

求解高维昂贵多目标问题的约束型 Dropout代理辅助进化算法

张 睿,白晓露,潘理虎

(太原科技大学计算机科学与技术学院,山西太原 030024)

摘要: 现实优化问题常需同时对多个冲突目标进行评估优化,由于评估过程多依赖大量复杂的仿真实验,从而产生计算代价昂贵的问题.代理模型辅助下的进化算法可在计算资源有限的情况下为此类问题提供有效的最优解集.但随着问题决策以及目标空间维度变高,则会衍生出诸多负面因素限制代理模型的预测精度.对此,提出一种基于约束型 Dropout 神经网络的代理辅助进化算法(Constrained Dropout Neural Network based surrogate-assisted Evolution Algorithm, CDNNEA),以约束型 Dropout 神经网络作为一种可扩展方案来增强代理模型在高维空间中的适用性,在模型管理部分中,构建种群个体收敛性以及多样性评判指标,自适应地选取引导代理模型更新的代表性个体.通过在 DTLZ 基准测试问题上进行实验,CDNNEA 显示出相较于其它先进算法性能表现最优或近似最优,同时将其应用于高维参数自适应优化的现实问题中.实验表明:提供的代表性解决方案投入实际应用时识别准确率均较优且模型计算量可控,验证出 CDNNEA 在高维昂贵类优化问题上的有效性.

关键词: 代理模型;昂贵问题;高维多目标优化;模型管理;进化算法

基金项目: 山西省基础研究计划(自由探索类)项目(No.20210302123216);山西省装备数字化与故障预测工程研究中心开放课题(No.ZBPHM20201104);山西省研究生教育改革研究课题项目(No.2021YJJG244);山西省研究生教育创新项目(No.2021Y699)

中图分类号: TP273

文献标识码: A

文章编号: 0372-2112(2023)07-1859-09

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20220694

Constrained Dropout Surrogate-Assisted Evolutionary Algorithm for Expensive Many-Objective Problems

ZHANG Rui, BAI Xiao-lu, PAN Li-hu

(School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan, Shanxi 030024, China)

Abstract: Realistic optimization problems often need to evaluate and optimize multiple conflicting targets at the same time. Because the evaluation process depends on a large number of complex simulation experiments, the calculation cost is expensive. The evolutionary algorithms assisted by the surrogate models can provide effective optimal solution sets for such problems with limited computational resources. However, as the dimensions of decision space and target space become higher, many negative factors will be derived to limit the prediction accuracy of the surrogate model. Therefore, a surrogate-assisted evolution algorithm based on the constrained dropout neural network (CDNNEA) is proposed. The constrained dropout neural network is used as a scalable scheme to enhance the applicability of surrogate models in high-dimensional space. In the model management part, the convergence and diversity evaluation indicators of the population individuals are constructed, in order to adaptively select the representative individuals to guide the surrogate model to update. Experiments on DTLZ benchmark problems show that CDNNEA performs the best or approximates the best compared with other advanced algorithms. Meanwhile, CDNNEA is applied to the high-dimensional parameter adaptive optimization of practical problems. The experimental results show the proposed representative solutions have better recognition accuracy and controllable model computation in practical application, which verifies the effectiveness of CDNNEA in high-dimensional expensive optimization problems.

Key words: surrogate models; expensive problem; multi-objective optimization; model management; evolutionary

algorithm

Foundation Item(s): Fundamental Research Program of Shanxi Province (No.20210302123216); Shanxi Province Engineering Research Center for Equipment Digitization and PHM (No.ZBPHM20201104); Postgraduate Education Reform Research Project of Shanxi Province (No.2021YJJG244); Postgraduate Education Innovation Project of Shanxi Province (No.2021Y699)

1 引言

在现实问题优化过程中经常存在多个难以权衡且相互冲突的目标,很难直接确定一套理想的优化方案,对于具有两个或两个以上目标的此类评估问题,可划分为多目标优化问题(Multi-objective Optimization Problem, MOP). 针对 MOP 的复杂特性,多目标进化算法(Multi-Objective Evolutionary Algorithms, MOEAs)凭借其高维空间动态编码的优势,相比传统数学优化方法有更大概率得到问题的封闭解,已被广泛应用于相关领域. 在求解 MOP 时,进化算法通过计算目标函数解集进行问题的优化评估,并反复迭代至评估条件结束,因此会对问题解产生大量的真实计算. 然而,许多现实优化问题进行大量真实评估所造成的时间代价将是十分昂贵的,即存在昂贵多目标优化问题(Expensive Multi-objective Optimization Problem, EMOP). 代理辅助进化算法通过建立代理模型施加模型管理准则进行辅助评估逐渐成为求解 EMOP 的有效手段.

代理模型技术即通过建立计算廉价的代理模型代替昂贵的真实计算过程,利用已知样本对未知设计点进行插值与拟合获取到目标估值. 常用的代理模型包括 Kriging 模型、径向基函数模型和神经网络模型等. 由于缺乏代理模型选择方面的理论准则,目前应用于 EMOP 的代理辅助进化算法,大多会优先选择可以为每个预测点提供不确定性估计的 Kriging 模型,如: Knowles 在单目标有效全局优化算法上扩展出的 ParEGO^[1], Zhang 等人立足于减少建模开销所提出的 MOEA/D-EGO^[2], Chugh 等人利用自适应参考向量选择结合 Kriging 模型辅助逼近目标函数所开发的 KRVEA^[3]. 这些领域内代表性算法在求解较低维 MOP 上已被验证其算法有效性,但在求解高维 MOP 时将变得难以计算,原因在于:(1) 随着决策空间维度变高, Kriging 模型需要充足的训练样本来保证其高精度特性,由此将造成模型复杂度呈指数倍增长;(2) Kriging 模型的更新必须完全重建,将很难适用于需频繁更新模型导向的高维问题空间中;(3) Kriging 模型属单输出回归模型,在面对高维目标的 MOP 时,选用 Kriging 模型可能无法准确逼近目标函数.

