



计算机科学

COMPUTER SCIENCE

演化循环神经网络研究综述

胡中源, 薛羽, 查加杰

引用本文

胡中源, 薛羽, 查加杰. 演化循环神经网络研究综述[J]. 计算机科学, 2023, 50(3): 254-265.

HU Zhongyuan, XUE Yu, ZHA Jiajie. Survey on Evolutionary Recurrent Neural Networks[J]. Computer Science, 2023, 50(3): 254-265.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于迁移学习和多视图特征融合提高RNA碱基相互作用预测](#)

Improving RNA Base Interactions Prediction Based on Transfer Learning and Multi-view Feature Fusion

计算机科学, 2023, 50(3): 164-172. <https://doi.org/10.11896/jsjcx.211200186>

[面向频谱接入深度强化学习模型的后门攻击方法](#)

Backdoor Attack Against Deep Reinforcement Learning-based Spectrum Access Model

计算机科学, 2023, 50(1): 351-361. <https://doi.org/10.11896/jsjcx.220800269>

[基于云平台日志的故障检测和复杂构件系统即时可靠性度量研究](#)

Study on Anomaly Detection and Real-time Reliability Evaluation of Complex Component System Based on Log of Cloud Platform

计算机科学, 2022, 49(12): 125-135. <https://doi.org/10.11896/jsjcx.220200106>

[基于机器学习的剩余使用寿命预测实证研究](#)

Empirical Research on Remaining Useful Life Prediction Based on Machine Learning

计算机科学, 2022, 49(11A): 211100285-9. <https://doi.org/10.11896/jsjcx.211100285>

[基于少样本的太阳射电爆发事件检测研究](#)

Study on Solar Radio Burst Event Detection Based on Transfer Learning

计算机科学, 2022, 49(11A): 210900198-7. <https://doi.org/10.11896/jsjcx.210900198>

演化循环神经网络研究综述

胡中源 薛羽 查加杰

南京信息工程大学软件学院 南京 210044

(20211220009@nuist.edu.cn)

摘要 演化计算利用生物演化过程中的自然选择机制和遗传规律求解优化问题,循环神经网络的精度和效率依赖其参数以及结构的优化效果,采用演化计算解决循环神经网络中的参数与结构自适应优化问题是自动化深度学习领域的研究热点。文中针对结合演化计算和循环神经网络的算法进行了详细的调研。首先,简要介绍了演化算法的传统类别、常见算法和优点,以及循环神经网络模型的结构及特点,并对影响循环神经网络性能的因素进行了分析;其次,分析了演化循环神经网络的算法框架,并分别从权重优化、超参数优化和结构优化方面分析了当前演化循环神经网络的研究进展;然后,对演化循环神经网络的一些其他工作进行了分析;最后,指出了演化循环神经网络面临的挑战以及发展趋势。

关键词: 循环神经网络;演化计算;权重优化;超参数优化;结构优化;集成学习;迁移学习

中图分类号 TP183

Survey on Evolutionary Recurrent Neural Networks

HU Zhongyuan, XUE Yu and ZHA Jiajie

School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China

Abstract Evolutionary computation utilizes natural selection mechanisms and genetic laws in the process of biological evolution to solve optimization problems. The accuracy and efficiency of the evolutionary recurrent neural network model depends on the optimization effect of parameters and the structures. The utilization of evolutionary computation to solve the problem of adaptive optimization of parameters and structures in recurrent neural networks is a hot spot of automated deep learning. This paper summarizes the algorithms that combine evolutionary algorithms and recurrent neural networks. Firstly, it briefly reviews the traditional categories, common algorithms, and advantages of evolutionary computation. Next, it briefly introduces the structures and characteristics of the recurrent neural network models and analyzes the influencing factors of recurrent neural network performance. Then, it analyzes the algorithmic framework of evolutionary recurrent neural networks, and the current research development of evolutionary recurrent neural networks from weight optimization, hyperparameter optimization and structure optimization. Besides, other work on evolutionary recurrent neural networks is analyzed. Finally, it points out the challenges and the development trend of evolutionary recurrent neural networks.

Keywords Recurrent neural network, Evolutionary computation, Weight optimization, Hyperparameter optimization, Optimization of structure, Ensemble learning, Transfer learning

1 引言

物联网、大数据等新技术、新应用的不断涌现,促进了深度学习的创新发展。深度学习是由多层神经网络发展而来的,是大数据处理和主流方法^[1-2]。深度学习的概念

起源于人工神经网络的研究,其通过提取底层特征形成更抽象的高层属性类别或特征表示机制,从而发现数据的分布式特征表示^[3]。循环神经网络是深度学习的基础模块,是一种反映特征映射关系的网络模型,用于学习具有时序性的输入与具有时序性的输出之间的对应关系和规则。循环神经网络

到稿日期:2022-05-31 返修日期:2022-09-27

基金项目:国家自然科学基金(61876089);数据科学与智慧软件江苏省重点实验室开放课题基金(2019DS302);江苏省自然科学基金(BK20141005);江苏省高校自然科学研究项目(14KJB520025);江苏省研究生科研与实践创新计划(KYCX22_1206)

This work was supported by the National Natural Science Foundation of China(61876089), Opening Project of Jiangsu Key Laboratory of Data Science and Smart Software(2019DS302), Natural Science Foundation of Jiangsu Province(BK20141005), Natural Science Foundation of the Jiangsu Higher Education Institutions of China(14KJB520025) and Postgraduate Research & Practice Innovation Program of Jiangsu Province(KYCX22_1206).

通信作者:薛羽(xueyu@nuist.edu.cn)

(Recurrent Neural Network, RNN)在解决时间序列预测问题上具有优势,其最大的特点就是神经元在某时刻的输出可以作为输入再次输入到神经元。这种串联的网络结构对具有时序性的数据有很强的学习能力,能够提取出这类数据的前后次序所反映出的信息^[4]。

神经网络模型的性能受诸多因素影响,如基本神经元的结构、网络深度和权重等。这些因素的调整都有各自的难点,通过人工去调节这些因素往往需要大量的专业知识,且需要专业人员的多次试错来逐步提升模型性能。

神经网络因其在时间序列预测问题上的优异表现,而被广泛应用于各类时序问题,如时间序列预测^[5]、自然语言处理^[6]、机器翻译^[7]和图像标注^[8]等。其中,神经网络的(超)参数和结构的优化是应用神经网络处理复杂任务时的重要研究对象。1)如何高效地训练神经网络的权重是一个重要研究课题,目前神经网络一般使用时序反向传播算法(Back Propagation Through Time, BPTT)^[9]进行训练,因此存在陷入局部最优解的问题^[10]和梯度消失及爆炸的问题^[11]。2)如何自适应地优化神经网络的超参数是另一个重要问题,机器学习模型的超参数能够根据具体任务自适应地调整是自动化机器学习的一项基本任务^[12-13]。3)如何优化神经网络的结构(如神经元的结构等)是亟需研究的问题^[14-16],传统神经网络结构的设计基于人类经验,对现有网络结构的设计过程缺乏可解释性。由于结构的潜在空间是无限的,因此试错法得到的结构往往不是所有潜在可能的结构中最好的结构。4)如何将演化神经网络与其他机器学习知识相结合也是一个重要的问题。

综上所述,神经网络在参数或结构的优化中面临着诸多的问题。演化神经网络是将演化计算的方法融入到神经网络的参数优化或者网络框架的优化设计中,利用演化计算所擅长的领域来解决上述问题。本文首先简要介绍了演化算法的传统类别、常见算法和优点,并介绍了神经网络模型的结构、特点和性能的影响因素;其次,分析了演化神经网络的算法框架和当前演化神经网络的主要研究进展;然后,分析了演化神经网络的一些其他工作;最后,指出了演化神经网络的发展趋势以及面临的挑战。

2 基础知识

2.1 演化计算

演化计算的灵感来源于生物演化中的“优胜劣汰、适者生存”的自然法则和遗传基因的传递规律。传统的演化计算包括遗传算法、遗传规划、演化策略和演化规划^[17-18],演化计算的一般流程如图1所示。其中,遗传算法较为成熟并已得到广泛应用^[19]。

