doi.org/10.1109/cvpr.2015.7298594)]
- [32] Huang G, Liu Z, van der Maaten L, Weinberger K. Densely connected convolutional networks. In: *Proc. of the 2017 IEEE Conf. on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 2261–2269. [doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243)]
- [33] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: *Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 2818–2826. [doi: [10.1109/CVPR.2016.308](https://doi.org/10.1109/CVPR.2016.308)]
- [34] Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. Inception-v4, Inception-ResNet and the impact of residual connections on learning. In: *Proc. of the 31st AAAI Conf. on Artificial Intelligence*. San Francisco: AAAI Press, 2017. 4278–4284.
- [35] El-Melegy MT. Model-wise and point-wise random sample consensus for robust regression and outlier detection. *Neural Networks*, 2014, 59: 23–35. [doi: [10.1016/j.neunet.2014.06.010](https://doi.org/10.1016/j.neunet.2014.06.010)]
- [36] Gribonval R, Jenatton R, Bach F. Sparse and spurious: Dictionary learning with noise and outliers. *IEEE Trans. on Information Theory*, 2015, 61(11): 6298–6319. [doi: [10.1109/TIT.2015.2472522](https://doi.org/10.1109/TIT.2015.2472522)]
- [37] Dai W, Liu Q, Chai TY. Particle size estimate of grinding processes using random vector functional link networks with improved robustness. *Neurocomputing*, 2015, 169: 361–372. [doi: [10.1016/j.neucom.2014.08.098](https://doi.org/10.1016/j.neucom.2014.08.098)]

- [38] Wang DH, Li M. Robust stochastic configuration networks with kernel density estimation for uncertain data regression. *Information Sciences*, 2017, 412–413: 210–222. [doi: [10.1016/j.ins.2017.05.047](https://doi.org/10.1016/j.ins.2017.05.047)]
- [39] Xie J, Zhou P. Robust stochastic configuration network multi-output modeling of molten iron quality in blast furnace ironmaking. *Neurocomputing*, 2020, 387: 139–149. [doi: [10.1016/j.neucom.2020.01.030](https://doi.org/10.1016/j.neucom.2020.01.030)]
- [40] Li M, Huang CQ, Wang DH. Robust stochastic configuration networks with maximum correntropy criterion for uncertain data regression. *Information Sciences*, 2019, 473: 73–86. [doi: [10.1016/j.ins.2018.09.026](https://doi.org/10.1016/j.ins.2018.09.026)]
- [41] Dai W, Li DP, Chen QX, Chai TY. Data driven particle size estimation of hematite grinding process using stochastic configuration network with robust technique. *Journal of Central South University*, 2019, 26(1): 43–62. [doi: [10.1007/s11771-019-3981-2](https://doi.org/10.1007/s11771-019-3981-2)]
- [42] Wu RZ, Lv BY, Dai CM, Wang WG. Bayesian stochastic configuration networks for robust data modeling. *Concurrency and Computation: Practice and Experience*, 2022, 34(1): e6495. [doi: [10.1002/cpe.6495](https://doi.org/10.1002/cpe.6495)]
- [43] Lu J, Ding JL. Mixed-distribution-based robust stochastic configuration networks for prediction interval construction. *IEEE Trans. on Industrial Informatics*, 2020, 16(8): 5099–5109. [doi: [10.1109/TII.2019.2954351](https://doi.org/10.1109/TII.2019.2954351)]
- [44] Lu J, Ding JL, Liu CX, Chai TY. Hierarchical-Bayesian-based sparse stochastic configuration networks for construction of prediction intervals. *IEEE Trans. on Neural Networks and Learning Systems*, 2022, 33(8): 3560–3571. [doi: [10.1109/TNNLS.2021.3053306](https://doi.org/10.1109/TNNLS.2021.3053306)]
