

基于迁移代理模型的黑盒优化算法

王志强^{1,2}, 乔振阳³, 张云鹏^{1*}, 傅为农³

(1. 中国科学院深圳先进技术研究院, 广东 深圳 518055;

2. 中国科学院大学, 北京 101408;

3. 深圳理工大学 计算机科学与技术工程学院, 广东 深圳 518107)

Black-box Optimization Algorithm based on Transfer Surrogate Model

Wang Zhiqiang^{1,2}, Qiao Zhenyang³, Zhang Yunpeng^{1*}, Fu Weinong³

(1. Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China;

2. University of Chinese Academy of Sciences, Beijing 101408, China;

3. Faculty of Computer Science and Control Engineering, Shenzhen University of Advanced Technology, Shenzhen 518107, China)

Abstract: [Objective] Black-box optimization algorithms are widely used in various complex engineering design scenarios. However, their inherent computational complexity and internal operational opacity result in intensive computing consumption. In the optimal design of motors, black-box optimization algorithms require numerous iterative computations and finite element simulations to screen optimal structural parameters, which significantly extends the design cycle and restricts engineering application efficiency. Although traditional surrogate models can reduce computational costs, they exhibit poor adaptability and struggle to construct universal databases for multiple operating conditions. Slight changes in design requirements or application scenarios necessitate data recollection and model retraining, leading to inherent engineering drawbacks including weak generalization capability and low data reusability. [Methods] To solve the above problems, a black-box optimization algorithm based on transfer surrogate model (TSM) was proposed. Firstly, a novel TSM was constructed based on transfer learning theory. By autonomously screening high-quality and effective samples from the source domain, the inherent data features were deeply mined to improve the prediction accuracy and generalization ability of the model under the small-sample conditions of the target domain. Subsequently, the optimized TSM was deeply integrated with the iterative optimization algorithm to eliminate redundant computational steps during

the optimization process and further enhance the execution efficiency of the optimization task. Finally, simulation tests were conducted to verify the practical effectiveness of the proposed algorithm. [Results] The simulation results showed that, compared with the traditional optimization algorithm, the proposed optimization algorithm based on TSM effectively reduced computational consumption, significantly improved the overall efficiency of motor parameter optimization, and accelerated the iterative convergence speed of the model. The proposed algorithm exhibited superior stability under variable motor design conditions. [Conclusion] The proposed optimization algorithm based on TSM addresses the inherent shortcomings of traditional black-box optimization and conventional surrogate models from the perspectives of data utilization and algorithm integration. It breaks through the limitations of poor scenario adaptability and high retraining costs existing in traditional models. This scheme can efficiently adapt to the variable design requirements of practical engineering scenarios. It provides reliable and efficient technical support for the intelligent optimal design of motors and other complex engineering black-box optimization problems, and possesses excellent engineering applicability and popularization prospects.

Key words: black-box optimization; surrogate model; transfer learning; optimal design of motor

摘要: [目的] 黑盒优化算法广泛应用于各类复杂工程设计场景,但其固有的运算过程复杂、内部逻辑不透明,导致算力消耗大。在电机优化设计过程中,黑盒优化算法通过反复迭代运算和有限元仿真筛选最优结构参数,

基金项目: 国家自然科学基金(52407010)

National Natural Science Foundation of China(52407010)

大幅拉长了设计周期,制约工程应用效率。传统代理模型虽能降低算力成本,但适配性较差,难以搭建适用于多工况的通用数据库,设计需求和应用场景出现小幅变动后,需重新采集数据、训练模型,存在泛化能力弱、数据复用率低的工程缺陷。【方法】针对上述问题,本文提出一种基于迁移代理模型(TSM)的黑盒优化算法。首先,依托迁移学习理论搭建全新TSM,通过模型自主筛选源域优质有效样本,深度挖掘数据内在特征,提升模型在目标域小样本条件下的预测精度与泛化能力。然后,将优化后的TSM与迭代优化算法深度融合,删减优化过程中的冗余计算步骤,进一步提升目标优化任务的执行效率。最后,通过仿真测试所提算法的实际应用效果。【结果】仿真结果表明,相较于传统优化算法,本文所提基于TSM的优化算法可有效降低算力损耗,显著提升电机参数优化的整体工作效率,同时能够加快模型迭代收敛速度,在多变的电机设计工况中具备更稳定的表现。【结论】本文所提基于TSM的优化算法,从数据利用与算法融合层面解决了传统黑盒优化算法和常规代理模型的固有短板,突破了传统模型场景适配性差、重复训练成本高的局限。所提算法能高效适配实际工程的多变设计需求,为电机智能优化设计以及同类复杂工程黑盒优化问题提供可靠、高效的技术支撑,具备极高的工程落地价值与推广前景。