同时,代理模型辅助进化算法中另一项关键环节在于对模型的管理,即依据填充准则筛选出可用于真实函数计算的个体以及更新扩充训练样本等. 因此,如何基于自身进化算法来设计配套的准则填补解集进而

修正代理模型,是相关领域中学者们致力探寻的. 如孙超利^[4]等人对种群的多样性、个体的收敛性以及估值的不确定度进行综合考虑,基于个体间角度惩罚距离以及估值不确定度提出一种期望计算方法,促进代理模型更有效地辅助优化算法搜索问题的非支配解集. Pan^[5]等人以代理模型作为分类器,将进化算法得到的解分成非支配解和支配解,根据分类器在测试集上的表现结合阈值大小去选择非支配解或是支配解当作真实评估个体. 在上述研究中管理准则大多是基于人为主观偏好设定的,仅通过填充符合设定预期的优越解样本来引导种群偏向某一角度,很有可能会导致算法搜索方向错误. 因此,如何减少人为干预,以算法本身特性为参考,自适应地增加动态更新概念以平衡算法的勘探和开发能力还需要进一步探索.

鉴于此,本文提出一种基于约束型 Dropout 神经网络的代理辅助进化算法,重点解决高维特性下的代理模型扩展与管理问题. 该算法提出一种约束型 Dropout 神经网络(Constrained dropout neural network, C-dropout)作为解决 EMOP 的可扩展代理模型. 通过对 Dropout 神经网络损失项施加约束,克服原先网络的相异随机子模型输出不一致问题,区别于传统在高维目标空间中建立多个代理模型进行预测精度评估的方式, C-dropout 利用便捷的单一网络来建立优化问题的求解模型,新增样本填充机制以及损失约束项来减轻模型的复杂度并提升模型的可信度. 同时,提出一种新颖的模型管理策略用来选择代表个体进行精确评估. 利用成就尺度函数确定候选解中的关键点,以此计算前沿曲率从而实现候选解的收敛性评估. 通过高效的相似度量方式,在无需预先设定参数的前提下衡量种群个体的多样性. 依靠迭代前后个体自身的性能表现自适应选择更新导向,科学选择真实计算个体的同时平衡了算法的勘探和开发能力. 确保优化结果的高效,精准性.

2 基于约束型 Dropout 神经网络的代理辅助进化算法

CDNNEA 算法总体框架如算法 1 所示. 算法中各步骤详细介绍如下.

2.1 初始化

在优化开始前, CDNNEA 使用拉丁超立方体采样方式在决策空间中生成 $11d-1$ 大小的均匀采样点,并对其使用待优化的目标函数进行真实计算,所获得的决

策变量以及相应目标解将被作为数据集训练代理模型. 采样次数设置为 $11d-1$ 的适宜性可参考文献[6].

2.2 代理模型构建

性能优越的代理模型是辅助算法在尽可能少的函数评估次数下快速收敛得到满意解的关键前提. 针对前文讨论中传统代理模型在高维空间中存在的计算复杂度高、拟合能力差、泛化能力低等缺陷, 本文开发并使用了一种新颖的可扩展代理模型作为替代方案, 即约束型 Dropout 神经网络.

算法 1 CDNNEA 总体框架

输入: MaxFE(最大评价次数), FE(评价次数), P (种群大小), d (决策变量个数), f (代价函数), σ (真实评价解的数量), ρ (评判比率)
输出: (X, Y) (真实问题的非支配解集)
1: $X = \text{LatinHypercubeSample}(11d-1)$, $Y = f(X)$; //生成初始解
2: $\text{FE} = 11d-1$;
3: WHILE $\text{FE} \leq \text{MaxFE}$ DO
4: $C\text{-dropout} = \text{TrainingData}(X, Y)$; //代理模型构建
5: $(X_1, Y_1, \rho_1, \rho_2) = \text{Estimate}(P, C\text{-dropout})$; //模型非支配解集计算
6: $X_2 = \text{Selection}(X_1, Y_1, \rho_1, \rho_2, C\text{-dropout}, \sigma)$, $Y_2 = f(X_2)$; //真实计算个体选择
7: $X' = X \cup X_2$, $Y' = Y \cup Y_2$;
8: $(X, Y) = \text{Update}(X', Y', 11d-1, \sigma)$; //代理模型训练集更新
9: $\text{FE} = \text{FE} + 1$, $\rho_1 = \rho_2$;
10: END WHILE

将原始 Dropout 技术^[7]直接用作代理模型会由于网络的复杂性导致与寻优算法无法适配. 因此本文提出的 C-dropout 仅包含有 2 层前馈神经网络 $\{l_{y_1}, l_{y_2}\}$, 每层包含 40 个神经元, 且均使用 Relu 函数进行激活以克服误差梯度消失并加快模型学习效率. 训练过程中采用批次学习方式, 批次大小设置与决策向量个数一致, 保证输入为方阵形式从而避免高维空间中的部分不可逆运算. 对于输入的决策变量值 x_i , 通过正向传播得到预测输出值 \tilde{y}_o . 过程如式(1)~(3), 其中 y_1, y_2 分别表示第一与第二层神经元输出, w_1, w_2, w_3 与 b_1, b_2, b_3 分别表示各层求解所需的权值矩阵与偏置向量, f_R 代表 Relu 函数激活方式.

$$y_1 = f_R(w_1(r_1 \circ x_i) + b_1) \quad (1)$$

$$y_2 = f_R(w_2(r_2 \circ y_1) + b_2) \quad (2)$$

$$\tilde{y}_o = w_3 y_2 + b_3 \quad (3)$$

\tilde{y}_o 计算完成后, 参照训练集中 x_i 的真实输出值 y_o , 获取近似误差进行反向传播并调整对应的权值与偏置. 在网络计算传播过程中, 原始 Dropout 技术对部分神经元随机丢弃的操作将导致每次产生的子模型都是互异的. 考虑到代理模型估值准确性对优化算法寻优朝向的关键作用, 将随机性较强的 Dropout 神经网络直接用作代理模型显然是不适宜的, 因此本文以缩减不