广义的演化计算还包括差分演化算法^[20]、粒子群算法^[21]和文化基因算法^[22]等。多目标演化算法^[23]常用于解决较复杂的多目标优化问题。

演化计算将随机搜索与并行式的加强近邻搜索有效结合,具有通用性、并行性和全局最优性等特点。再者,演化

计算采用自然演化机制表现复杂的现象,能够快速有效地解决复杂问题^[24]。演化计算已在自动化深度学习领域得到了应用^[25]。

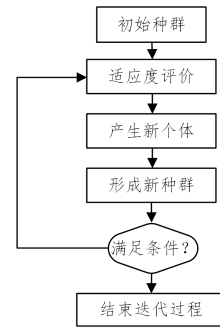


图1 演化算法的一般流程

Fig. 1 General flow of evolutionary algorithms

2.2 神经网络

神经网络主要用于解决时间序列问题。在神经网络的前向传播中,前一时刻的隐藏层的输出值可以作为下一时刻隐藏层的一个输入。这表明隐藏层的输出同时取决于当前时刻的输入值和经过前一时刻的隐藏层处理而得到的值^[26]。

如图2所示^[27],根据具体问题的不同,神经网络框架的隐藏层单元数量、输入数量和输出数量也会有相应的调整。同一目标任务下,这些数量的调整对模型的精度和收敛速度也会有相应的影响。如目标任务适应的结构是多输入对应单输出,其中一种代替方案如下:先采用多输入对应多输出的方式,接着在多输出的结果上施加注意力机制,最终得到单输出,这种替换方式能在一定程度上提升模型的精度^[28]。如图3所示,结构中输出的数量和位置是可进行优化调整的超参数。

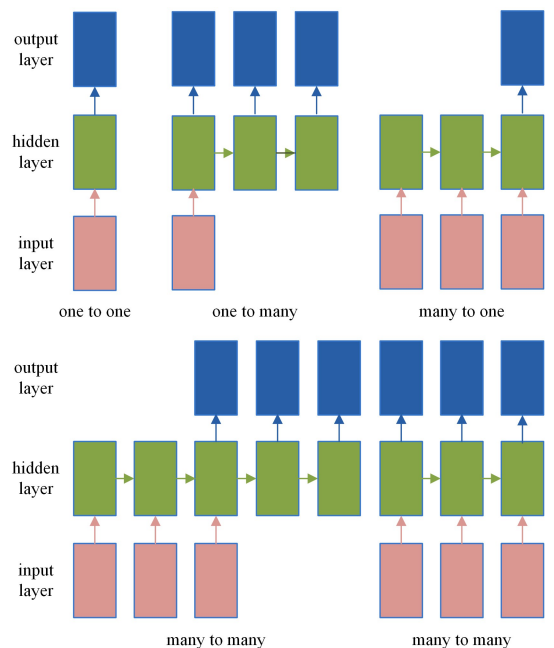


图2 神经网络的常见结构

Fig. 2 Common structure of recurrent neural networks

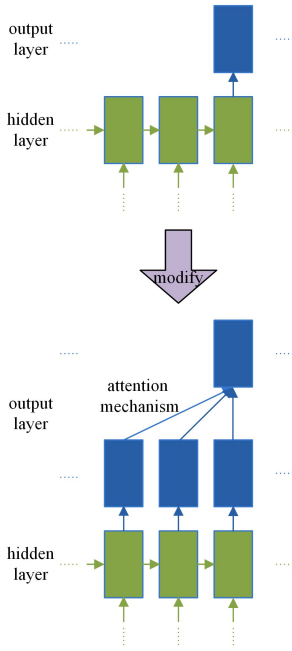


图3 一种循环神经网络结构调整方式

Fig. 3 Method of restructuring recurrent neural networks

循环神经网络的神经元结构对循环神经网络的性能有极其重要的影响。图4给出了一种简单的循环神经网络单元结构,这种循环神经网络在训练过程中容易出现梯度爆炸和梯度消失的问题,导致在训练时梯度的传递性不高,即梯度不能在较长序列中传递,使得这种循环神经网络的长期记忆能力较差^[11]。

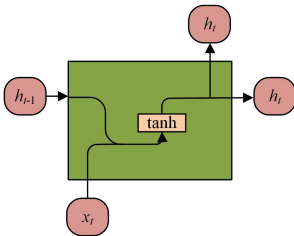


图4 简单循环神经网络单元的结构

Fig. 4 Structure of simple recurrent neural networks units

图4中单元的计算式如下:

$$h_t = \tanh(w_x x_t + w_h h_{t-1} + b) \quad (1)$$

其中, w_x 和 w_h 分别为对应的权重矩阵; b 为偏置; h_{t-1} 为 $t-1$ 时刻隐藏层单元的输出; x_t 为 t 时刻输入。

一些循环神经网络变体^[29-31]缓解了上述问题,其中长短期记忆网络(Long Short-term Memory, LSTM)^[29]的研究和应用最为广泛。图5给出了LSTM单元的结构。

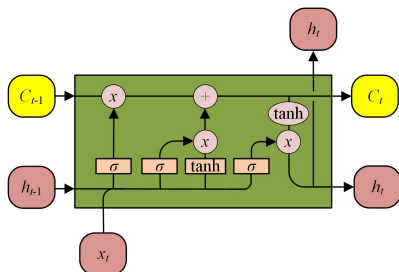


图5 长短期记忆网络单元的结构

Fig. 5 Structure of long short-term memory units

LSTM单元的前向传播计算式如下:

$$i_t = \sigma(w_{xi} x_t + w_{hi} h_{t-1} + b_i) \quad (2)$$

$$f_t = \sigma(w_{xf} x_t + w_{hf} h_{t-1} + b_f) \quad (3)$$

$$o_t = \sigma(w_{xo} x_t + w_{ho} h_{t-1} + b_o) \quad (4)$$

$$g_t = \tanh(w_{xg} x_t + w_{hg} h_{t-1} + b_g) \quad (5)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (6)$$

$$h_t = o_t \odot \tanh(c_t) \quad (7)$$

其中, x_t 为 t 时刻的输入; h_{t-1} 为 t 时刻前一时刻的RNN单元输出; w_{xi} , w_{xf} , w_{xo} 和 w_{xg} 分别为对应的输入权重矩阵; w_{hi} , w_{hf} , w_{ho} 和 w_{hg} 分别为前一层隐藏层输出的权重矩阵; b_i , b_f , b_o 和 b_g 分别为对应的偏置; \odot 表示哈达玛积; σ 为sigmoid函数。

循环神经网络单元之间的连接方式对循环神经网络的性能也有极其重要的影响。跳跃连接能缓解循环神经网络梯度消失问题^[32]。图6和图7所示的两种循环神经网络单元数和层数完全一致,但连接方式不同使得其效果往往也有差异。图6所示的双向循环神经网络中的任意中间时间节点的输出结果,不仅依赖于该时间节点之前的输入,也依赖于该时间节点之后的输入^[33]。图7所示的堆叠循环神经网络^[34]则是期望通过增加网络的深度来增加模型的精度,但这也增大了计算量,加大了陷入局部最优解的风险。