关键词: 黑盒优化;代理模型;迁移学习;电机优化设计

0 引言

黑盒优化是一类无明确数学解析表达式的优化问题,研究人员仅能获取并分析目标函数的输入与输出数据。由于黑盒优化问题的内在机理无法直接解析,黑盒优化算法只能依靠反复迭代更新种群,逐步筛选出更优质的子代个体。与此同时,这类优化问题通常需要依托复杂的仿真分析与数值计算,运算成本极高,大幅限制了黑盒优化算法在实际工程场景中的落地应用^[1-2]。为改善这一现状,目前已有大量研究围绕算法性能优化展开探索。部分研究通过拆解优化问题,利用不同种群之间的交互协作加快算法收敛速度^[3]。文献[4]根据优化变量的灵敏度差异完成变量分类,针对不同敏感度等级的变量匹配对应的优化策略,有效提升了电机几何结构的优化效率。除上述优化思路外,诸多研究通过引入代理模型实现优化加速。现代优化算法的核心模块可划分为生成器、评估器与判别器三部分^[5],根据作用机制

的不同,代理模型在优化算法中的应用形式主要分为两大类^[6]。

第一类应用形式是将代理模型作为算法的评估模块,替代仿真计算与数值分析,拟合目标函数的映射关系。由于代理模型的计算速度远优于传统仿真方法,能够显著压缩迭代优化的耗时^[7-9]。但这类代理模型辅助优化的方式仍存在诸多缺陷:(1)模型训练质量稳定性较差,其精度高度依赖训练数据的数量与质量,数据量不足或数据质量偏低,都会大幅降低代理模型的预测精度;(2)代理模型泛化能力薄弱,即便在某一场景下训练完成的高精度模型,仅优化变量新增少量低影响因子也难以适配全新的工况场景;(3)即便代理模型具备较高的整体精度,细微的预测误差仍可能导致优质个体被淘汰,这一问题在算法迭代收敛后期尤为突出。因此,当前研究重点聚焦于第二类应用方式,即利用代理模型优化子代生成策略、改进个体留存机制,以此替代传统优化算法中的生成器与判别器模块^[10]。

现有相关研究中,有学者将优化过程划分为全局搜索与局部搜索两个阶段,在局部搜索阶段将随机响应面模型求解的极小值点作为全新子代个体,提升寻优效率^[10]。也有研究通过多代理模型融合与分层设计策略,强化算法的最优子代寻优能力^[11]。还有部分研究借助人工神经网络挖掘数据特征,生成优质子代个体^[12-14],但人工神经网络可解释性较差,无法阐明其生成优质子代的内在机理。文献[15]采用径向基函数筛选并剔除黑盒优化算法生成的劣质子代,有效减少了黑盒目标函数的调用次数,该方法兼具高效性与良好的可解释性。

代理模型的引入能够显著提升黑盒优化算法的整体性能,但数据稀缺问题始终制约着代理模型优化算法的工程应用^[16]。现阶段多数研究聚焦于迁移优化,以此提升算法优化效率^[17],但现有方法大多适用于多任务场景的跨任务知识迁移或是动态任务迭代过程中的阶段迁移^[18-19],并未探究如何利用已有历史数据提升单一优化任务的运行效率。与之不同,人工智能领域的相关研究十分注重数据的高效利用,迁移学习技术已发展成熟,能够有效降低目标域任务的数据依赖度^[20-21]。即便源域与目标域的数据分布存在较

大差异,迁移学习仍可依托源域数据和学习特征降低目标域的训练难度^[22-23]。因此,将迁移学习与代理优化算法相结合,能够有效缓解代理模型的数据依赖问题,提升单一工程任务的优化效率。

在跨域知识迁移过程中,迁移学习可通过分析不同域数据的统计特征,针对性设计适配的迁移策略^[24-25]。针对源域与目标域差异较大的场景,还可引入中间域数据搭建迁移桥梁,实现知识的高效传递^[26-27]。但迁移学习存在负迁移风险,部分源域无效数据不仅无法辅助目标域任务优化,还会干扰模型训练,降低任务精度^[28-29]。因此,在复用历史数据的过程中,代理模型必须精准筛选有效样本,规避负迁移问题带来的负面影响。