- [45] Wipf DP, Rao BD, Nagarajan S. Latent variable Bayesian models for promoting sparsity. *IEEE Trans. on Information Theory*, 2011, 57(9): 6236–6255. [doi: [10.1109/TIT.2011.2162174](https://doi.org/10.1109/TIT.2011.2162174)]
- [46] Dietterich TG. Ensemble methods in machine learning. In: *Proc. of the 1st Int'l Workshop on Multiple Classifier Systems*. Cagliari: Springer, 2000. 1–15. [doi: [10.1007/3-540-45014-9_1](https://doi.org/10.1007/3-540-45014-9_1)]
- [47] Chen TQ, Guestrin C. XGBoost: A scalable tree boosting system. In: *Proc. of the 22nd ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*. San Francisco: Association for Computing Machinery, 2016. 785–794. [doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785)]
- [48] Breiman L. Bagging predictors. *Machine Learning*, 1996, 24(2): 123–140. [doi: [10.1007/BF00058655](https://doi.org/10.1007/BF00058655)]
- [49] Wang YY, Wang DJ, Geng N, Wang YZ, Yin YQ, Jin YC. Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection. *Applied Soft Computing*, 2019, 77: 188–204. [doi: [10.1016/j.asoc.2019.01.015](https://doi.org/10.1016/j.asoc.2019.01.015)]
- [50] Lu J, Ding JL, Dai XW, Chai TY. Ensemble stochastic configuration networks for estimating prediction intervals: A simultaneous robust training algorithm and its application. *IEEE Trans. on Neural Networks and Learning Systems*, 2020, 31(12): 5426–5440. [doi: [10.1109/TNNLS.2020.2967816](https://doi.org/10.1109/TNNLS.2020.2967816)]
- [51] Li K, Wang W, Wang YP. Application of ensemble stochastic configuration network in aquaculture water quality monitoring. *Trans. of the Chinese Society of Agricultural Engineering*, 2020, 36(4): 220–226 (in Chinese with English abstract). [doi: [10.11975/j.issn.1002-6819.2020.04.026](https://doi.org/10.11975/j.issn.1002-6819.2020.04.026)]
- [52] Qu HQ, Feng TL, Wang YP, Zhang Y. AdaBoost-SCN algorithm for optical fiber vibration signal recognition. *Applied Optics*, 2019, 58(21): 5612–5623. [doi: [10.1364/AO.58.005612](https://doi.org/10.1364/AO.58.005612)]
- [53] Qu HQ, Feng TL, Zhang Y, Wang YP. Ensemble learning with stochastic configuration network for noisy optical fiber vibration signal recognition. *Sensors*, 2019, 19(15): 3293. [doi: [10.3390/s19153293](https://doi.org/10.3390/s19153293)]
- [54] Pratama M, Wang DH. Deep stacked stochastic configuration networks for lifelong learning of non-stationary data streams. *Information Sciences*, 2019, 495: 150–174. [doi: [10.1016/j.ins.2019.04.055](https://doi.org/10.1016/j.ins.2019.04.055)]
- [55] Wang DH, Cui CH. Stochastic configuration networks ensemble with heterogeneous features for large-scale data analytics. *Information Sciences*, 2017, 417: 55–71. [doi: [10.1016/j.ins.2017.07.003](https://doi.org/10.1016/j.ins.2017.07.003)]
- [56] Geman S, Bienenstock E, Doursat R. Neural networks and the bias/variance dilemma. *Neural Computation*, 1992, 4(1): 1–58. [doi: [10.1162/neco.1992.4.1.1](https://doi.org/10.1162/neco.1992.4.1.1)]
- [57] Hansen LK, Salamon P. Neural network ensembles. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1990, 12(10): 993–1001. [doi: [10.1109/34.58871](https://doi.org/10.1109/34.58871)]