同子模型输出间分布的差异性为出发点, 新增样本填充机制以及损失约束项来提升 EMOP 求解过程中代理模型的可信度. 对于批次大小为 d (与决策变量大小一致) 的样本组成的训练集 $T = \{(X_i, Y_i)\} (i=1, 2, \dots, d)$, 原始进行反向传播的目的即最小化式(4)所示的均方误差函数, 式中 y_o 表示样本真实值, \bar{y}_o 表示样本预测值. 而在 C-dropout 过程中, 每个批次的训练数据 X_i 都将被复制后堆叠为新样本再输入, 其目的在于模拟相同数据的两次正向传播过程以获取模型预测结果的两个分布 $P_1 = P(X_i | Y_i)$ 与 $P_2 = P'(X_i | Y_i)$, 同时凭借 Dropout 网络结构动态变化的特性, 堆叠新样本的方式可对昂贵优化问题中稀缺的真实计算样本进行一定程度上的数据扩充. 此时误差损失最小化的目标则变化为式(5), 且对于模型差异性的缩减问题便转化为如何对输出后的分布 P_1 与 P_2 进行约束. 结合斯皮尔曼相关系数对变量间相关程度的有效度量^[8], 本文将其应用于衡量两次输出的不一致性, 即构造式(6)所示的约束项, 并与 l_2 联立成最终的训练损失函数如式(7)所示, 式中 α 表示约束项所占比例系数, 以此降低原始网络空间中参数的自由度. 求取误差后, 模型将按照所设迭代次数, 结合由梯度与链式法更新后的权值和偏差进行反向传播, 最终完成训练阶段.

$$l_1 = \frac{1}{2} \left(\sum_{i=1}^d (\bar{y}_o - y_o) \right) \quad (4)$$

$$l_2 = \frac{1}{2} \left(\sum_{i=1}^d (\bar{y}_o - y_o) + \sum_{j=1}^d (\bar{y}'_o - y'_o) \right) \quad (5)$$

$$l_c = \text{Spearman}(P_1, P_2) = \frac{\sum_{i=1}^d (P_1 - \bar{P}_1)(P_2 - \bar{P}_2)}{\sqrt{\sum_{i=1}^d (P_1 - \bar{P}_1)^2 \sum_{i=1}^d (P_2 - \bar{P}_2)^2}} \quad (6)$$

$$l = l_2 + \alpha \cdot l_c, \alpha = \left| \frac{l_2 - l_1}{l_2} \right| \quad (7)$$

在优化算法寻优过程中, 除填充真实评估个体外的所有原问题的求解过程都将利用 C-dropout 测试阶段进行拟合逼近. 值得注意的是, 在对 I_{test} 个解决方案目标值近似等价时, 若采用传统测试阶段的 Dropout 关闭操作, 将额外保存 I_{test} 个神经网络的权重参数且需随着训练集数据的改变而频繁更新权重信息. 因此, C-dropout 中的 Dropout 方式在训练阶段和测试阶段均处于开启状态. 当测试阶段迭代完成后, 种群中相应个体 β 的适应度 \tilde{y}_β 以及相应的置信度 \tilde{s}_β 计算方式如式(8).

$$\tilde{y}_\beta = \frac{\sum_{i=1}^{I_{\text{test}}} \tilde{y}_i}{I_{\text{test}}}, \tilde{s}_\beta = \sqrt{\frac{1}{I_{\text{test}}} \sum_{i=1}^{I_{\text{test}}} (\tilde{y}_i)^2 - (\tilde{y}_\beta)^2} \quad (8)$$

2.3 模型非支配解集计算

为了最大化高维问题中代理模型所求解集与原始费时问题需求解的等价程度,提高算法中非支配最优解集求解效率,本文使用具有良好高维多目标问题求解性能的PeEA算法^[9]进行代理模型的最优解集搜索.区别于标准求解多目标优化问题的进化算法流程. PeEA在自身环境选择策略中首先利用成就尺度函数定位最能表征Pareto前沿曲率的关键点,计算点间距离的比值来近似获取前沿曲率.之后根据估计出的曲率值由构建的自适应尺度函数来选择合适的指标实现算法的最大化收敛.同时利用独特的相似度量方式提升高维环境下存在异常值时解集的一致性.在PeEA中,对于 m 目标维度优化问题中的前沿面上点 $x=(x_1, x_2, \dots, x_m)$,则有式(9)的Pareto前沿形状估计.

$$\sum_{i=1}^m x_i^q = 1 \quad (9)$$

其中 q 为代表前沿曲率的正参数,对于 $0 < q < 1$ 、 $q = 1$ 、 $q > 1$ 分别表示前沿属凹形、线性、凸形.在确定 q 值时,PeEA首先利用每个目标上最小值点与极值点实现目标标准化,以期提供与目标轴等距的超平面作为基平面.其中极值点由令式(10)所示的成就尺度函数最小的个体解确定.

$$S_\infty(x|W, z^*) = \max_{j=1}^m \frac{f_j'(x)}{w_{ij}} \quad (10)$$

式(10)中 z^* 表示在 m 维度目标空间中搜索到的最小值点, $f_j'(x)$ 个体对应目标解在 z^* 上各个方向上的投影. $W=(w_{11}, w_{12}, \dots, w_{ij}, w_{im})$ 代表相应目标轴的方向,且当 $i=j$ 时 $w_{ij}=1$,其它情景下 w_{ij} 设为极小标量 10^{-6} . $f_j'(x)$ 与不同的 w 作量化比较则计算出个体在何种目标上所占比重大,最后选取式(10)中最小的个体作为最具曲率代表意义的点即极值点.如图1所示,此时Pareto前沿曲率则可由最小值点 $O(z^*$ 归一化后形成的原点)到基平面间的距离 d^\perp 与 O 到极值点 D 的距离 d 的比值确定.若 $d < d^\perp$,则 $0 < q < 1$,此Pareto前沿可被看作是凹形.若 $d > d^\perp$,则 $q > 1$,则前沿面将是凸形. $q = 1$ 时,前沿面则呈线性.