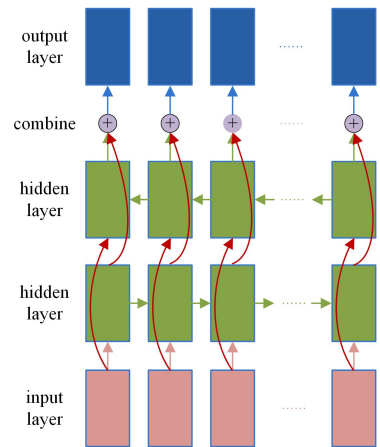


图6 双向循环神经网络结构

Fig. 6 Structure of bidirectional recurrent neural networks

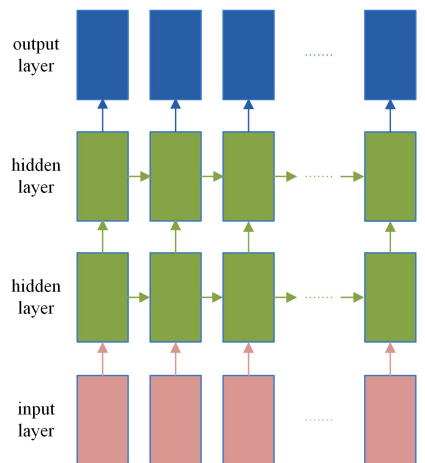


图7 堆叠循环神经网络结构

Fig. 7 Structure of stacked recurrent neural networks

3 演化循环神经网络

3.1 演化循环神经网络基本思路

演化循环神经网络是将演化计算和循环神经网络相结合的一种算法。这种算法结合了群体智能中自适应演化的机制和神经网络的学习特性,有效缓解了传统循环神经网络的缺陷。

演化循环神经网络属于演化神经网络的子方向,主要分

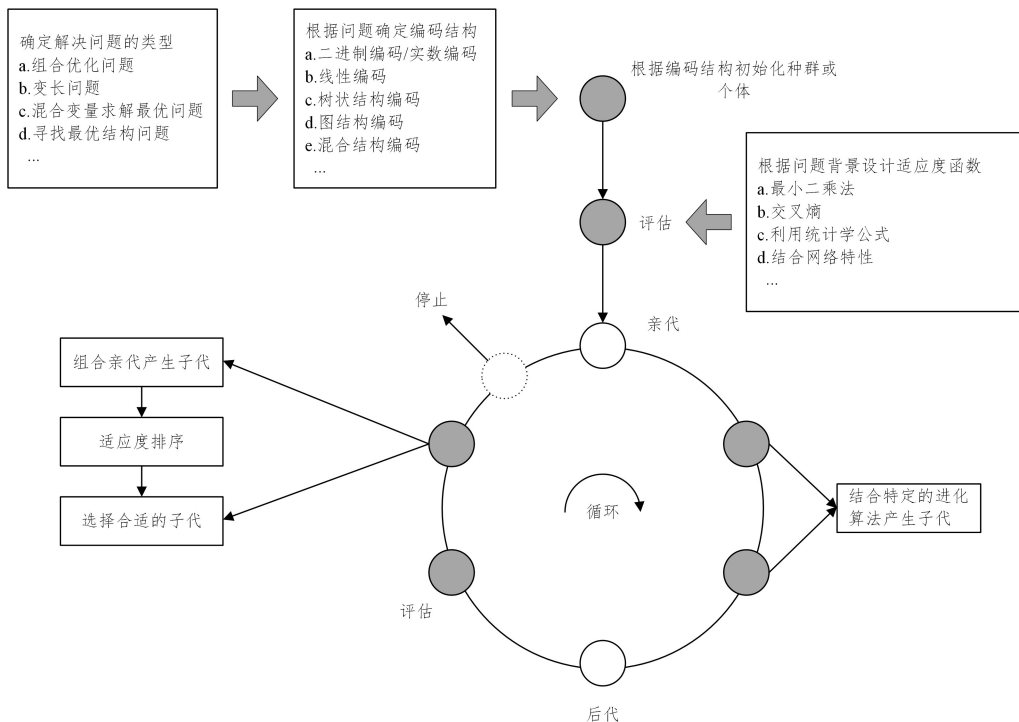


图8 演化循环神经网络的基本思路

Fig. 8 Basic thinking of evolutionary recurrent neural networks

3.2 演化循环神经网络基本类别

3.2.1 权重优化

循环神经网络的权重优化通常是使用基于梯度的方法^[9]。一方面,梯度下降法常常会造成局部最优解的问题^[10];另一方面,在使用 BPTT 方法训练时,由于时序传播,梯度与小值的相

为 4 类:权重优化、超参数优化、结构优化和与其他机器学习知识的结合。基本思路是将神经网络的优化问题抽象为演化计算中的适应度函数优化问题,分为以下 3 步:首先,将循环神经网络模型存在的问题转换为一个最优化求解的问题;其次,根据相应的问题设计合适的编码方式;最后,选取合适的演化算法^[35]。根据文献^[35]对演化神经网络的总结,演化循环神经网络的思路如图 8 所示。

乘会造成梯度消失的问题,即数据会因为传播的时间过长而被遗忘,同时,该算法也有发生梯度爆炸问题的概率^[11]。LSTM 缓解了梯度消失的问题,但未解决局部最优解的问题。演化计算能突破梯度下降法的限制,以收敛时间为代价同时解决这些问题。表 1 列出了按年份顺序汇总相关文献的结果。

表 1 循环神经网络权值演化优化的相关研究

Table 1 Research on the evolutionary optimization of recurrent neural network weights

Year	Literature	Method	Model	Taskordataset	Comparison
2001	[45]	GA	RNN	Fuzzygrammatical inference	RTRL
2007	[36]	PSO	ENN	Ultrasonic Motors speedidentification and control	GD
2007	[37]	PSO-EA	ENN	Prediction of the missing values from the given time series	PSO,EA
2007	[38]	PSO-BP	ENN	Impedance identificationof power electronic-systems	PSO,BP
2012	[44]	PSO-GA	RNN	Control of athree-dimensional tower crane	PSO,GA
2013	[39]	PSO-EA	ENN	Solar radiation prediction	-
2015	[52]	GGABC	RNN	Breast cancer prediction	ABC,Guided ABC,Global ABC
2015	[48]	CS-BP	ENN	7-Bit Parity,Selected UCI benchmarkclassification datasets	ABC-BP,ABC-LM,ABC-NN, BP-NN,CS-BP-RNN
2017	[47]	DE	NARX	Prediction of the contact area of the tire	GD
2017	[40]	PSO-BP	ENN	QTECG dataset, Two NASA space shuttle datasets, MGH/MF dataset	GD

(续表)

Year	Literature	Method	Model	Taskordataset	Comparison
2018	[41]	PSO	RNN	Voltage instability prediction	GD
2018	[50]	HS,GWO, SCA, ALOA	LSTM	Breast cancer wisconsin dataset, Epileptic seizure recognition dataset	—
2019	[49]	MA	NAR,NARX, ENN,LSTM	Predict of future energy consumption in buildings	Sequential and parallel memetic algorithm
2019	[42]	PSO-GWO	ENN	Endometrial samples classification	GD, PSO, GWO
2019	[51]	ALO A	RNN	Energy management of micro grid connected system	GA, ABC, BFA
2019	[46]	GA	NARX	SoC Estimation for Lithium-Ion Battery	GD
2019	[57]	EA	LSTM	Beijing PM2.5, SML 2010, MSR Action3D	Traditional models, GD
2020	[53]	WOA	MEEMD-LSTM	Carbon price forecasting	Traditional models, GD
2020	[54]	WOA	BRNN	Nine imbalanced datasets; yeast-2_vs_4, yeast-1_vs_7 and ECBDL 14, et al, Datasets of balloon, breastcancer, iris and heart	With and without SMOTE, Traditional models
2021	[43]	PSO	ENN	Ten benchmark classification problems of breastcancer, heart and hepatitis, et al	PSO-NN
2021	[56]	ES	LSTM	Character-level language modeling	BPTT, RTRL

粒子群算法^[36-44]在循环神经网络权值演化优化领域最受欢迎。遗传算法是应用最为广泛的演化计算方法^[44-46]。差分进化算法^[47]、布谷鸟搜索算法^[48]、文化基因算法^[49]、灰狼优化算法^[42,50]、蚂蚁狮子优化算法^[50-51]、人工蜂群算法^[52]和鲸鱼优化算法^[53-54]在此领域也有相应的工作。文献[50]分别使用和声搜索、灰狼优化算法、正余弦优化算法和蚂蚁狮子优化算法4种算法训练循环神经网络,并在两种数据集上进行了对比实验。

不同的演化算法有不同的优势和局限性,一些改进的方法应运而生^[55]。一种是对原本的演化算法进行一定的改进,能有效提升循环神经网络的性能^[36,47,49,53,56-57];另一种是将不同的算法用合理的方式结合以得到扬长避短的效果,同样能有效提升网络的性能^[37-40,42,44,48]。