- [58] Huang CQ, Li M, Wang DH. Stochastic configuration network ensembles with selective base models. *Neural Networks*, 2021, 137: 106–118. [doi: [10.1016/j.neunet.2021.01.011](https://doi.org/10.1016/j.neunet.2021.01.011)]
- [59] Lu YM, Roychowdhury V, Vandenberghe L. Distributed parallel support vector machines in strongly connected networks. *IEEE Trans. on Neural Networks*, 2008, 19(7): 1167–1178. [doi: [10.1109/TNN.2007.2000061](https://doi.org/10.1109/TNN.2007.2000061)]
- [60] Ai W, Wang DH. Distributed stochastic configuration networks with cooperative learning paradigm. *Information Sciences*, 2020, 540: 1–16. [doi: [10.1016/j.ins.2020.05.112](https://doi.org/10.1016/j.ins.2020.05.112)]
- [61] Ge XW, Ai W, Zhou YJ. Distributed event-triggered learning for stochastic configuration networks. In: *Proc. of the 2020 Chinese Automation Congress*. Shanghai: IEEE, 2020. 3384–3388. [doi: [10.1109/CAC51589.2020.9326565](https://doi.org/10.1109/CAC51589.2020.9326565)]
- [62] Zhao LJ, Zou SD, Huang MZ, Wang GG. Distributed regularized stochastic configuration networks via the elastic net. *Neural*

- Computing and Applications, 2021, 33(8): 3281–3297. [doi: [10.1007/s00521-020-05178-x](https://doi.org/10.1007/s00521-020-05178-x)]
- [63] Zhou YJ, Ai W. ADMM-based sparse distributed learning for stochastic configuration networks. In: Proc. of the 2020 Chinese Automation Congress. Shanghai: IEEE, 2020. 4354–4358. [doi: [10.1109/CAC51589.2020.9326861](https://doi.org/10.1109/CAC51589.2020.9326861)]
- [64] Dai W, Zhou XY, Li DP, Zhu S, Wang XS. Hybrid parallel stochastic configuration networks for industrial data analytics. IEEE Trans. on Industrial Informatics, 2022, 18(4): 2331–2341. [doi: [10.1109/TII.2021.3096840](https://doi.org/10.1109/TII.2021.3096840)]
- [65] Zhang CL, Ding SF, Zhang J, Jia WK. Parallel stochastic configuration networks for large-scale data regression. Applied Soft Computing, 2021, 103: 107143. [doi: [10.1016/j.asoc.2021.107143](https://doi.org/10.1016/j.asoc.2021.107143)]
- [66] Wang QJ, Dai W, Lu Q, Fu XR, Ma XP. A sparse learning method for SCN soft measurement model. Control and Decision, 2022, 37(12): 3171–3182 (in Chinese with English abstract). [doi: [10.13195/j.kzyjc.2021.1047](https://doi.org/10.13195/j.kzyjc.2021.1047)]
- [67] Zhao LJ, Zou SD, Guo S, Huang MZ. Ball mill load condition recognition model based on regularized stochastic configuration networks. Control Engineering of China, 2020, 27(1): 1–7 (in Chinese with English abstract). [doi: [10.14107/j.cnki.kzgc.20190437](https://doi.org/10.14107/j.cnki.kzgc.20190437)]
- [68] Wang QJ, Yang CY, Ma XP, Zhang CF, Peng SM. Underground airflow quantity modeling based on SCN. Acta Automatica Sinica, 2021, 47(8): 1963–1975 (in Chinese with English abstract). [doi: [10.16383/j.aas.c190602](https://doi.org/10.16383/j.aas.c190602)]
- [69] Li W, Zeng ZQ, Qu HQ, Sun CB. A novel fiber intrusion signal recognition method for OFPS based on SCN with dropout. Journal of Lightwave Technology, 2019, 37(20): 5221–5230. [doi: [10.1109/JLT.2019.2930624](https://doi.org/10.1109/JLT.2019.2930624)]