确定曲率值后,PeEA将以 z^* 作为参考点,结合曲率信息,采用式(11)所示的 L_p -metric评估出解的收敛性 \tilde{c} .通过测量参考点与可行目标区域间的距离,提供适宜的标量值来辅助解的适应度估计.

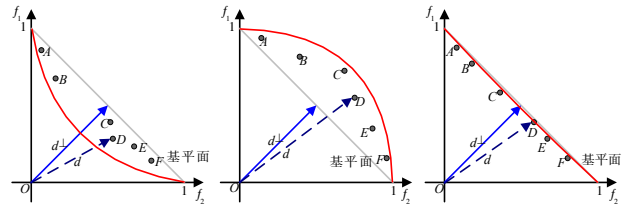
$$\tilde{c}(x) = \left(\sum_{i=1}^m |f_i(x) - z_i^*|^p \right)^{\frac{1}{p}} \quad (11)$$

式中 p 的值随先验知识 q 的不同而改变.当 $0 < q \leq 1$ 时,则 $p = 1$ 且式(11)将转换为式(12)的线性缩放形式,抽象在二维空间中,即横纵坐标和相等的点汇聚而成的面,用以衡量前沿面近似为凹型或线型的情况.当 $q > 1$ 时,

则 $p = \infty$ 且式(11)将转换为式(13)的切比雪夫距离,抽象在二维空间中,即取横、纵坐标中最大值的点,用以衡量前沿面近似为凸型的情况.

$$\tilde{c}(x) = \sum_{i=1}^m (f_i(x) - z_i^*) \quad (12)$$

$$\tilde{c}(x) = \max_{i=1,2,\dots,m} (f_i(x) - z_i^*) \quad (13)$$



(a) 前沿凹形 (b) 前沿凸形 (c) 前沿线形

图1 曲率估计示意图(红色线代表Pareto前沿大致形状,A-F为帕累托上部分解决方案)

衡量解集多样性时,PeEA引入了一种计算成本较低的相似度量方式,通过计算解集相互间的维度边际距离,缩减维度增加过程中各维度上解的差异对评估结果的影响.假设存在两个向量 $a, b \in S$,则 a, b 间的维度边际距离定义如式(14).

$$d(a, b) = \sum_{i=1}^m \frac{\max(a_i, b_i) - \min(a_i, b_i)}{1 + \min(a_i, b_i)} \quad (14)$$

对于一组大小为 N 的解集 X ,PeEA将计算出 X 中某个解与其余 $\sqrt{|N|}$ 个最近解的维度边际距离 $d(x_N)$.之后此解的多样性 \tilde{d} 计算则如式(15)所示.

$$\tilde{d}(x) = \frac{1}{d(x_N) + 2} \quad (15)$$

求取 \tilde{c} 与 \tilde{d} 后,PeEA则利用两者之和作为新的适应度评估函数进行随后的环境选择,种群与后代的联合个体中和值较小的解则被挑选出作为候选解进行种群下一轮的迭代优化.

2.4 模型管理

模型管理部分即从PeEA搜索到的候选解中选择若干个体进行真实评估,经评估后的解一方面作为填充个体更新训练集从而提升代理模型在EMOP中的近似精度,另一方面则充当EMOP中最优非支配解集的潜在候选个体.

在选择真实计算个体方面.首先利用曲率相关信息以及维度边际距离获得初始种群的收敛性信息 \tilde{c}_1 与多样性信息 \tilde{c}_2 ,并求得两者之比 ρ_1 .同时统计PeEA搜索得到种群 P_c 时最后一代个体的收敛性信息 \tilde{c}_2 与多样性信息 \tilde{d}_2 ,以及个体适应度 \tilde{y}_2 以及相应的置信度 \tilde{s}_2 ,计算出 \tilde{c}_2 与 \tilde{d}_2 之比 ρ_2 .之后使用Mini Batch K-Means聚类^[10]算法将种群 P_c 分为 σ 簇,分簇的目的在于过滤掉

属性较相似的解,使得从每簇中选择出的个体更具代表性,而使用 Mini Batch K-Means 聚类算法则更能在仅使用小规模数据子集的前提下迅速将种群个体划分至所属类别,进一步缩减高维 EMOP 在时间成本上的消耗.接着通过对比 ρ_1 与 ρ_2 两项值的大小,若 $\rho_1 > \rho_2$ 时,则表明种群 P_c 较初始种群的收敛性较好,此时则选择每簇中置信度 \hat{s}_2 较大的个体进行精确评估从而提升种群的多样性.若 $\rho_1 < \rho_2$ 时,则表明种群 P_c 较初始种群的多样性较好,应对其收敛性进行提升,此时则利用式(16)选择每簇中离目标空间原点欧式距离 d_o 最小的个体进行精确评估.通过所建立的个体填充方式,可在无需设定额外参数的情况下自适应的动态兼顾收敛性和多样性两项指标,从而协调快速开发和深入勘探之间的平衡,以确保高效、精准的优化结果.

$$d_o = \sqrt{\sum_{i=1}^m (\tilde{y}_2^i)^2} \quad (16)$$

对于代理模型训练集 (X, Y) 的更新,在填充时需优先考虑从 σ 组簇中选出的真实计算代表个体.由于训练集大小限制为 $11d-1$,故对其余 $11d-1-\sigma$ 数量个体选择时,将以式(17)近似 PeEA 算法所获非支配解集目标值 f_p^i 与 σ 大小的真实计算个体目标值 f_t^i 间的夹角大小 θ , θ 最大的解则会被选取以保证训练集在目标空间中的均匀分布.由此随着算法不断迭代,训练集中的样本会逐渐靠近真实的 Pareto 前沿,进而提升代理模型对真实问题的拟合性能.

$$\theta = \min \left[\frac{\sum_{i=1}^m f_p^i f_t^i}{\sqrt{\sum_{i=1}^m (f_p^i)^2} \sqrt{\sum_{i=1}^m (f_t^i)^2}} \right], t = 1, 2, \dots, \sigma \quad (17)$$

3 实验结果与分析

3.1 测试问题集

实验选用 DTLZ^[11] 测试问题作为基准测试问题.对于每个测试问题,最大的决策变量设置为 100. 本文实验均在 Intel Core i5-9400F CPU, Microsoft Windows 10 操作系统的计算机上由 MATLAB R2020b 软件设计实现,其中对比算法的结果在 PlatEMO 平台^[12] 上运行得到.