3.2.2 超参数优化

循环神经网络的超参数影响着网络模型的性能,因此需要对超参数进行合理的设置或有效的优化。神经网络超参数优化是自动化深度学习的难点:1)超参数优化算法需要对离散、有序和连续的变量同时进行优化,因此是一个组合优化问题,无法使用梯度方法优化;2)待优化的超参数与最终的评价指标之间可能是非线性关系,使得优化规律较为复杂;3)每次调整出一组超参数都要训练一次神经网络,对时间和计算资源的需求较大^[58-59]。演化计算具有很好的全局搜索能力,不依靠梯度信息就能对可行域进行探索。表2列出了按年份顺序汇总相关文献的结果,表3列出了循环神经网络的常见超参数,图9给出了循环神经网络超参数演化优化的一般流程。

表2 循环神经网络超参数演化优化的相关研究

Table 2 Research on the evolutionary optimization of recurrent neural network hyperparameters

Year	Literature	Method	Model	Taskordataset
2016	[60]	CMA-ES	LSTM	Vocabulary speech recognition
2018	[61]	DE	LSTM	Electricity price prediction
2018	[62]	GA	LSTM	Daily Koreastock price index data
2018	[63]	GA, ES	RNN	Car parkoccupancy prediction
2019	[64]	GA	LSTM	Predicting housing price in Shenzhen
2019	[65]	ES	LSTM	Waste generationprediction
2019	[66]	ES	LSTM	Waste generationprediction
2019	[67]	GA	LSTM-FCN	Emotionestimation
2019	[68]	GA	LSTM	Building energy consumption forecasting
2020	[69]	FOA	LSTM	NN3, Energyconsumption prediction
2020	[70]	LSTM	GA, PSO	Electric load forecasting
2020	[71]	GA	LSTM-FCN	Blind fastener installation
2020	[72]	DE	LSTM	MNIST, IMDB, UCI HAR
2020	[73]	CMA-ES	LSTM	Short-termwind speedprediction
2020	[74]	SCOA	LSTM	Buildingenergy consumption forecasting
2020	[75]	GA	LSTM	Human activity recognition, Solar flare prediction, Air quality prediction
2021	[76]	GA	LSTM-FCN	Industrial prognosis under changing conditions
2021	[77]	GA	LSTM	Sentenceclassification
2021	[78]	BO-ES	GRU	Edgeresource usage prediction
2022	[79]	GA	LSTM	Prediction of geophysical flows

表3 循环神经网络的常见超参数

Table 3 Common hyperparameters of recurrent neural networks

Classification	Hyperparameters	Description
Value	Cell number of per layer	The depth of RNN
Type	Activation function	Type of nonlinear activation function
Value	Number of layers	Number of memory cells that stores the temporal information
Value	Dropout	Probability randomly zeroingout the partial weights or outputs of the hidden layers
Type	Optimizer	Type of weight update method
Value	Parameters of optimizer	Setting of relevant parameters
Value	Epochs	Number of weights training
Value	Number of batch size	Number of samples pergradient update
Type	Loss function	Function that measures the degree of difference between the predicted value and the true value
Value	Window size	Input number of the time series data
Type	Weight initialization	The way the initial weights are set

当前的众多研究使用了不同的演化计算方法来对超参数的不同类型的组合进行优化。文献[62,68]利用遗传算法调整 LSTM 的时间窗口大小和单元数。Peng 等^[61]使用差分进化算法来优化 LSTM 的输入长度、单元数、隐藏层层数和批量大小,随后使用果蝇优化算法做了进一步的研究^[69],相比文献[61,64]的演化方法取得了更好的效果。Somu 等^[74]使用改进的正弦余弦优化算法搜索 LSTM 的隐藏层层数、学习率、动量和权重衰减系数的最佳组合。Bouktif 等^[70]分别使用遗传算法和粒子群算法搜索出 LSTM 的输入数据个数、输入序列长度、序列起点、单元数量、批量大小、激活函数和权重优化方法的最佳组合。Rogers 等^[77]使用遗传算法优化了 LSTM 各隐藏层的维度和各层的 dropout 值。文献[67,71,76]则对 LSTM 和 FCN 的混合模型的超参数进行了搜索优化。

多目标演化算法中,协方差自适应演化策略(Covariance Matrix Adaptation-evolution Strategy, CMA-ES)也常被作为优化超参数的方法^[60,73]。在与其他算法结合方面,Violos 等^[78]将演化策略与贝叶斯优化结合起来进行超参数的优化。

在编码策略方面,Viswambaran 等^[75]进行了相应的研究,提出使用一种可变长的编码方式对超参数进行优化,实验得出使用基于可变长编码的遗传算法进行循环神经网络的超参数优化可以使此循环网络在大多数数据集上的性能超过当时最优秀的网络模型。

激活函数是神经网络的重要超参数之一,却往往被忽略。从已存的激活函数中选择出最契合的激活函数是常见做法^[75,79]。Vijayaprabakaran 等^[72]使用差分进化算法搜索出一种新的激活函数,能有效取代长短期记忆网络中传统的激活函数。

缩短优化时间将会是未来研究的重点方向。Tanaka 等^[60]使用的 CMA-ES 方法能将原本的计算时间缩短 22.7%。Camero 等^[63,65-66]着重关注循环神经网络的隐藏层层数和单元数,文献[65]在神经网络训练评估时通过减少一定比例的数据量来缩短时间,文献[66]使用了平均绝对误差随机抽样的方法来评估网络,这种方法可使所需的时间缩短一半。

3.2.3 结构优化

循环神经网络的结构优化源于一个原始的问题:什么结构的神经网络性能是最好的? 传统的研究工作往往依据人工

的方式对网络结构进行试错性设计。这种方式不仅耗时耗力,而且对专业人员的需求也较大。本文将循环神经网络的单元内部结构和这些单元之间的连接方式的优化分类为循环神经网络的结构优化。基于此,将演化计算作为搜索策略自动搜索出优秀的循环神经网络结构是亟需研究的问题之一。循环神经网络拓扑结构优化的流程与其超参数优化的流程基本一致,同样如图 9 所示。

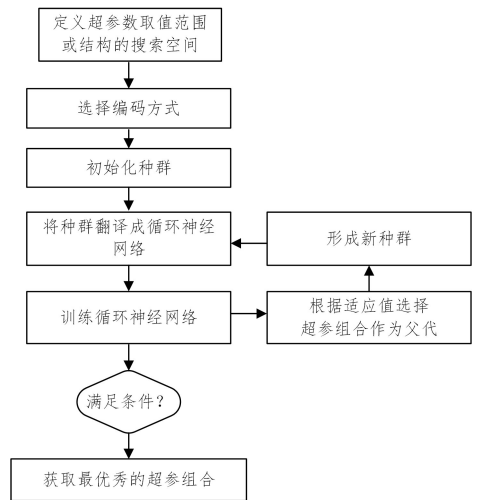


图9 循环神经网络超参数或结构演化优化的一般流程

Fig. 9 General flow of hyperparameter or structure evolutionary optimization of recurrent neural network

早期文献[80]认为遗传算法不适用于搜索网络结构,并提出了一种可获取泛化循环连接(GeNeralized Acquisition of Recurrent Links, GNARL)的演化程序。该程序同时演化出循环神经网络的结构和权重,可搜索出在人工搜索中容易被排除的拓扑结构。

Juang^[81]提出了一种混合遗传算法和粒子群优化的算法来自动设计循环神经网络。演化过程中,新一代的个体不仅通过遗传算法中的交叉和变异操作产生,还通过粒子群算法增强群体中的精英产生。该算法设计出了一个完全连通的循环神经网络,并在对比实验中被证明比遗传算法和粒子群优化算法更优越。

Bayer 等^[82]使用 NSGA-II 多目标演化算法对 LSTM 单元结构进行变异演化,并使用新的单元结构构建出新的循环

神经网络,用于 T-Maze 任务。尽管这样构建出的网络可用于解决部分可观察马尔可夫决策过程(Partially Observable Markov Decision Process, POMDP),但其深度却不足以用于现实的人工智能任务。

Wang 等^[83]基于图像标注任务提出了一种演化策略,发现了 100 多个 RNN 结构变体,以及许多新的子结构,如新的激活函数、新的 dropout 层等。

Rawal 等^[84]将神经演化算法 NEAT^[85]应用于 LSTM 神经单元,并将信息最大化作为目标来对 LSTM 结构进行演化优化。Rawal 等^[86]还采用了基于树编码的演化算法来演化拓扑结构,第一个实验在一个固定的循环神经网络架构中使用遗传规划演化出优秀的单元结构,第二个实验则使用演化算法将不同类型的节点组合到一个隐藏层中,该文还使用代理模型来缩短评估时间,降低资源消耗。