- [70] Pan CY, Xu JX, Weng YP. A fault identification method of chemical process based on manifold regularized stochastic configuration network. Chinese Journal of Scientific Instrument, 2021, 42(5): 219–226 (in Chinese with English abstract). [doi: [10.19650/j.cnki.cjsi.J2107541](https://doi.org/10.19650/j.cnki.cjsi.J2107541)]
- [71] Pan JQ, Luan F, Gao YQ, Wei YJ. FPGA-based implementation of stochastic configuration network for robotic grasping recognition. IEEE Access, 2020, 8: 139966–139973. [doi: [10.1109/ACCESS.2020.3012819](https://doi.org/10.1109/ACCESS.2020.3012819)]
- [72] Li WT, Tao H, Li H, Chen KQ, Wang JP. Greengage grading using stochastic configuration networks and a semi-supervised feedback mechanism. Information Sciences, 2019, 488: 1–12. [doi: [10.1016/j.ins.2019.02.041](https://doi.org/10.1016/j.ins.2019.02.041)]
- [73] Zhang Q, Li WT, Li H, Wang JP. Self-blast state detection of glass insulators based on stochastic configuration networks and a feedback transfer learning mechanism. Information Sciences, 2020, 522: 259–274. [doi: [10.1016/j.ins.2020.02.058](https://doi.org/10.1016/j.ins.2020.02.058)]
- [74] Liu W, Ren C, Xu Y. PV generation forecasting with missing input data: A super-resolution perception approach. IEEE Trans. on Sustainable Energy, 2021, 12(2): 1493–1496. [doi: [10.1109/TSTE.2020.3029731](https://doi.org/10.1109/TSTE.2020.3029731)]
- [75] Sheng ZY, Zeng ZQ, Qu HQ, Zhang Y. Optical fiber intrusion signal recognition method based on TSVD-SCN. Optical Fiber Technology, 2019, 48: 270–277. [doi: [10.1016/j.yofte.2019.01.023](https://doi.org/10.1016/j.yofte.2019.01.023)]
- [76] Dai W, Li DP, Zhou P, Chai TY. Stochastic configuration networks with block increments for data modeling in process industries. Information Sciences, 2019, 484: 367–386. [doi: [10.1016/j.ins.2019.01.062](https://doi.org/10.1016/j.ins.2019.01.062)]
- [77] Tian Q, Yuan SJ, Qu HQ. Intrusion signal classification using stochastic configuration network with variable increments of hidden nodes. Optical Engineering, 2019, 58(2): 026105. [doi: [10.1117/1.OE.58.2.026105](https://doi.org/10.1117/1.OE.58.2.026105)]
- [78] Cao WP, Xie ZW, Li JQ, Xu ZW, Ming Z, Wang XZ. Bidirectional stochastic configuration network for regression problems. Neural Networks, 2021, 140: 237–246. [doi: [10.1016/j.neunet.2021.03.016](https://doi.org/10.1016/j.neunet.2021.03.016)]
- [79] Wang QJ, Dai W, Ma XP, Shang ZG. Driving amount based stochastic configuration network for industrial process modeling. Neurocomputing, 2020, 394: 61–69. [doi: [10.1016/j.neucom.2020.02.029](https://doi.org/10.1016/j.neucom.2020.02.029)]
- [80] Li K, Wang W, Lin SH. Soft measurement of ammonia nitrogen concentration based on GA-SCN. In: Proc. of the 2018 IEEE Symp. on Product Compliance Engineering-Asia (ISPCE-CN). Shenzhen: IEEE, 2018. 1–4. [doi: [10.1109/ISPCE-CN.2018.8805767](https://doi.org/10.1109/ISPCE-CN.2018.8805767)]
- [81] Zhang CL, Ding SF. A stochastic configuration network based on chaotic sparrow search algorithm. Knowledge-based Systems, 2021, 220: 106924. [doi: [10.1016/j.knosys.2021.106924](https://doi.org/10.1016/j.knosys.2021.106924)]
- [82] Zhu XL, Feng XC, Wang WW, Jia XX, He RQ. A further study on the inequality constraints in stochastic configuration networks. Information Sciences, 2019, 487: 77–83. [doi: [10.1016/j.ins.2019.02.066](https://doi.org/10.1016/j.ins.2019.02.066)]