基于上述研究背景,本文设计了一种基于迁移代理模型(Transfer Surrogate Model, TSM)的优化算法。TSM 可自动从已有数据集中筛选有效优质样本,辅助模型完成训练迭代。优化算法通过交叉、变异算子生成大量子代个体后,先由 TSM 筛选出优质候选个体,再开展性能评估,大幅提升整体优化效率。

1 TSM

定义 S 为带标签的源域数据, T 为优化迭代过程中生成的目标域数据。源域与目标域具备同构标签体系,但变量结构存在异构差异。其中,源域数据的标签物理含义与目标优化任务的优化目标完全对应。 S 可以是单一数据源,也可包含多个相互独立的数据源,当存在多组源域数据时,本文将它们整合为一个整体统一参与运算。 N_s 为源域的总样本数量, N_t 为目标域的总样本数量。每组样本均包含一组变量参数 x 以及对应标签 y 。对于目标域而言,变量 x 为待优化参数,标签 y 为对应的优化目标。 S 可取自过往各类优化试验或采样试验的历史数据。TSM 的核心作用是根据 x 精准预测 y 。通过自适应筛选源域中的有效样本并挖掘深层特征信息,能够显著提升 TSM 的预测精度与泛化能力。源域与目标域的数据分布往往存在明显差异,若强行利用源域数据辅助代理模型训练,极易引发负迁移问题。此时,模型从源域习得的知识不仅无法提升其在目标任务中的性

能,反而会导致预测精度下降。为有效规避负迁移风险,本文所提 TSM 引入自编码器结构,完成源域优质样本筛选与特征提取工作,TSM 结构如图 1 所示。该自编码器通过编码模块提取输入数据的深层特征,再利用解码模块实现输入数据重构。除此之外,由于不同数据域的维度不统一,自编码器增设多层线性映射层,实现跨域数据的维度对齐与特征分离。针对维度不同的多组源域数据,每组源域均配置独立的线性映射层,完成与目标域数据的维度匹配。综上,该自编码器的整体工作流程为

$$h = f_e(x)$$

$$\hat{x} = f_d(h) \quad (1)$$

式中: h 为编码器从优化变量中提取的特征向量; $\hat{x}$ 为解码器重构得到的变量结果; $f_e()$ 、 $f_d()$ 分别为编码器、解码器的映射函数。

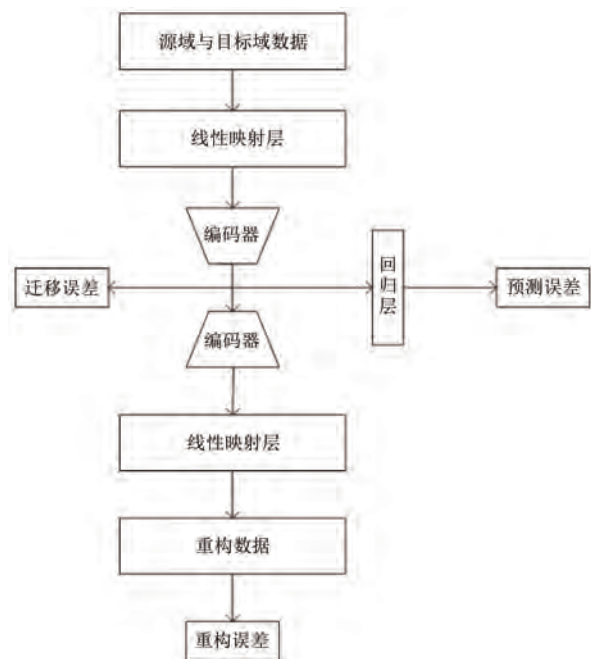


图 1 TSM 结构

Fig. 1 The structure of TSM

为使自编码器能够自主筛选源域中的有效样本,本文采用自编码器的重构误差作为评价指标,量化源域数据与目标域数据的差异程度。现有研究普遍通过设定重构误差阈值^[22],筛选适配目标任务的优质样本,但该方法会导致模型无法利用最终残差完成梯度回传。为此,本文引入 Sigmoid 函数,取消固定阈值设置,保障梯度正常传递,相