文献[87-89]通过基于蚁群优化算法选择 RNN 和 LSTM 中应使用的连接来预测飞行参数。特别地,文献[88]被证明可以将可训练连接的数量减少一半,同时显著改善发动机振动的预测。然而,这种方法在固定的循环神经网络体系结构中工作,不能演化出整体的循环神经网络结构,并且这些工作尚局限于具有单一周期时间跨度和简单连接的小型循环神经网络。文献[90]首次利用蚁群算法来优化具有不同时间跨度和连接复杂的循环神经网络。

Desell 等^[91]提出了 EXALT 方法,同时进行了单元和连接的演化优化,演化的单元基于简单的 RNN 单元和 LSTM 单元,子代个体依据预设的概率对父代个体进行变异(包括节点变异和边变异)、交叉和复制操作后产生。在评估加速方面,先使用父代的权重初始化子代的权重,然后对子代进行反向传播训练。在此基础上,Ororbia 等^[92]提出了 EXAMM 算法,该算法能够使用多种现存的单元结构(Δ -RNN^[93], GRU^[31], LSTM, MGU^[94] 和 UGRNN^[95])来演化循环神经网络,并进一步细化了 EXALT 的突变操作来减少超参数。相比 EXALT 只使用一个种群,EXAMM 在此基础上使用岛(每次训练 10 个岛,每个岛包括 5 个种群)的方法^[96]扩大了搜索规模,并利用并行机制加快了计算速度,还添加了岛间交叉操作来加强岛之间的信息交流。该团队随后继续使用 EXAMM 算法进行了广泛的实验^[97],实验证明具有深度循环连接(时间跨度较长的连接)的循环神经网络的性能明显优于没有这种连接的网络,在某些情况下,性能优异的循环神经网络仅由简单的神经元结构和深度循环连接组成。

显而易见,循环神经网络的大规模神经架构搜索对时间和计算资源的耗费是巨大的(如 EXAMM 算法共循环了 1210 次,演化出了 2420000 个循环神经网络,在高性能的计算集群器上共运行了 12100 个 CPU 小时),一些研究因此致力于加速的工作。Mo 等^[98]采用了早停的方法(通过减少训练次数来缩短时间),并将已经被评估的结构保存了下来,但是早停会造成一些前期训练表现差但训练至收敛后性能优秀的网络被淘汰。文献[86]则使用一种基于 LSTM 的序列到序列的代理模型预测新结构的适应度,但这种方法需要先训练足够数量的样本来满足代理模型的训练量,以确保代理

模型的预测能达到足够的精度。Camero 等^[99]提出了一种基于随机误差抽样的神经演化算法,在网络结构优化过程中,使用这种无训练的方法来预测循环神经网络的性能可将耗时缩短一半。文献[100]则对 EXAMM 算法^[92]中的权重初始化和权重继承做了进一步的对比实验研究,结果表明拉马克权重策略相比其他方法的表现更好,能更大程度地减少循环神经网络的训练次数。

4 其他相关研究

4.1 演化循环神经网络的集成学习

研究表明,集成模型相比于单一模型有更好的泛化性能^[101]。演化循环神经网络在集成学习中也有相应的研究工作。

循环神经网络的集成是演化神经网络集成^[102]的子方向,文献[102]总结了演化神经网络集成的两种形式:1)先用演化算法设计一定数量的神经网络,然后采用某种方式把这些神经网络集成为一个最终系统;2)先设计好一定数量的神经网络,再采用演化算法把它们集成为一个最终系统。从整理的相关文献可知,目前演化循环神经网络的集成学习大多采用第一种形式。图 10 给出了演化循环神经网络集成学习的一般流程。

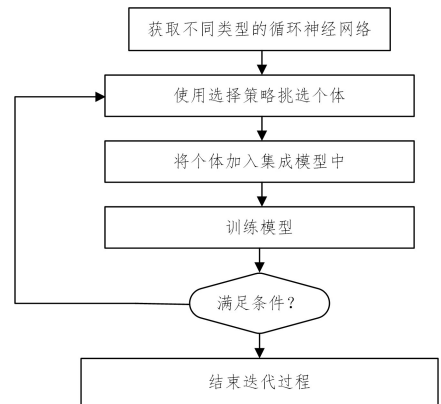


图 10 演化循环神经网络集成学习的一般流程

Fig. 10 General flow of evolutionary recurrent neural network ensemble learning

Smith 等^[103]提出了一种混合多目标演化算法,用于训练和优化循环神经网络的结构,并对 3 种不同的选择方法(选择帕雷托前沿中低于阈值的个体、基于训练误差选择个体和根据个体预测值的多样性选择个体)进行了对比实验,最后集成模型的预测值,取所有被选择个体预测值的平均值。

Ai 等^[104]将具有相同或相似结构的 LSTM 网络作为集成模型中的个体,使用演化算法不断地迭代更新集成模型来提高预测精度,输出结果取所有 LSTM 预测值的平均值。

Viswambaran 等^[105]基于 NSGA-II 提出了一种新的演化多目标的算法,用于自动设计出一组具备多样性和高精度的 LSTM。集成方法是先使用一定数量的精度最高的 LSTM 来初始化集成模型,然后不断加入能使集成模型的预测精度有效提升的 LSTM 个体(迭代过程中使用集成学习方法中的 Bagging 来训练集成模型),直到加入个体使集成模型的精度

下降,则停止加入这个个体并得到最终的集成模型。

Wiswambaran 等^[106] 随后还提出了另一种方法来完成 LSTM 的集成:第一阶段,自动演化出性能优异的 LSTM 个体,过程中使用权重继承的方式来加速 LSTM 个体的训练时间;第二阶段,选择合适的 LSTM 作为集成的成员,并不限定选择的个数,然后使用 Bagging 的方法来实现集成。

4.2 演化循环神经网络的迁移学习

循环神经网络的迁移学习指在源数据集上训练循环神经网络,然后将这个训练好的循环神经网络的结构、超参数和权重保存下来并进行适当的调整后,将其用于目标数据集(目标数据集与源数据集不同但有一定相似性)的预测任务。相关迁移学习的知识可参考综述文献^[107-108],本文不做过多赘述。

迁移学习存在以下问题:1)循环神经网络的迁移学习很少受到研究者的关注;2)迁移学习很少对所迁移的网络模型进行大规模的改进^[109]。

Ororbia 等采用此前提出的 EXAMM 神经演化算法^[92] 对循环神经网络的迁移学习做了一系列的研究^[109-111]。文献^[109]首次将神经演化用于循环神经网络的迁移学习领域(甚至是整个迁移学习的领域),并首次提出能够适应不同输入输出维度的算法框架。该算法可以根据需要添加和删除节点来改动网络模型(已经在源数据集上得到训练)的输入和输出以适应目标数据集,然后将改进后的模型用于目标数据集并在目标数据集上继续进行结构的演化优化,其对比实验结果表明,该方法使得循环神经网络在目标数据集上的演化和训练要快于从头演化和训练循环神经网络,并且有效地提高了循环神经网络的泛化性能。文献^[110-111]则对上述研究工作做了进一步的改进,并取得了更好的效果。

5 总结与展望

目前,演化计算与循环神经网络的结合算法在权重优化、超参数优化、结构优化、集成学习和迁移学习方面都有了相关的研究,而这些不同的研究方向都存在自身的问题,并有待进一步的探究和优化,未来的研究工作可参考以下方面。

(1) 没有统一数据集进行对比实验的问题

当前的演化循环神经网络应用较为广泛,大体可分为序列到序列模型和序列到标签模型,大量的研究工作没有统一的数据集,因此很难对这些算法进行统一的比较,如在结构优化中,需要比较整个运算所消耗的运算资源和时间以及最终搜索出的网络结构的精度和运算量等。