- [83] Niu HY, Wei JM, Chen YQ. Optimal randomness for stochastic configuration network (SCN) with heavy-tailed distributions. Entropy, 2021, 23(1): 56. [doi: [10.3390/e23010056](https://doi.org/10.3390/e23010056)]
- [84] Zhao LJ, Wang J, Huang MZ, Wang GG. Estimation of effluent quality index based on partial least squares stochastic configuration networks. CIESC Journal, 2020, 71(12): 5672–5680 (in Chinese with English abstract). [doi: [10.11949/0438-1157.20200604](https://doi.org/10.11949/0438-1157.20200604)]
- [85] Wang W, Li K, Guo G. Seawater ammonia nitrogen concentration modelling via RS-SCN. In: Proc. of the 2019 Chinese Automation Congress. Hangzhou: IEEE, 2019. 2008–2011. [doi: [10.1109/CAC48633.2019.8996654](https://doi.org/10.1109/CAC48633.2019.8996654)]
- [86] Wang W, Jia Y, Yu W, Pang HS, Cai KW. On-line ammonia nitrogen measurement using generalized additive model and stochastic configuration networks. Measurement, 2021, 170: 108743. [doi: [10.1016/j.measurement.2020.108743](https://doi.org/10.1016/j.measurement.2020.108743)]

- [87] Felicetti MJ, Wang DH. Deep stochastic configuration networks with optimised model and hyper-parameters. *Information Sciences*, 2022, 600: 431–441. [doi: [10.1016/j.ins.2022.04.013](https://doi.org/10.1016/j.ins.2022.04.013)]
- [88] Alçın M, Pehlivan İ, Koyuncu İ. Hardware design and implementation of a novel ANN-based chaotic generator in FPGA. *Optik*, 2016, 127(13): 5500–5505. [doi: [10.1016/j.ijleo.2016.03.042](https://doi.org/10.1016/j.ijleo.2016.03.042)]
- [89] Sahin S, Becerikli Y, Yazici S. Neural network implementation in hardware using FPGAs. In: *Proc. of the 13th Int'l Conf. on Neural Information Processing*. Hong Kong: Springer, 2006. 1105–1112. [doi: [10.1007/11893295_122](https://doi.org/10.1007/11893295_122)]
- [90] Lei P, Liang JW, Guan ZY, Wang J, Zheng T. Acceleration of FPGA based convolutional neural network for human activity classification using millimeter-wave radar. *IEEE Access*, 2019, 7: 88917–88926. [doi: [10.1109/ACCESS.2019.2926381](https://doi.org/10.1109/ACCESS.2019.2926381)]
- [91] Liu Q, Liang T, Huang Z, Dinavahi V. Real-time FPGA-based hardware neural network for fault detection and isolation in more electric aircraft. *IEEE Access*, 2019, 7: 159831–159841. [doi: [10.1109/ACCESS.2019.2950918](https://doi.org/10.1109/ACCESS.2019.2950918)]
- [92] Gao YQ, Luan F, Pan JQ, Li X, He YD. FPGA-based implementation of stochastic configuration networks for regression prediction. *Sensors*, 2020, 20(15): 4191. [doi: [10.3390/s20154191](https://doi.org/10.3390/s20154191)]
- [93] Wang YP, Zhang YB, Zhang Y, Fan J, Qu HQ. Stochastic configuration network-based SAR image target classification approach. *The Journal of Engineering*, 2019, 2019(21): 8121–8124. [doi: [10.1049/joe.2019.0683](https://doi.org/10.1049/joe.2019.0683)]
- [94] Niu HY, Wang D, Chen YQ. Estimating actual crop evapotranspiration using deep stochastic configuration networks model and UAV-based crop coefficients in a pomegranate orchard. In: *Proc. of the 2020 Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping V*. SPIE, 2020. 76–82. [doi: [10.1117/12.2558221](https://doi.org/10.1117/12.2558221)]