3.2 参数设置

(1)独立运行次数:对于每个测试问题,每种算法独立运行 20 次.

(2)最大真实评估次数:除初始真实采样的 $11d-1$ 个训练数据外,另外采用 120 次真实求解来测试比较算法的性能,因此最大真实评估次数设置为 $11d+119$.

(3)每代选取真实计算解的数量:根据 Mini Batch K-Means 聚类簇数设置为 5.

(4)C-dropout 相关参数:两层神经元失活概率值 p_1 与 p_2 分别为 0.2 与 0.5,学习率大小设为 0.01,批次大小设置与决策变量个数一致,模型训练迭代次数 I_{train} 设为 1×10^4 ,测试迭代次数 I_{test} 设为 100.

(5)PeEA 算法参数:种群大小设置为 50,对代理模型最大评估次数设置为 20 次.

(6)变异算子:采用多项式突变方式,设变异概率为 $1/d$,分布指数为 20.

(7)交叉算子:采用模拟二进制交叉方式,设交叉概率为 1.0,分布指数为 20.

3.3 评价指标

实验选用反向世代距离评价指标 (Inverted Generational Distance, IGD)^[13] 作为评价指标.相关计算公式如式(18)所示.

$$\text{IGD}(P, Q) = \frac{\sum_{v \in P} \text{dist}(v, Q)}{|P|} \quad (18)$$

式中, P 为均匀分布在真实 Pareto 前沿上的目标向量集合, Q 为经算法计算得到的目标向量集合. $\text{dist}(v, Q)$ 表示 P 中的一个目标向量 $v \in P$ 到 Q 中最近向量的欧式距离.

同时,本文采用置信度水平为 0.05 的 Wilcoxon 秩和检验来对 CDNNEA 与其它对比算法的性能差异性进行统计检验,符号“+”、“-”以及“=”分别表示 CDNNEA 算法性能相较于对比算法好、差以及没有明显差异.

3.4 CDNNEA 与先进算法的对比实验结果

本节验证 CDNNEA 性能时所对比的算法包括两种未使用代理辅助的进化算法 (ARMOEA^[14], PeEA) 以及四种具备代表性的代理辅助进化算法 (EDN-ARMOEA, ParEGO, K-RVEA, MOEA/D-EGO). 其中代理辅助进化算法初始化方法上均设置为真实采样 $11d-1$ 大小的训练数据, ARMOEA 与 PeEA 算法设置与代理辅助进化算法相同的最大真实评估次数即 $11d+119$. 性能比较主要针对决策空间维度对算法的影响,故在 DTLZ 测试集上开展了固定目标数量 $m=3$ 并改变决策变量维度 d (20, 40, 60, 100) 的实验. 七种算法最终取得的 IGD 平均值结果如表 1 所示. 可以发现,面对共 28 组测试问题, CDNNEA 在 14 组上相比其余对比算法有最优的性能表现,在 7 组上性能表现近似最优,且相关情形均属测试问题中决策变量较高维,表明 CDNNEA 相比其它算法在高维问题空间中具备较优的求解能力.