(2) 时间和计算资源耗费巨大的问题

超参数优化、结构优化、集成学习和迁移学习都需要对大量的循环神经网络的结构进行训练,以评估其性能,而这些过程都需要耗费大量的计算资源和时间。关于如何降低这些消耗,有权重继承、使用代理模型预测性能和降低数据集规模等方法。在权重优化方面,演化计算的方法通常不如梯度下降算法的收敛速度快。未来工作应继续考虑如何在保证网络精度的前提下不断提升速度,譬如将当前阶段神经架构搜索研究中已存的各种加速方法迁移到循环神经网络超参数或结构

演化的加速方面。演化循环神经网络的加速演化将会是这一方向永恒的课题,这一点类似于网速的加速研究。

(3) 自动化不足的问题

一方面,演化计算选用的方法都是人为设定的,如何根据具体问题的不同选择最适合的演化算法也是未来需要研究的问题,并且解决这一问题尚缺少庞大的数据。另一方面,在循环神经网络的结构演化优化中,往往无法同时实现超参数自动优化,循环神经网络的结构和超参数的组合方式的自适应也需要进一步的研究。

(4) 集成学习和迁移学习研究较少问题

在某些领域,数据集的获取有一定难度,小样本学习依旧是深度学习的热点问题。集成学习和迁移学习能有效提升预测模型的泛化能力,利用少量的训练数据就能得到相对较高精度的模型。目前,演化循环神经网络的集成学习很少采用文献^[102]中提及的第二种形式,即先设计好一定数量的神经网络,然后采用演化算法把它们集成为一个最终系统。因此,尝试使用演化计算的方法来完成循环神经网络的集成过程也会是未来需要研究的方向之一。在研究自动化机器学习和深度学习的时代背景下,集成学习和迁移学习的自动化实现也应当受到足够的重视。

结束语 本文聚焦于演化计算与循环神经网络的结合这一人工智能方向,介绍了演化循环神经网络的基本思路和框架,然后分析并整理了使用演化计算优化循环神经网络的权重、超参数和结构的相关研究文献,接着对演化循环神经网络的集成学习和迁移学习也做了相应的分析和总结,最后针对当前研究阶段存在的问题提出了一些建议。演化循环神经网络应当是当前自动化机器学习和深度学习研究的一个必要课题,值得各界学者进行进一步的深入研究。

参 考 文 献

- [1] WANG X W, LI J, TAN Z H, et al. The state of the art and future tendency of "Internet+" oriented network technology[J]. Journal of Computer Research and Development, 2016, 53(4): 729-741.
- [2] HAO L, SHAO Z P, ZHANG J H, et al. Review of deep learning-based action recognition algorithm[J]. Computer Science, 2020, 47(S1): 139-147.
- [3] KAVIANI S, SOHN I. Influence of random topology in artificial neural networks: A survey[J]. ICT Express, 2020, 6(2): 145-150.
- [4] LIPTON Z C, BERKOWITZ J, ELKAN C. A critical review of recurrent neural networks for sequence learning [J]. arXiv: 1506.00019, 2015.
- [5] CHE Z, PURUSHOTHAM S, CHO K, et al. Recurrent neural networks for multivariate time series with missing values[J]. Scientific reports, 2018, 8(1): 1-12.
- [6] MIKOLOV T, KARAFIÁT M, BURGET L, et al. Recurrent neural network based language model[C]// Interspeech. 2010: 1045-1048.
- [7] KALCHBRENNER N, BLUNSOM P. Recurrent continuous

- translation models[C]//Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, 2013;1700-1709.
- [8] YOU Q, JIN H, WANG Z, et al. Image captioning with semantic attention[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016;4651-4659.
- [9] WERBOS P J. Backpropagation through time; what it does and how to do it[C]//Proceedings of the IEEE, 1990;1550-1560.
- [10] BIANCHINI M, GORI M, MAGGINI M. On the problem of local minima in recurrent neural networks[J]. IEEE Transactions on Neural Networks, 1994, 5(2):167-177.
- [11] PASCANU R, MIKOLOV T, BENGIO Y. On the difficulty of training recurrent neural networks[C]//International Conference on Machine Learning. PMLR, 2013;1310-1318.
- [12] YAO Q, WANG M, CHEN Y, et al. Taking human out of learning applications: A survey on automated machine learning[J]. arXiv:1810.13306, 2018.
- [13] SPARKS E R, TALWALKAR A, HAAS D, et al. Automating model search for large scale machine learning[C]//Proceedings of the Sixth ACM Symposium on Cloud Computing, 2015;368-380.
- [14] SCHRIMPF M, MERITY S, BRADBURY J, et al. A flexible approach to automated rnn architecture generation [J]. arXiv:1712.07316, 2017.
- [15] ZOPH B, LE Q V. Neural architecture search with reinforcement learning[J]. arXiv:1611.01578, 2016.
- [16] PHAM H, GUAN M, ZOPH B, et al. Efficient neural architecture search via parameters sharing[C]//International Conference on Machine Learning. PMLR, 2018;4095-4104.
- [17] BACK T, HAMMEL U, SCHWEFEL H P. Evolutionary computation; Comments on the history and current state[J]. IEEE Transactions on Evolutionary Computation, 1997, 1(1):3-17.
- [18] KOU G J, MA Y Y, YUE J, et al. Survey of bio-inspired natural computing[J]. Computer Science, 2014, 41(S1):37-41.
- [19] KATOCH S, CHAUHAN S S, KUMAR V. A review on genetic algorithm: past, present, and future[J]. Multimedia Tools and Applications, 2021, 80(5):8091-8126.
- [20] DAS S, SUGANTHAN P N. Differential evolution: A survey of the state-of-the-art [J]. IEEE Transactions on Evolutionary Computation, 2010, 15(1):4-31.
- [21] KENNEDY J, EBERHART R. Particle swarm optimization [C]//International Conference on Neural Networks (ICNN'95), 1995;1942-1948.
- [22] KRASNOGOR N, SMITH J. A tutorial for competent memetic algorithms; model, taxonomy, and design issues[J]. IEEE transactions on Evolutionary Computation, 2005, 9(5):474-488.
- [23] ZHOU A, QU B Y, LI H, et al. Multiobjective evolutionary algorithms: A survey of the state of the art[J]. Swarm and evolutionary computation, 2011, 1(1):32-49.
- [24] YAO X, CHEN G L, XU H M, et al. A survey of evolutionary algorithms[J]. Chinese Journal of Computers, 1995 (9):694-706.
- [25] XIE X, LIU Y, SUN Y, et al. BenchENAS: A Benchmarking Platform for Evolutionary Neural Architecture Search[J]. IEEE Transactions on Evolutionary Computation, 2022, 26(6):1473-1485.
- [26] ELMAN J L. Finding structure in time[J]. Cognitive science, 1990, 14(2):179-211.
- [27] KARPATHY A. The unreasonable effectiveness of recurrent neural networks[EB/OL]. (2015-05-21) [2016-08-13]. <http://karpathy.github.io/2015/05/21/rnn-effectiveness>, 2016.
- [28] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[J]. arXiv:1409.0473, 2014.
- [29] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural computation, 1997, 9(8):1735-1780.
- [30] JAEGER H, HAAS H. Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication [J]. Science, 2004, 304(5667):78-80.
- [31] CHUNG J, GULCEHRE C, CHO K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [J]. arXiv:1412.3555, 2014.
- [32] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553):436-444.
- [33] GRAVES A, FERNÁNDEZ S, SCHMIDHUBER J. Bidirectional LSTM networks for improved phoneme classification and recognition[C]//International Conference on Artificial Neural Networks, 2005;799-804.
- [34] GRAVES A, MOHAMED A, HINTON G. Speech recognition with deep recurrent neural networks[C]//2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 2013;6645-6649.
- [35] MA L B, LI N, CHENG S. Evolutionary neural networks: principles, models and methods[J]. Journal of Shaanxi Normal University(Natural Science Edition), 2021, 49(5):30-38, 133.
- [36] GE H W, LIANG Y C, MARCHESE M. A modified particle swarm optimization-based dynamic recurrent neural network for identifying and controlling nonlinear systems[J]. Computers & Structures, 2007, 85(21/22):1611-1622.
- [37] CAI X, ZHANG N, VENAYAGAMOORTHY G K, et al. Time series prediction with recurrent neural networks trained by a hybrid PSO-EA algorithm[J]. Neurocomputing, 2007, 70(13/14/15):2342-2353.
- [38] XIAO P, VENAYAGAMOORTHY G K, CORZINE K A. Combined training of recurrent neural networks with particle swarm optimization and backpropagation algorithms for impedance identification[C]//2007 IEEE Swarm Intelligence Symposium, 2007;9-15.
- [39] ZHANG N, BEHERA P K, WILLIAMS C. Solar radiation prediction based on particle swarm optimization and evolutionary algorithm using recurrent neural networks[C]//2013 IEEE International Systems Conference(SysCon). IEEE, 2013;280-286.
- [40] KANG Q, LIAO W K, AGRAWAL A, et al. A hybrid training algorithm for recurrent neural network using particle swarm optimization-based preprocessing and temporal error aggregation [C]//2017 IEEE International Conference on Data Mining