关计算式如式(2)~(4)所示:

$$g(a, b) = \|a - b\|_2^2 \quad (2)$$

$$v(a, b) = 2 - \frac{2}{1 + e^{-g(a, b)}} \quad (3)$$

$$J_1 = \frac{1}{N_T} \sum_{i=1}^{N_T} g(\hat{x}_T^i, x_T^i) + \frac{1}{N_S} \sum_{i=1}^{N_S} g(\hat{x}_S^i, x_S^i) \cdot v(\hat{x}_S^i, x_S^i) - \frac{1}{N_S} \sum_{i=1}^{N_S} v(\hat{x}_S^i, x_S^i) \quad (4)$$

式中: $g()$ 为均方误差算子, a, b 为算子的输入参数; $v()$ 为改进型 Sigmoid 函数; J_1 为整体重构损失函数。

权重系数 v 越接近 1,表明样本与目标域数据的相似度越高,自编码器会为该样本分配更高的训练权重;反之, v 越接近 0,表明该样本大概率会引发负迁移,应当予以剔除。 J_1 由三项构成。第一项为目标域数据的重构误差,所有目标域数据均可直接参与训练,无需额外筛选。第二项为源域数据的重构误差,通过 v 实现优质样本的筛选,当源域与目标域数据相似度较高时, v 趋近于 1,第二项误差值随之减小;当源域与目标域数据差异较大时, v 趋近于 0,第二项误差值趋近于极小值。因此在 TSM 的训练过程中,为使整体误差最小化,容易出现将源域权重全部置零,完全舍弃源域数据的训练倾向。为避免该问题、充分挖掘源域数据的迁移价值,本文在损失函数中增设第三项约束项,引导自编码器充分利用源域数据。该约束项的残差会随源域有效筛选样本数量的增加而减小,本质是依托源域数据对自编码器实施辅助优化。与此同时, Sigmoid 函数具备梯度裁剪效应,能够有效防止自编码器对源域数据出现过拟合问题。

TSM 能够依托源域的变量知识辅助开展目标优化任务。但与图像、文本数据不同,变量本身承载的有效信息有限。为实现效果更优的迁移学习,还需充分利用源域的标签信息。源域与目标域的标签结构虽然同构,但标签数据的分布存在差异,不同域的标签数值往往相差较大,增加了迁移学习的训练难度,导致模型整体性能下降。采用普通回归层学习源域数据,会造成源域数据过拟合,致使模型在目标域的表现变差。为在充分利用标签数据的同时,保证回归层的预测性能, TSM 引入了余弦相似度损失函数。该损失函数通

过捕捉特征向量之间的方向关联关系,使模型学习源域标签数据中蕴含的有效知识,损失函数表达式为

$$C_{ij} = \begin{cases} 1 - \cos(\mathbf{h}_i, \mathbf{h}_j), & |y_i - y_j| \leq \lambda_a \\ \max[0, \cos(\mathbf{h}_i, \mathbf{h}_j) - \lambda_m], & \text{其他} \end{cases} \quad (5)$$

式中: $\mathbf{h}_i, \mathbf{h}_j$ 分别为编码器从输入数据 x_i, x_j 中提取的特征向量,输入数据可取自源域或目标域; y_i, y_j 分别为 x_i, x_j 对应的标签; $\cos(\mathbf{h}_i, \mathbf{h}_j)$ 算子用于计算两组特征向量之间的余弦相似度; λ_a 为超参数,用于判定两组特征向量是否具有相似特征; λ_m 为余弦相似度阈值。

λ_a 能够根据标签的差异,指导损失函数选择不同的计算方式。针对多维标签,可通过计算标签各元素的平均差值完成差异对比。当标签足够相似时,损失函数会指导模型拉近样本特征间的差距;当标签不够相似时,损失函数会指导模型区分不同样本。为了防止模型过度惩罚标签不相似的样本对,损失函数引入阈值 λ_m ,只有特征相似度超过 λ_m 时,才需要惩罚;如果特征已经足够不相似,就不再惩罚。

在计算出所有样本两两特征向量之间的余弦相似度损失后,可求得平均损失 J_2 :

$$J_2 = \frac{1}{(N_T + N_S)^2} \sum_{i=1}^{N_T+N_S} \sum_{j=1}^{N_T+N_S} C_{ij} \quad (6)$$

式(6)为代理模型学习标签数据提供了全面的评价依据,能够充分捕捉标签中蕴含的内在知识。通过最小化平均损失 J_2 ,可使 TSM 更好地挖掘源域的相关知识并将其迁移至目标域,从而提升模型的整体性能。