表1 CDNNEA与六种对比算法在 $m=3$ DTLZ测试问题上获得的平均IGD值

Problem	d	ARMOEA	PeEA	EDN -ARMOEA	ParEGO	K-RVEA	MOEA/D -EGO	CDNNEA
DTLZ1	20	3.039 6e+2 (3.28e+1)=	3.244 6e+2 (4.30e+1)=	3.214 6e+2 (3.37e+1)=	3.277 5e+2 (5.19e+1)=	3.396 0e+2 (1.82e+1)=	2.991 5e+2 (5.46e+1)=	3.008 9e+2 (3.40e+1)
	40	1.169 0e+3 (2.12e+1)+	8.998 8e+2 (4.62e+1)=	8.599 3e+2 (5.83e+1)=	9.922 3e+2 (5.82e+1)+	9.874 1e+2 (3.17e+1)+	8.927 3e+2 (5.26e+1)=	8.766 1e+2 (5.83e+1)
	60	1.9312e+3 (1.18e+1)+	1.9024e+3 (1.61e+1)+	1.4752e+3 (2.59e+1)+	1.5091e+3 (1.55e+1)+	1.4394e+3 (4.81e+1)=	1.4941e+3 (1.16e+2)+	1.3995e+3 (4.81e+1)
	100	3.3671e+3 (7.51e+2)+	3.6548e+3 (6.66e+2)+	2.2793e+3 (4.94e+1)-	2.4894e+3 (7.31e+1)=	2.7303e+3 (8.11e+1)+	2.7144e+3 (8.59e+1)+	2.4891e+3 (7.24e+1)
DTLZ2	20	8.933 6e-1 (8.95e-2)+	8.541 8e-1 (9.31e-2)+	6.974 0e-1 (6.47e-2)=	7.119 3e-1 (6.68e-2)=	5.207 1e-1 (8.94e-2)-	6.677 5e-1 (9.41e-2)-	7.0417e-1 (6.32e-2)
	40	2.533 6e+0 (3.51e-1)+	2.979 3e+0 (3.69e-1)+	1.972 8e+0 (1.26e-1)=	2.577 6e+0 (2.20e-1)+	2.073 1e+0 (1.37e-1)+	2.109 3e+0 (1.49e-1)+	1.861 5e+0 (1.18e-1)
	60	3.618 4e+0 (2.51e-1)+	3.593 8e+0 (2.97e-1)+	3.296 7e+0 (1.33e-1)=	3.519 4e+0 (2.14e-1)+	3.405 7e+0 (1.41e-1)+	3.422 2e+0 (1.07e-1)+	3.173 9e+0 (1.32e-1)
	100	6.640 8e+0 (4.09e-1)+	6.726 3e+0 (4.50e-1)+	5.744 6e+0 (1.29e-1)=	5.761 3e+0 (2.29e-1)=	5.700 8e+0 (1.16e-1)=	6.127 3e+0 (2.09e-1)+	5.746 8e+0 (1.47e-1)
DTLZ3	20	1.554 8e+3 (2.94e+2)+	1.436 4e+3 (1.51e+2)+	1.094 6e+3 (1.47e+2)+	1.141 0e+3 (2.35e+2)+	1.163 3e+3 (2.14e+2)+	6.034 0e+2 (8.49e+2)-	8.656 1e+2 (1.36e+2)
	40	2.938 6e+3 (2.16e+2)+	2.974 9e+3 (2.14e+2)+	2.788 4e+3 (1.66e+2)+	2.852 7e+3 (1.27e+2)+	2.992 0e+3 (1.19e+2)+	2.851 4e+3 (1.27e+2)+	2.578 8e+3 (6.70e+1)
	60	4.836 6e+3 (2.73e+2)+	4.738 6e+3 (3.95e+2)+	4.548 0e+3 (2.04e+2)+	4.638 1e+3 (2.77e+2)=	4.434 2e+3 (3.30e+2)=	4.546 6e+3 (1.01e+2)+	4.305 5e+3 (1.14e+2)
	100	8.832 8e+3 (2.01e+2)+	9.152 7e+3 (2.67e+2)+	8.615 3e+3 (3.97e+2)+	8.113 6e+3 (2.94e+2)+	8.043 7e+3 (2.09e+2)+	8.263 0e+3 (1.41e+1)+	7.891 4e+3 (1.81e+2)
DTLZ4	20	1.145 7e+0 (7.26e-1)=	1.244 2e+0 (7.96e-1)=	8.477 4e-1 (1.83e-1)-	1.130 8e+0 (6.09e-1)=	9.936 4e-1 (7.80e-1)=	1.173 4e+0 (6.86e-1)=	1.101 5e+0 (8.72e-1)
	40	2.855 4e+0 (2.37e-1)+	2.913 7e+0 (2.09e-1)+	2.195 8e+0 (1.74e-1)=	2.660 8e+0 (1.96e-1)+	2.638 1e+0 (1.06e-1)+	2.620 7e+0 (2.42e-1)+	2.222 1e+0 (1.84e-1)
	60	3.685 1e+0 (2.42e-1)+	3.922 7e+0 (2.13e-1)+	3.608 2e+0 (1.30e-1)+	3.402 0e+0 (1.87e-1)=	3.851 0e+0 (1.22e-1)+	3.744 7e+0 (1.26e-1)+	3.360 9e+0 (1.99e-1)
	100	6.608 2e+0 (4.96e-1)+	6.116 3e+0 (1.77e-1)=	6.047 1e+0 (1.37e-1)=	6.098 2e+0 (5.16e-1)=	6.068 8e+0 (2.07e-1)=	6.103 6e+0 (1.77e-1)=	6.065 3e+0 (1.84e-1)
DTLZ5	20	5.174 9e-1 (6.16e-2)+	5.132 5e-1 (6.46e-2)+	4.464 5e-1 (7.86e-2)+	4.082 7e-1 (8.09e-2)+	4.778 5e-1 (7.91e-2)+	5.260 2e-1 (6.47e-2)+	3.207 6e-1 (5.22e-2)
	40	2.222 6e+0 (7.83e-1)+	2.081 5e+0 (7.60e-1)+	2.062 3e+0 (1.77e-1)+	2.174 0e+0 (1.16e-1)+	1.873 1e+0 (1.44e-1)=	2.038 8e+0 (9.14e-1)+	1.793 6e+0 (9.44e-1)
	60	3.208 0e+0 (2.11e-1)+	3.526 3e+0 (2.09e-1)+	3.221 7e+0 (1.60e-1)+	3.332 6e+0 (1.93e-1)+	3.353 3e+0 (1.77e-1)+	3.237 7e+0 (1.61e-1)+	3.023 3e+0 (1.74e-1)
	100	7.017 2e+0 (3.94e-1)=	7.130 8e+0 (6.56e-1)=	5.608 6e+0 (1.70e-1)=	5.638 0e+0 (2.66e-1)=	5.741 7e+0 (1.63e-1)=	5.577 9e+0 (1.92e-1)=	5.606 5e+0 (1.16e-1)
DTLZ6	20	1.960 8e+1 (6.13e-1)+	1.975 7e+1 (4.07e-1)+	1.330 9e+1 (7.07e-1)=	7.011 5e+0 (4.52e-1)-	1.199 4e+1 (2.70e-1)=	6.974 4e+0 (3.47e-1)-	1.391 2e+1 (7.06e-1)
	40	3.803 1e+1 (7.61e-1)+	3.747 0e+1 (7.11e-1)+	3.153 6e+1 (4.30e-1)=	3.372 9e+1 (4.77e-1)+	2.906 6e+1 (4.90e-1)-	3.298 8e+1 (7.16e-1)+	3.145 1e+1 (4.17e-1)
	60	5.260 8e+1 (6.16e-1)+	5.211 3e+1 (7.29e-1)+	4.952 9e+1 (7.15e-1)+	5.112 6e+1 (5.16e-1)+	5.098 5e+1 (4.42e-1)+	5.084 9e+1 (5.36e-1)+	4.789 9e+1 (6.55e-1)