- Workshops(ICDMW). IEEE,2017;812-817.
- [41] IBRAHIM A M,EL-AMARY N H. Particle Swarm Optimization trained recurrent neural network for voltage instability prediction[J]. *Journal of Electrical Systems and Information Technology*,2018,5(2):216-228.
- [42] HU H,WANG H,BAI Y,et al. Determination of endometrial carcinoma with gene expression based on optimized Elman neural network[J]. *Applied Mathematics and Computation*,2019,341:204-214.
- [43] AB AZIZ M F,MOSTAFA S A,FOOZY C F M,et al. Integrating Elman recurrent neural network with particle swarm optimization algorithms for an improved hybrid training of multidisciplinary datasets[J]. *Expert Systems with Applications*,2021,183:115441.
- [44] DUONG S C,UEZATO E,KINJO H,et al. A hybrid evolutionary algorithm for recurrent neural network control of a three-dimensional tower crane[J]. *Automation in Construction*,2012,23:55-63.
- [45] BLANCO A,DELGADO M,PEGALAJAR M C. A real-coded genetic algorithm for training recurrent neural networks[J]. *Neural Networks*,2001,14(1):93-105.
- [46] GUO C X,YUAN G,ZHU C Z,et al. SoC estimation for lithium battery using recurrent NARX neural network and genetic algorithm[C]// *IOP Conference Series: Materials Science and Engineering*. IOP Publishing,2019.
- [47] DUCHANOY C A,MORENO-ARMENDÁRIZ M A,URBINA L,et al. A novel recurrent neural network soft sensor via a differential evolution training algorithm for the tire contact patch [J]. *Neurocomputing*,2017,235:71-82.
- [48] NAWI N M,KHAN A,REHMAN M Z,et al. Weight optimization in recurrent neural networks with hybrid metaheuristic Cuckoo search techniques for data classification[J]. *Mathematical Problems in Engineering*,2021,18:115441.
- [49] RUÍZ L G B,CAPEL M I,PEGALAJAR M C. Parallel memetic algorithm for training recurrent neural networks for the energy efficiency problem[J]. *Applied Soft Computing*,2019,76:356-368.
- [50] RASHID T A,FATTAH P,AWLA D K. Using accuracy measure for improving the training of LSTM with metaheuristic algorithms[J]. *Procedia Computer Science*,2018,140:324-333.
- [51] ROY K,MANDAL K K,MANDAL A C. Ant-Lion Optimizer algorithm and recurrent neural network for energy management of micro grid connected system[J]. *Energy*,2019,167:402-416.
- [52] SHAH H,CHIROMA H,HERAWAN T,et al. An efficient bio-inspired bees colony for breast cancer prediction[C]// *Proceedings of the International Conference on Data Engineering 2015(DaEng-2015)*. Singapore:Springer,2019:597-608.
- [53] YANG S,CHEN D,LI S,et al. Carbon price forecasting based on modified ensemble empirical mode decomposition and long short-term memory optimized by improved whale optimization algorithm[J]. *Science of the Total Environment*,2020,716:137117.
- [54] HASSIB E,EL-DESOUKY A,LABIB L,et al. WOA+BRNN: An imbalanced big data classification framework using Whale optimization and deep neural network [J]. *Soft Computing*,2020,24(8):5573-5592.
- [55] GROSAN C,ABRAHAM A. Hybrid evolutionary algorithms: methodologies, architectures, and reviews[M]// *Hybrid evolutionary Algorithms*. 2007:1-17.
- [56] VICOL P,METZ L,SOHL-DICKSTEIN J. Unbiased gradient estimation in unrolled computation graphs with persistent evolution strategies [C]// *International Conference on Machine Learning*. PMLR,2021:10553-10563.
- [57] LI Y,ZHU Z,KONG D,et al. EA-LSTM: Evolutionary attention-based LSTM for time series prediction [J]. *Knowledge-Based Systems*,2019,181:104785.
- [58] BERGSTRA J,BARDENET R,BENGIO Y,et al. Algorithms for hyper-parameter optimization[C]// *International Conference on Neural Information Processing Systems*. 2011.
- [59] LORENZO P R,NALEPA J,KAWULOK M,et al. Particle swarm optimization for hyper-parameter selection in deep neural networks[C]// *Proceedings of the Genetic and Evolutionary Computation Conference*. 2017:481-488.
- [60] TANAKA T,MORIYA T,SHINOZAKI T,et al. Automated structure discovery and parameter tuning of neural network language model based on evolution strategy[C]// *2016 IEEE Spoken Language Technology Workshop (SLT)*. IEEE,2016:665-671.
- [61] PENG L,LIU S,LIU R,et al. Effective long short-term memory with differential evolution algorithm for electricity price prediction[J]. *Energy*,2018,162:1301-1314.
- [62] CHUNG H,SHIN K. Genetic algorithm-optimized long short-term memory network for stock market prediction[J]. *Sustainability*,2018,10(10):1-18.
- [63] CAMERO A,TOUTOUH J,STOLFI D H,et al. Evolutionary deep learning for car park occupancy prediction in smart cities [C]// *International Conference on Learning and Intelligent Optimization*. Cham:Springer,2018:386-401.
- [64] LIU R,LIU L. Predicting housing price in China based on long short-term memory incorporating modified genetic algorithm [J]. *Soft Computing*,2019,23(22):11829-11838.
- [65] CAMERO A,TOUTOUH J,FERRER J,et al. Waste generation prediction under uncertainty in smart cities through deep neuroevolution[J]. *Revista Facultad de Ingeniería Universidad de Antioquia*,2019(93):128-138.
- [66] CAMERO A,TOUTOUH J,ALBA E. A specialized evolutionary strategy using mean absolute error random sampling to design recurrent neural networks[J]. *arXiv:1909.02425*,2019.
- [67] CHUNG C C,LIN W T,ZHANG R,et al. Emotion estimation by joint facial expression and speech tonality using evolutionary deep learning structures[C]// *2019 IEEE 8th Global Conference on Consumer Electronics (GCCE)*. IEEE,2019:221-224.
- [68] ALMALAQ A,ZHANG J J. Evolutionary deep learning-based energy consumption prediction for buildings[J]. *IEEE Access*,2018,7:1520-1531.
- [69] PENG L,ZHU Q,LV S X,et al. Effective long short-term me-