自编码器完成迁移学习任务后,还需构建专属预测模型以实现目标域的标签预测。采用全连接神经网络(Fully Connected Neural Network, FCNN)搭建(与自编码器所用网络结构一致)预测模型,将训练学习得到的特征向量映射至目标空间。FCNN 是核心功能模块,负责将自编码器提取的抽象特征表征转换映射到目标域标签空间。TSM 以目标域的预测误差作为优化指标,对回归层参数进行迭代优化。该优化方式能够确保模型的参数调整完全依据其在目标域中的实际性能表现,如式(7)所示:

$$\hat{y}_T^i = f_i(\mathbf{h}_T^i)$$

$$J_3 = \frac{1}{N_T} \sum_{i=1}^{N_T} g(\hat{y}_T^i, y_T^i) \quad (7)$$

式中: $\mathbf{h}_T^i$ 为自编码器从输入数据中提取的特征向量; f_i 为回归层; $\hat{y}_T^i$ 为回归层输出的预测目标标签; J_3 为仅聚焦于目标域的损失函数, 可有效规避源域数据带来的干扰。

模型预测值与真实标签之间的误差依旧采用均方误差进行量化。结合回归层与自编码器的损失约束, TSM 的整体损失为

$$\min_{\theta} J = J_1 + J_2 + J_3 \quad (8)$$

式中: θ 为 TSM 的所有参数, 涵盖自编码器、线性映射层以及回归层的全部参数。

该整体损失同时融合源域与目标域的约束信息, 一方面利用自编码器的重构误差与余弦相似度从源域中筛选有效样本; 另一方面依托回归层完成目标域的预测任务。因此, TSM 能够充分学习源域中的有效知识, 从而更好地完成目标域任务。

由于目标域数据样本数量有限, TSM 需要高效利用这类稀缺数据资源, 同时防止对局部目标域数据产生过拟合。因此训练过程需要采取均衡策略, 在充分挖掘现有数据价值的同时, 保障模型在整个目标域的泛化能力。为解决此问题, 本文在 TSM 的训练流程中引入交叉验证与早停机制。

模型训练流程图如图 2 所示。图中, N_{epoch} 为模型最大训练轮次; N_{stop} 为早停判定轮次; $\min N$ 为综合验证损失 J_{val} 连续未改善的训练轮次; N_L 为 TSM 中 FCNN 的层数; J_{min} 为最小综合验证损失。TSM 训练时, 将目标域数据划分为 K 个子集, 训练过程中轮流选取其中一个子集作为验证集, 用于评估模型性能, 剩余子集则与源域数据共同组成训练集, 完成代理模型的迭代训练。第 i 个子集作为验证集时得到的预测误差记为 J_i , 其计算方式与训练过程中的损失项 J_3 保持一致。所有子集对应的预测误差共同构成本轮训练的综合验证损失 J_{val} , 并以此作为模型性能的评价依据。

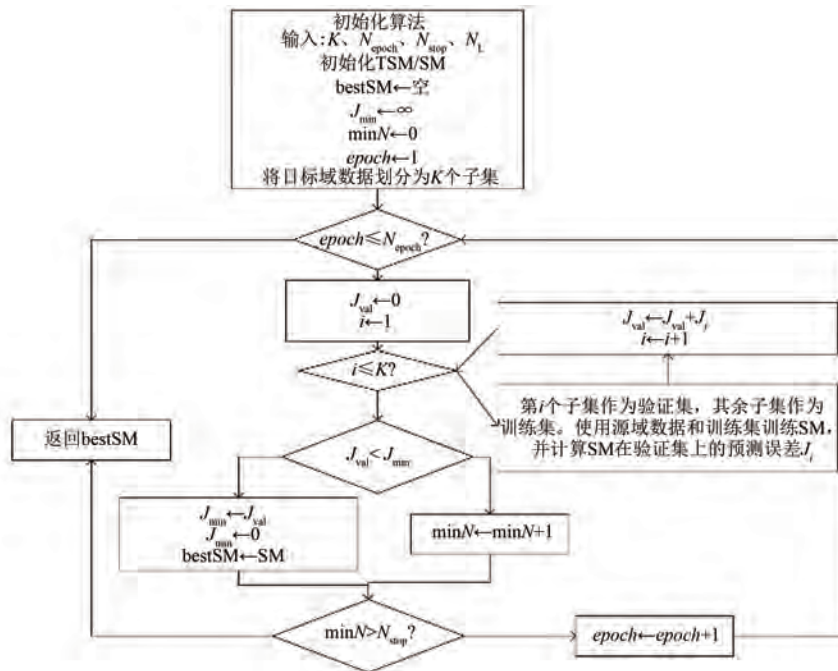


图 2 模型训练流程图

Fig. 2 Flowchart of the model training

受目标域数据总量限制, 若 K 值过大, 各子集的样本数量会过少, 导致代理模型无法充分学习数据特征, 引发训练不稳定以及局部子集过拟合问题; 反之, 若 K 值过小, 会降低 TSM 的验证精度, 同样会造成模型过拟合。由此可见, 选取合适