续表

Problem	<i>d</i>	ARMOEA	PeEA	EDN -ARMOEA	ParEGO	K-RVEA	MOEA/D -EGO	CDNNEA
	100	8.927 1e+1 (7.37e-1)+	8.782 0e+1 (6.93e-1)+	8.477 2e+1 (6.28e-1)=	8.819 2e+1 (6.77e-1)+	8.764 9e+1 (2.68e-1)+	8.252 7e+1 (4.10e-1)-	8.443 7e+1 (6.24e-1)
DTLZ7	20	2.323 6e+0 (7.51e-1)=	2.680 1e+0 (5.53e-1)=	2.958 8e+0 (7.94e-1)=	6.900 1e+0 (7.42e-1)+	3.845 7e-1 (1.16e-1)+	6.113 0e+0 (8.96e-1)+	2.659 6e+0 (3.71e-1)
	40	8.711 8e+0 (7.37e-1)+	8.605 5e+0 (8.96e-1)+	4.556 2e+0 (9.50e-1)-	8.132 7e+0 (8.16e-1)+	8.260 8e+0 (9.52e-1)+	9.162 3e+0 (8.17e-1)+	5.606 4e+0 (8.94e-1)
	60	9.736 9e+0 (3.66e-1)+	9.460 8e+0 (5.70e-1)+	7.522 6e+0 (5.83e-1)-	9.441 5e+0 (5.42e-1)+	8.407 3e+0 (5.77e-1)=	8.398 4e+0 (8.10e-1)=	8.257 7e+0 (6.55e-1)
	100	6.802 4e+0 (5.84e-1)+	6.721 0e+0 (5.28e-1)+	6.029 8e+0 (6.35e-1)+	6.281 4e+0 (4.27e-1)+	5.801 9e+0 (4.68e-1)=	6.013 6e+0 (3.16e-1)+	5.756 5e+0 (4.37e-1)
		+/-/=	24/0/4	22/0/6	11/4/13	18/1/9	15/2/11	18/4/6

3.5 算法在实际问题上的可行性

考虑到 CDNNEA 面对高维昂贵类问题时表现出的优越性能, 以及其与 Pareto 前沿未知的神经网络结构参数优化问题具备的较强适配性, 将 CDNNEA 应用于对 MNIST 手写数字数据集识别模型^[15]的高维参数自适应优化问题中以验证算法解决实际问题的能力。

首先对于 MNIST 手写数字数据集则按 6:1 比例划分好训练集与测试集并预设如图 2 所示的三层卷积神经网络模型。之后明确 CDNNEA 算法的决策变量维度即网络模型中待优化的参数。对于 MNIST 数据集识别模型共包含 17 个参数, 即: 三层卷积核个数(Channel_1、Channel_2、Channel_3), 三层卷积核尺寸(C_s_1、C_s_2、C_s_3), 卷积层中三个激活函数(Act_1、Act_2、Act_3), 三层池化方式(pool_1、pool_2、pool_3), 两层全连接层节点数(f_1、f_2), 梯度下降函数(Op), 学习率(L_rate), 批次大小(batch)。其中, 除激活函数、池化方式、学习率、梯度下降函数外的待定参数由于信息的整数规划问题, 需要对其第一位小数进行取整操作。接着

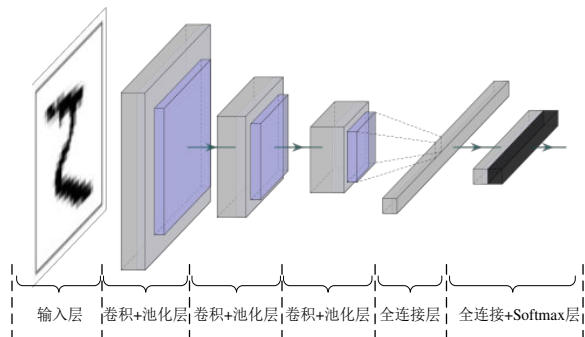


图2 MNIST 识别模型结构图

进入到 CDNNEA 算法优化流程, 在每次迭代中, 对网络参数进行优化, 以寻得网络模型的测试时间、计算量、测试误差即三个目标的最小值。实例中, 测试时间即对测试集识别过程所需的总时间, 计算量则以模型的每秒浮点运行次数(Floating Point Operations Per second, FLOPs)来衡量, 测试误差以式(19)计算, 有

$$l_{val} = \frac{1}{\text{batch}} \sum_{i=1}^{\text{batch}} \frac{|t(i) - p(i)|}{t(i)} \quad (19)$$

式(19)中, $t(i)$ 代表真实标签, $p(i)$ 为预测标签, batch 为批量大小。实验设置 CDNNEA 算法种群总数为 20, 最大评估次数为 216, MNIST 数据集识别模型迭代轮数设为 20。图 3 为算法经过最后一次迭代后在识别模型中搜索到的解决方案分布图。从图中可以明显观察得到, 在 A 点处, 网络测试时间最短, B 点处, 网络计算量最小, C 点处, 网络有着最低的测试误差。以 A、B、C 三项解决方案为代表。将三组方案带入识别模型中得到的测试精度曲线如图 4 所示, 可以发现在 MNIST 数据集识别模型优化过程中 A、B、C 识别准确度均处于

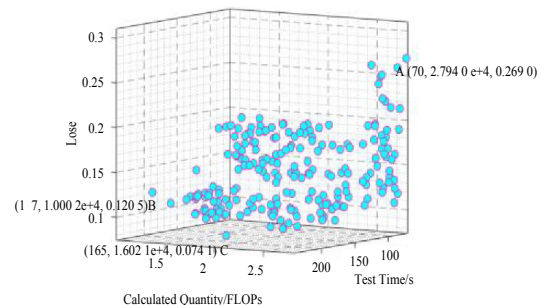


图3 识别模型参数优化解决方案分布图

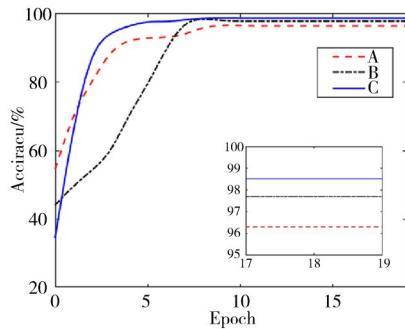


图4 三组解决方案的识别精度曲线

表2 与六种对比算法应用于识别模型上的性能对比

评价指标	ARMOEA	PeEA	EDN-ARMOEA	ParEGO	K-RVEA	MOEA/DEGO	CDNNEA
计算时间/h	17.14	14.26	5.82	7.15	7.24	8.24	5.13
最高精度/%	97.90	96.85	97.20	97.75	98.70	97.25	98.40