- [95] Niu HY, Wang D, Chen YQ. Estimating crop coefficients using linear and deep stochastic configuration networks models and UAV-based normalized difference vegetation index (NDVI). In: *Proc. of the 2020 Int'l Conf. on Unmanned Aircraft Systems*. Athens: IEEE, 2020. 1485–1490. [doi: [10.1109/ICUAS48674.2020.9213888](https://doi.org/10.1109/ICUAS48674.2020.9213888)]
- [96] Agarwal SK, Shah S, Kumar R. Classification of mental tasks from EEG data using backtracking search optimization based neural classifier. *Neurocomputing*, 2015, 166: 397–403. [doi: [10.1016/j.neucom.2015.03.041](https://doi.org/10.1016/j.neucom.2015.03.041)]
- [97] Pang LP, Guo L, Zhang J, Wanyan XR, Qu HQ, Wang X. Subject-specific mental workload classification using EEG and stochastic configuration network (SCN). *Biomedical Signal Processing and Control*, 2021, 68: 102711. [doi: [10.1016/j.bspc.2021.102711](https://doi.org/10.1016/j.bspc.2021.102711)]
- [98] Wang YF, Guo HF, Li SY, Wang LD, Song X, Zhao XK. Identify risk factors and predict the postoperative risk of ESCC using ensemble learning. *Biomedical Signal Processing and Control*, 2021, 68: 102784. [doi: [10.1016/j.bspc.2021.102784](https://doi.org/10.1016/j.bspc.2021.102784)]
- [99] Deutsch J, He D. Using deep learning-based approach to predict remaining useful life of rotating components. *IEEE Trans. on Systems, Man, and Cybernetics: Systems*, 2018, 48(1): 11–20. [doi: [10.1109/TSMC.2017.2697842](https://doi.org/10.1109/TSMC.2017.2697842)]
- [100] Li YB, Xu MQ, Wei Y, Huang WH. A new rolling bearing fault diagnosis method based on multiscale permutation entropy and improved support vector machine based binary tree. *Measurement*, 2016, 77: 80–94. [doi: [10.1016/j.measurement.2015.08.034](https://doi.org/10.1016/j.measurement.2015.08.034)]
- [101] Pang LP, Zhang J, Liu M, Qu HQ, Wang J. Thermal models for avionics pod cabin based on stochastic configuration network (SCN). *Energy and Built Environment*, 2020, 1(1): 93–105. [doi: [10.1016/j.enbenv.2019.10.001](https://doi.org/10.1016/j.enbenv.2019.10.001)]
- [102] Huang DR, Li YQ, Guan SY, Zhang X, Tang M. A novel collaborative diagnosis approach of incipient faults based on VMD and SCN for rolling bearing. *Optimal Control Applications and Methods*, 2021: 1–15. [doi: [10.1002/oca.2820](https://doi.org/10.1002/oca.2820)]
- [103] Liu JN, Hao RJ, Zhang TL, Wang XZ. Vibration fault diagnosis based on stochastic configuration neural networks. *Neurocomputing*, 2021, 434: 98–125. [doi: [10.1016/j.neucom.2020.12.080](https://doi.org/10.1016/j.neucom.2020.12.080)]
- [104] Chen Y, Jiang T, Liu HL, Li YH, Yu ZH. Event recognition system based on fiber bragg grating and mRMR-CWCs-SCN. *IEEE Sensors Journal*, 2021, 21(22): 26132–26139. [doi: [10.1109/JSEN.2021.3119993](https://doi.org/10.1109/JSEN.2021.3119993)]
- [105] Ding CX, Yan AJ. Fault detection in the MSW incineration process using stochastic configuration networks and case-based reasoning. *Sensors*, 2021, 21(21): 7356. [doi: [10.3390/s21217356](https://doi.org/10.3390/s21217356)]