的 K 值对模型性能至关重要。 K 值可根据计算机硬件可承载的最大数据批处理量以及优化过程生成的目标域数据总量综合确定。除交叉验证外, 还引入早停机制, 以节约算力资源、加快训练速度。当代理模型在连续多个训练轮次内, 目标域性能不

再提升时,训练过程将提前终止。训练停止后,算法将回存并输出训练全程性能最优的代理模型。

2 基于 TSM 的优化算法

优化算法需要完成种群的生成、评估与保留工作。若优化算法引入 TSM,就需要明确模型的调用时机与使用方式。本文所提优化算法利用 TSM 生成性能更优的全新子代个体,以此加快优化过程的收敛速度。

基于 TSM 的优化算法框架如图 3 所示。该优化算法整体流程基于非支配排序多目标遗传算法 (Non-dominated Sorting Genetic Algorithm II, NSGA-II) 框架进行搭建。首先,在优化迭代开始前,需准备源域数据用于辅助训练 TSM,且该源域数据的标签物理含义需与目标任务的优化目标保持一致。优化启动后,采用拉丁超立方抽样生成初始种群,并通过评估函数计算该初始种群的目标函数值。完成初始种群评估后,计算初代优化目标的统计特征参数,依托该参数对源域数据以及优化过程中生成的所有数据做标准化处理。迭代优化过程中,算法通过拥挤距离计算、非支配排序以及精英保留策略,筛选出性能优异且多样性充足的子代个体。当优化迭代代数小于阈值 N_2 时,算法采用模拟二进制交叉算子与多项式变异算子,生成包含 K_1 个个体的子代种群;当优化迭代代数大于 N_2 时,算法将 TSM 与交叉、变异算子相结合,生成全新子代个体。阈值 N_2 用于控制 TSM 介入辅助优化的时机,可有效规避因训练数据不足导致的 TSM 性能不佳、过拟合等问题。用

于辅助训练代理模型的源域数据与目标任务的相似度越高, N_2 的设定值越小。

优化算法在生成子代个体时,更关注父代种群的排序等级,而非目标函数的具体数值。若直接采用代理模型替代评估函数,即便模型仅产生微小预测误差,也会造成种群排序结果出错,进而生成性能更差的子代个体。此外,优化目标的数据分布会随迭代过程不断偏移,这就要求代理模型在优化过程中持续迭代训练。若将代理模型直接作为评估函数使用,模型的训练数据量会持续缩减,无法及时适配数据分布的动态变化。因此,代理模型不能直接替代优化算法中的评估模块。

为解决上述问题,充分发挥代理模型的辅助优化作用,本文所提优化算法利用代理模型筛选子代中的劣解个体,使更多算力资源集中于潜力更高的优质个体。具体流程如下:首先,通过交叉、变异算子生成 K_2 个子代个体,且 K_2 的取值需远大于 K_1 ;随后,依托 TSM 预测所有子代个体的优化目标,并根据预测结果对子代种群开展非支配排序;最后,从所有子代个体中筛选出 K_1 个最优个体,作为最终的有效子代种群。该代理模型可有效规避模型预测误差带来的负面影响。得益于精英保留策略,即便 TSM 存在一定预测误差,导致非支配排序结果出现小幅偏差,也基本不会影响优质精英个体的保留效果。

优化算法每一代的评估函数的调用次数不会减少,用于训练代理模型的数据集可不断纳入最新优化数据,保障模型能够根据数据分布偏移及时更新参数。除此之外,简化 TSM 的优化流程可进一步提升整体优化算法的性能。仅当优化过程出现全新最优解时,才对 TSM 进行参数更新;同时,在 TSM 训练初始化阶段,可复用历史模型参数替代随机初始化参数,加快模型收敛速度。上述方法均可有效降低优化算法的算力消耗,提升整体优化效率。