4 结论

本文提出了一种基于约束型 Dropout 神经网络的代理辅助进化算法(CDNNEA)用于解决昂贵多目标问题. CDNNEA 首先采用 C-dropout 作为代理模型, 保证模型高拟合性以及泛化性的同时降低了计算复杂度. 其次, 以具备良好高维多目标问题求解性能的 PeEA 算法进行代理模型的最优解集搜索, 在 Pareto 前沿形状、连续性等特点未知的情况下提供出可用性较强的候选解. 最后, 设计与代理模型匹配的模型管理策略, 依据种群个体自身的收敛性以及多样性表现自适应选择更新导向, 挑选出每代需使用真实函数计算的个体引导算法进行迭代更新. 在与其他六种多目标进化算法于多组基准测试问题上的实验对比表明, CDNNEA 具备良好的高维空间求解性能. 在应用于如神经网络参数寻优的昂贵类现实问题时, CDNNEA 能提供出可用性较强的多组解决方案, 具备实际应用能力.

算法中, 对 C-dropout 代理模型的训练在迭代次数上设置为固定的数值, 虽在大多数高维问题上性能优于传统模型, 但当样本量过多或过少时, 可能会出现计算冗余或与真实函数拟合不充分等问题. 为此, 如何进一步提升 C-dropout 在不均衡样本量背景下的性能需要进一步展开研究.

参考文献

- [1] KNOWLES J. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems[J]. IEEE Transactions on Evolutionary Computation, 2006, 10(1): 50-66.
- [2] ZHANG Q, LI H. MOEA/D: A multiobjective evolution-ary algorithm based on decomposition[J]. IEEE Transactions on Evolutionary Computation, 2007, 11(6): 712-731.
- [3] CHUGH T, JIN Y, MIETTINEN K, et al. A surrogate-assisted reference vector guided evolutionary algorithm for computationally expensive many-objective optimization [J]. IEEE Transactions on Evolutionary Computation, 2018, 22(1): 129-142.
- [4] 孙超利, 李贞, 等. 模型辅助的计算费时进化高维多目标优化[J]. 自动化学报, 2022, 48(4): 1119-1128.
SUN Chao-li, LI Zhen, et al. Surrogate-assisted expensive evolutionary many-objective optimization[J]. Acta Automatica Sinica, 2022, 48(4): 1119-1128. (in Chinese)
- [5] PAN L, HE C, TIAN Y, et al. A classification-based surrogate-assisted evolutionary algorithm for expensive many-objective optimization[J]. IEEE Transactions on Evolutionary Computation, 2019, 23(1): 74-88.
- [6] GUO D, WANG X, et al. Evolutionary optimization of high-dimensional multiobjective and many-objective expensive problems assisted by a dropout neural network[J]. IEEE Transactions on Systems, Man and Cybernetics, 2022, 52(4): 2084-2097.
- [7] MCKAY M, BECKMAN R, CONOVER W. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code[J]. Technometrics, 2000, 42(1): 55-61.
- [8] SUN W, HU X, ZHANG L. Moments of discounted aggregate claims with dependence based on Spearman copula[J]. Journal of Computational and Applied Mathematics, 2020, 377: 112889.
- [9] LI L, YEN G, SAHOO A, et al. On the estimation of Pare-

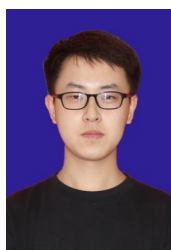
- to front and dimensional similarity in many-objective evolutionary algorithm[J]. Information Sciences, 2021, 563: 375-400.
- [10] SLAPAK E, GAZDA J, BUGAR G, et al. HetNet spatial topology design using mini-batch k -means clustering[C]// Proceedings of the 2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems. Piscataway: IEEE, 2019: 1-4.
- [11] 刘冰洁, 毕晓君. 一种基于角度信息的约束高维多目标进化算法[J]. 电子学报, 2021, 49(11): 2208-2216.
LIU Bing-jie, BI Xiao-jun. A constrained many-objective evolutionary algorithm based on angle information[J]. Acta Electronica Sinica, 2021, 49(11): 2208-2216. (in Chinese)
- [12] TIAN Y, CHENG R, ZHANG X, et al. PlatEMO: A MATLAB platform for evolutionary multi-objective optimization[J]. IEEE Computational Intelligence Magazines, 2017, 12(4): 73-87.
- [13] DEB K, GOYAL M. A combined genetic adaptive search for engineering design[J]. Computer Science and Informatics, 1996, 26(4): 30-45.
- [14] TIAN Y, CHENG R, ZHANG X, et al. An indicator-based multiobjective evolutionary algorithm with reference point adaptation for better versatility[J]. IEEE Transactions on Evolutionary Computation, 2018, 22(4): 609-622.
- [15] DENG L. The MNIST database of handwritten digit images for machine learning research [Best of the Web][J]. IEEE Signal Processing Magazine, 2012, 29(6): 141-142.

作者简介



张睿 男, 1987年2月出生于山西省太原市. 现为太原科技大学计算机科学与技术学院副教授、硕士生导师. 主要从事智能信息处理、自动机器学习与群体智能优化.

E-mail: zhangrui@tyust.edu.cn



白晓露(通讯作者) 男, 1999年2月出生于山西省大同市. 现为太原科技大学计算机科学与技术学院硕士研究生. 主要研究方向为群体智能优化算法.

E-mail: bxl17335866519@163.com