- [106] Wang W, Guo G. Soft measurement model for ammonia nitrogen concentration in marine aquaculture based on stochastic configuration networks. *Trans. of the Chinese Society for Agricultural Machinery*, 2020, 51(1): 214–220 (in Chinese with English abstract). [doi: [10.6041/j.issn.1000-1298.2020.01.023](https://doi.org/10.6041/j.issn.1000-1298.2020.01.023)]
- [107] Lu J, Ding JL. Construction of prediction intervals for carbon residual of crude oil based on deep stochastic configuration networks. *Information Sciences*, 2019, 486: 119–132. [doi: [10.1016/j.ins.2019.02.042](https://doi.org/10.1016/j.ins.2019.02.042)]
- [108] Tao JM, Niu H, Zhang YJ, Li XS. An intelligent modeling method for nonlinear systems based on random configuration networks. *Control and Decision*, 2022, 37(10): 2559–2564 (in Chinese with English abstract). [doi: [10.13195/j.kzyjc.2021.0436](https://doi.org/10.13195/j.kzyjc.2021.0436)]
- [109] Lian BS, Zhang QL, Li JN. Integrated sliding mode control and neural networks based packet disordering prediction for nonlinear networked control systems. *IEEE Trans. on Neural Networks and Learning Systems*, 2019, 30(8): 2324–2335. [doi: [10.1109/TNNLS.2018.2873183](https://doi.org/10.1109/TNNLS.2018.2873183)]

- [110] Wang YF, Wang MX, Wang DH, Chang YL. Stochastic configuration network based cascade generalized predictive control of main steam temperature in power plants. *Information Sciences*, 2022, 587: 123–141. [doi: [10.1016/j.ins.2021.12.006](https://doi.org/10.1016/j.ins.2021.12.006)]
- [111] Bi J, Yuan HT, Zhang LB, Zhang J. SGW-SCN: An integrated machine learning approach for workload forecasting in geo-distributed cloud data centers. *Information Sciences*, 2019, 481: 57–68. [doi: [10.1016/j.ins.2018.12.027](https://doi.org/10.1016/j.ins.2018.12.027)]
- [112] Bi J, Yuan HT, Zhou MC, Liu Q. Time-dependent cloud workload forecasting via multi-task learning. *IEEE Robotics and Automation Letters*, 2019, 4(3): 2401–2406. [doi: [10.1109/LRA.2019.2899224](https://doi.org/10.1109/LRA.2019.2899224)]
- [113] Huang CQ, Huang QH, Wang DH. Stochastic configuration networks based adaptive storage replica management for power big data processing. *IEEE Trans. on Industrial Informatics*, 2020, 16(1): 373–383. [doi: [10.1109/TII.2019.2919268](https://doi.org/10.1109/TII.2019.2919268)]
- [114] Cai KW, Alalibo BP, Cao WP, Liu Z, Wang ZQ, Li GF. Hybrid approach for detecting and classifying power quality disturbances based on the variational mode decomposition and deep stochastic configuration network. *Energies*, 2018, 11(11): 3040. [doi: [10.3390/en11113040](https://doi.org/10.3390/en11113040)]
- [115] Zhao YW, Li P, Sun YH, Shen X, Yang XH. Short-term power load forecasting based on phase space reconstruction and stochastic configuration networks. *Electric Power Construction*, 2021, 42(9): 120–128 (in Chinese with English abstract). [doi: [10.12204/j.issn.1000-7229.2021.09.013](https://doi.org/10.12204/j.issn.1000-7229.2021.09.013)]
- [116] Hao J, Zhang GS, Yang KH, Wu MZ, Zheng YQ, Hu W. Online unified solution for selective harmonic elimination based on stochastic configuration network and Levenberg-Marquardt algorithm. *IEEE Trans. on Industrial Electronics*, 2022, 69(10): 10724–10734. [doi: [10.1109/TIE.2021.3137591](https://doi.org/10.1109/TIE.2021.3137591)]