优化过程中,TSM 可自适应调整自编码器与回归层的结构参数。TSM 将线性映射层的输出维度设置为所有数据域中的最大维度,特征维度则匹配所有数据域中的最小维度。编码器、解码器及回归层中 FCNN 的神经元数量统一设置为目标域维度的两倍,且三类模块的 FCNN 层数保持一致,使整个 TSM 仅保留单一结构参数,大幅简化模型调优流程。值得说明的是,尽管本文最终的

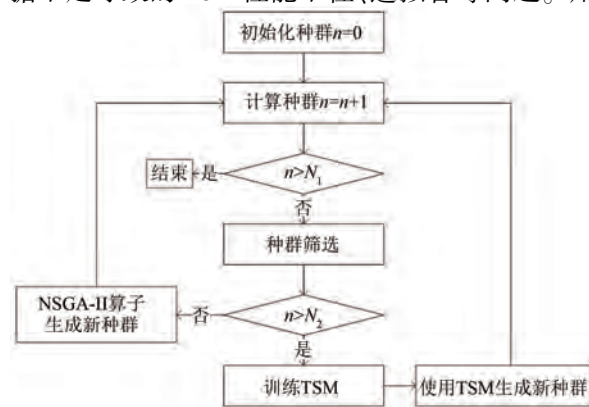


图 3 基于 TSM 的优化算法框架

Fig. 3 The framework of optimization algorithmml based on TSM

优化目标可简化为最大化输出转矩,但由于存在多类强非线性隐式约束(如电机几何干涉问题),采用普通单目标遗传算法配合固定罚函数极易导致可行域边界搜索失效。因此,本文选用具备约束处理能力与种群多样性保持能力的 NSGA-II 框架作为基础算法,以确保在复杂可行域边界上稳定逼近全局最优可行解。

3 试验验证

以电机设计为应用场景,将本文所提基于 TSM 的优化算法与传统 NSGA-II 进行对比,以验证所提算法的实际性能。电机设计的核心流程是根据工况需求匹配电机结构参数,并依托优化算法完成参数迭代优化。在实际设计过程中,单次优化迭代通常无法得到理想设计结果,往往需要根据优化结果

调整电机结构,反复迭代优化,最终确定最优设计方案。传统优化算法无法复用历次优化的无效样本数据,而电机设计场景存在大量可挖掘的历史迭代数据,因此非常适合作为基准测试案例,验证本文基于 TSM 的优化算法的性能。

本文选取四种不同结构的电机作为优化对象,如图 4 所示。四款电机的定子结构完全一致,转子结构存在差异。由于转子永磁体结构对电机性能的影响远大于定子槽型结构,四款电机的整体性能呈现明显差异。将电机的核心结构参数设为优化变量,以电机输出转矩作为优化目标,最大化指定型号电机的输出转矩。本文采用拉丁超立方抽样方法,为每款电机采集 8 000 组以上样本数据。针对任意一款待优化电机,其余三款电机的数据集均作为源域数据,用于辅助 TSM 的训练。

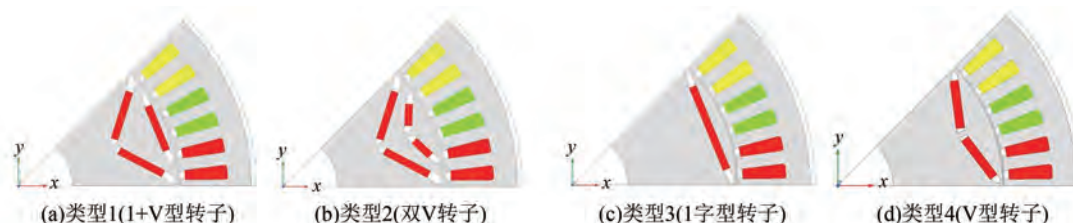


图 4 优化对象

Fig. 4 Optimization objects

试验中,除代理模型调用相关模块存在差异外,所有优化算法的其余参数完全统一。具体试验参数设置如下:算法总迭代次数为 100 代,每代生成 100 个种群个体,经精英保留策略筛选后,每代保留 30 个优质个体,算法的交叉算子与变异算子保持一致。所有电机性能仿真均采用 ANSYS 公司的 Maxwell 电磁仿真软件完成。优化迭代至第 5 代后,算法开始启用 TSM 辅助优化。在 TSM 训练参数中,相关参数分别设置为 0.01 和 0.5。启用 TSM 后,算法每代生成 1 000 个初始个体,通过 TSM 完成性能预测,筛选出最优 100 个个体作为正式子代种群。模型训练硬件平台为 RTX A2000 显卡。TSM 的线性映射层为单层线性结构,编码器、解码器及回归层均采用五层带残差连接的 FCNN,即 $N_L=5$ 。模型训练过程中,目标域数据被划分为 4 个子集用于交叉验证,即 $K=4$,每轮训练轮流选取一个子集作为验证集,其余子集作为训练集,循环遍历完成一轮完整训练。四轮交叉验证全部完成后,计算所有子集的综合验证损失 J_{val} 作为本轮