- memory with fruit fly optimization algorithm for time series forecasting[J]. *Soft Computing*, 2020, 24(19): 15059-15079.
- [70] BOUKTIF S, FIAZ A, OUNI A, et al. Multi-sequence LSTM-RNN deep learning and metaheuristics for electric load forecasting[J]. *Energies*, 2020, 13(2): 1-21.
- [71] ORTEGO P, DIEZ-OLIVAN A, DEL SER J, et al. Evolutionary LSTM-FCN networks for pattern classification in industrial processes[J]. *Swarm and Evolutionary Computation*, 2020, 54: 100650.
- [72] VIJAYAPRABAKARAN K, SATHIYAMURTHY K. Towards activation function search for long short-term model network: A differential evolution based approach[J]. *Journal of King Saud University-Computer and Information Sciences*, 2022, 34, 2637-2650.
- [73] NESHAT M, NEZHAD M M, ABBASNEJAD E, et al. An evolutionary deep learning method for short-term wind speed prediction: A case study of the lillgrund offshore wind farm[J]. *arXiv*, 2002. 09106, 2020.
- [74] SOMU N, MR G R, RAMAMRITHAM K. A hybrid model for building energy consumption forecasting using long short term memory networks[J]. *Applied Energy*, 2020, 261: 114131.
- [75] VISWAMBARAN R A, CHEN G, XUE B, et al. Evolving deep recurrent neural networks using a new variable-length genetic algorithm[C]//2020 IEEE Congress on Evolutionary Computation(CEC). IEEE, 2020: 1-8.
- [76] DIEZ-OLIVAN A, ORTEGO P, DEL SER J, et al. Adaptive dendritic cell-deep learning approach for industrial prognosis under changing conditions[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(11): 7760-7770.
- [77] ROGERS B, NOMAN N, CHALUP S, et al. Evolutionary Hyperparameter Optimisation for Sentence Classification [C] // 2021 IEEE Congress on Evolutionary Computation (CEC). IEEE, 2021: 958-965.
- [78] VIOLOS J, TSANAKAS S, THEODOROPOULOS T, et al. Hypertuning GRU Neural Networks for Edge Resource Usage Prediction[C]//2021 IEEE Symposium on Computers and Communications(ISCC). IEEE, 2021: 1-8.
- [79] PAWAR S A, SAN O, YEN G G. Hyperparameter Search using the Genetic Algorithm for Surrogate Modeling of Geophysical Flows[C]//AIAA SCITECH 2022 Forum. 2022.
- [80] ANGELINE P J, SAUNDERS G M, POLLACK J B. An evolutionary algorithm that constructs recurrent neural networks[J]. *IEEE Transactions on Neural Networks*, 1994, 5(1): 54-65.
- [81] JUANG C F. A hybrid of genetic algorithm and particle swarm optimization for recurrent network design[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B(Cybernetics)*, 2004, 34(2): 997-1006.
- [82] BAYER J, WIERSTRA D, TOGELIUS J, et al. Evolving memory cell structures for sequence learning[C]//International Conference on Artificial Neural Networks. Berlin: Springer, 2009: 755-764.
- [83] WANG H, WANG H, XU K. Evolutionary recurrent neural network for image captioning[J]. *Neurocomputing*, 2020, 401: 249-256.
- [84] RAWAL A, MIIKKULAINEN R. Evolving deep LSTM-based memory networks using an information maximization objective [C]//Proceedings of the Genetic and Evolutionary Computation Conference 2016. 2016: 501-508.
- [85] STANLEY K O, MIIKKULAINEN R. Evolving neural networks through augmenting topologies[J]. *Evolutionary computation*, 2002, 10(2): 99-127.
- [86] RAWAL A, MIIKKULAINEN R. From nodes to networks: Evolving recurrent neural networks [J]. *arXiv*: 1803. 04439, 2018.
- [87] DESELL T, CLACHAR S, HIGGINS J, et al. Evolving deep recurrent neural networks using ant colony optimization[C]//European Conference on Evolutionary Computation in Combinatorial Optimization. 2015: 86-98.
- [88] ELSAID A E R, EL JAMIY F, HIGGINS J, et al. Optimizing long short-term memory recurrent neural networks using ant colony optimization to predict turbine engine vibration[J]. *Applied Soft Computing*, 2018, 73: 969-991.
- [89] ELSAID A E R, JAMIY F E, HIGGINS J, et al. Using ant colony optimization to optimize long short-term memory recurrent neural networks[C]//Proceedings of the Genetic and Evolutionary Computation Conference. 2018: 13-20.
- [90] ELSAID A E R A, ORORBIA A G, DESELL T J. The ant swarm neuro-evolution procedure for optimizing recurrent networks[J]. *arXiv*: 1909. 11849, 2019.
- [91] ELSAID A E R, BENSON S, PATWARDHAN S, et al. Evolving recurrent neural networks for time series data prediction of coal plant parameters[C]//International Conference on the Applications of Evolutionary Computation (Part of EvoStar). Cham: Springer, 2019: 488-503.
- [92] ORORBIA A, ELSAID A E R, DESELL T. Investigating recurrent neural network memory structures using neuro-evolution [C]//Proceedings of the Genetic and Evolutionary Computation Conference. 2019: 446-455.
- [93] ORORBIA II A G, MIKOLOV T, REITTER D. Learning simpler language models with the differential state framework[J]. *Neural Computation*, 2017, 29(12): 3327-3352.
- [94] ZHOU G B, WU J, ZHANG C L, et al. Minimal gated unit for recurrent neural networks[J]. *International Journal of Automation and Computing*, 2016, 13(3): 226-234.
- [95] COLLINS J, SOHL-DICKSTEIN J, SUSSILLO D. Capacity and trainability in recurrent neural networks[J]. *arXiv*: 1611. 09913, 2016.
- [96] ALBA E, TOMASSINI M. Parallelism and evolution-ary algorithms [J]. *IEEE transactions on evolutionary computation*, 2002, 6(5): 443-462.
- [97] DESELL T, ELSAID A E R A, ORORBIA A G. Investigating Deep Recurrent Connections and Recurrent Memory Cells Using Neuro-Evolution[M]//Deep Neural Evolution. 2020: 253-291.
- [98] MO H, CUSTODE L L, IACCA G. Evolutionary neural architecture search for remaining useful life prediction[J]. *Applied Soft Computing*, 2021, 108: 107474.

- [99] CAMERO A, TOUTOUH J, ALBA E. Random error sampling-based recurrent neural network architecture optimization[J]. *Engineering Applications of Artificial Intelligence*, 2020, 96: 103946.
- [100] LYU Z, ELSAID A E R, KARNAS J, et al. An Experimental Study of Weight Initialization and Weight Inheritance Effects on Neuroevolution[J]. *arXiv*; 2009. 09644, 2020.
- [101] HANSEN L K, SALAMON P. Neural network ensembles[J]. *IEEE transactions on pattern analysis and machine intelligence*, 1990, 12(10): 993-1001.
- [102] YAO W S, WAN Q, CHEN Z Q, et al. The researching overview of evolutionary neural networks[J]. *Computer Science*, 2004 (3): 125-129.
- [103] SMITH C, JIN Y. Evolutionary multi-objective generation of recurrent neural network ensembles for time series prediction[J]. *Neurocomputing*, 2014, 143: 302-311.
- [104] AI S, CHAKRAVORTY A, RONG C. Evolutionary ensemble lstm based household peak demand prediction[C]// 2019 International Conference on Artificial Intelligence in Information and Communication(ICAIC). IEEE, 2019: 1-6.
- [105] VISWAMBARAN R A, CHEN G, XUE B, et al. Evolutionary design of long short term memory (lstm) ensemble[C]// 2020 IEEE Symposium Series on Computational Intelligence(SSCI). IEEE, 2020: 2692-2698.
- [106] VISWAMBARAN R A, CHEN G, XUE B, et al. Two-Stage Genetic Algorithm for Designing Long Short Term Memory (LSTM) Ensembles[C]// 2021 IEEE Congress on Evolutionary Computation(CEC). IEEE, 2021: 942-949.
- [107] PAN S J, YANG Q. A survey on transfer learning [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2009, 22(10): 1345-1359.
- [108] ZHUANG F, QI Z, DUAN K, et al. A comprehensive survey on transfer learning[C]// *Proceedings of the IEEE*. 2020: 43-76.
- [109] ELSAID A E R, KARNAS J, LYU Z, et al. Neuro-evolutionary transfer learning through structural adaptation[C]// *International Conference on the Applications of Evolutionary Computation(Part of EvoStar)*. Cham; Springer, 2020: 610-625.
- [110] ELSAID A E R, KARNAS J, ORORBIA II A, et al. Neuro-evolutionary Transfer Learning of Deep Recurrent Neural Networks through Network-Aware Adaptation[J]. *arXiv*; 2006. 02655, 2020.
- [111] ELSAID A E R, KARNAS J, LYU Z, et al. Improving neuro-evolutionary transfer learning of deep recurrent neural networks through network-aware adaptation [C]// *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*. 2020: 315-323.



HU Zhongyuan, born in 1999, postgraduate. His main research interests include evolutionary computation and deep learning.



XUE Yu, born in 1981, Ph.D, professor. His main research interests include deep learning, evolutionary computation, machine learning and computer vision.

(责任编辑:喻藜)