- [117] Wang W, Wang DH. Prediction of component concentrations in sodium aluminate liquor using stochastic configuration networks. *Neural Computing and Applications*, 2020, 32(17): 13625–13638. [doi: [10.1007/s00521-020-04771-4](https://doi.org/10.1007/s00521-020-04771-4)]
- [118] Han Y, Song XP, Li K, Yan XX. Hybrid modeling for submergence depth of the pumping well using stochastic configuration networks with random sampling. *Journal of Petroleum Science and Engineering*, 2022, 208: 109423. [doi: [10.1016/j.petrol.2021.109423](https://doi.org/10.1016/j.petrol.2021.109423)]
- [119] Zhao JH, Hu TY, Zheng RF, Ba PH, Mei CL, Zhang QC. Defect recognition in concrete ultrasonic detection based on wavelet packet transform and stochastic configuration networks. *IEEE Access*, 2021, 9: 9284–9295. [doi: [10.1109/ACCESS.2021.3049448](https://doi.org/10.1109/ACCESS.2021.3049448)]
- [120] Tian ZD, Chen H. Multi-step short-term wind speed prediction based on integrated multi-model fusion. *Applied Energy*, 2021, 298: 117248. [doi: [10.1016/j.apenergy.2021.117248](https://doi.org/10.1016/j.apenergy.2021.117248)]

附中文参考文献:

- [1] 丁世飞, 徐晓, 王艳茹. 基于不相似性度量优化的密度峰值聚类算法. *软件学报*, 2020, 31(11): 3321–3333. <http://www.jos.org.cn/1000-9825/5813.htm> [doi: [10.13328/j.cnki.jos.005813](https://doi.org/10.13328/j.cnki.jos.005813)]
- [2] 张健, 丁世飞, 丁玲, 张成龙. 基于实值RBM的深度生成网络研究. *软件学报*, 2021, 32(12): 3802–3813. <http://www.jos.org.cn/1000-9825/6126.htm> [doi: [10.13328/j.cnki.jos.006126](https://doi.org/10.13328/j.cnki.jos.006126)]
- [51] 李康, 王魏, 王奕鹏. 集成随机配置网络在养殖水质监测中的应用. *农业工程学报*, 2020, 36(4): 220–226. [doi: [10.11975/j.issn.1002-6819.2020.04.026](https://doi.org/10.11975/j.issn.1002-6819.2020.04.026)]
- [66] 王前进, 代伟, 陆群, 辅小荣, 马小平. 一种随机配置网络软测量模型的稀疏学习方法. *控制与决策*, 2022, 37(12): 3171–3182. [doi: [10.13195/j.kzyjc.2021.1047](https://doi.org/10.13195/j.kzyjc.2021.1047)]
- [67] 赵立杰, 邹世达, 郭烁, 黄明忠. 基于正则化随机配置网络的球磨机工况识别. *控制工程*, 2020, 27(1): 1–7. [doi: [10.14107/j.cnki.kzgc.20190437](https://doi.org/10.14107/j.cnki.kzgc.20190437)]
- [68] 王前进, 杨春雨, 马小平, 张春富, 彭思敏. 基于随机配置网络的井下供给风量建模. *自动化学报*, 2021, 47(8): 1963–1975. [doi: [10.16383/j.aas.c190602](https://doi.org/10.16383/j.aas.c190602)]
- [70] 潘承燕, 徐进学, 翁永鹏. 一种基于流形正则化随机配置网络的化工过程故障识别方法. *仪器仪表学报*, 2021, 42(5): 219–226. [doi: [10.19650/j.cnki.cjsi.J2107541](https://doi.org/10.19650/j.cnki.cjsi.J2107541)]
- [84] 赵立杰, 王佳, 黄明忠, 王国刚. 基于偏最小二乘随机配置网络的污水水质指标估计. *化工学报*, 2020, 71(12): 5672–5680. [doi: [10.11949/0438-1157.20200604](https://doi.org/10.11949/0438-1157.20200604)]
- [106] 王魏, 郭戈. 基于随机配置网络的海水养殖氨氮浓度软测量模型. *农业机械学报*, 2020, 51(1): 214–220. [doi: [10.6041/j.issn.1000-1298.2020.01.023](https://doi.org/10.6041/j.issn.1000-1298.2020.01.023)]
- [108] 陶金梅, 牛宏, 张亚军, 李旭生. 基于随机配置网络的非线性系统智能建模方法. *控制与决策*, 2022, 37(10): 2559–2564. [doi: [10.13195/j.kzyjc.2021.0436](https://doi.org/10.13195/j.kzyjc.2021.0436)]
- [115] 赵允文, 李鹏, 孙煜皓, 沈鑫, 杨晓华. 基于相空间重构和随机配置网络的电力负荷短期预测. *电力建设*, 2021, 42(9): 120–128. [doi: [10.12204/j.issn.1000-7229.2021.09.013](https://doi.org/10.12204/j.issn.1000-7229.2021.09.013)]



张成龙(1992—), 男, 博士生, CCF 学生会会员, 主要研究领域为随机配置网络, 智能优化算法.



郭丽丽(1990—), 女, 博士, 讲师, CCF 专业会员, 主要研究领域为深度学习, 情感计算, 语音情感识别, 多模态情感识别.



丁世飞(1963—), 男, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究领域为人工智能与模式识别, 机器学习, 数据挖掘, 大数据智能分析.



张健(1990—), 男, 博士, 讲师, CCF 专业会员, 主要研究领域为机器学习, 深度学习.