训练的最终误差。模型最大训练轮次 N_{epoch} 设置为 20 轮,早停判定轮次 N_{stop} 设置为 5,即若综合验证损失 J_{val} 连续 5 轮无下降,则提前终止训练。

图 5 为本文所提基于 TSM 的优化算法与传统 NSGA-II 的性能对比结果。两种算法均完成 100 代迭代优化,每间隔 5 代记录一次电机最优输出转矩数据。

由图 5 可知,由于 TSM 从第 5 代才开始介入优化,因此前 5 代两种算法的迭代过程完全一致。从第 5 代开始,本文优化算法依托 TSM 的预测结果提供寻优指导,降低了对随机交叉、变异算子的依赖,实现了更高效、更具针对性的搜索空间探索。四款电机的优化试验结果均表明,引入 TSM 能够显著加快优化收敛速度。同时,试验结果证实,优化问题的复杂度越高,TSM 带来的性能提升效果越明显。传统优化算法在复杂搜索空间中寻优难度大、效率低,而人工神经网络具备优异的自适应学习能力,可从现有数据中深度挖掘有效特征与先验知识,使本文算法在复杂优化场景中获得更大的性能增益。

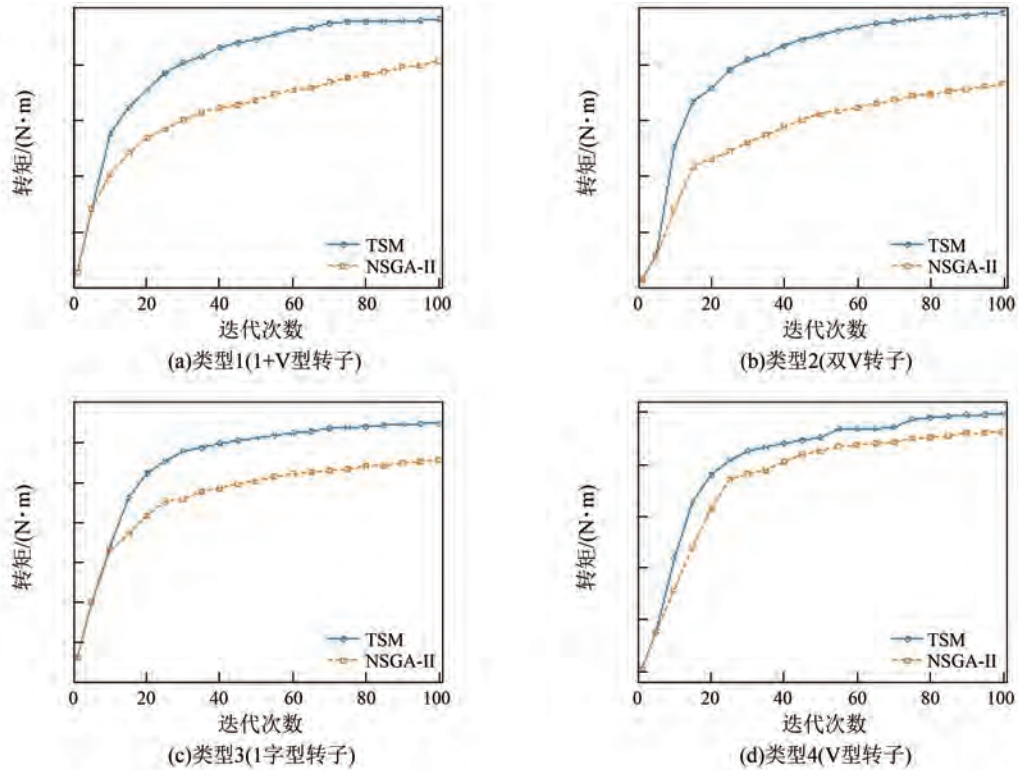


图5 两种优化算法的性能对比结果

Fig. 5 Performance comparison results of two optimization algorithms

TSM 通过引入源域已有数据,辅助构建适配目标任务的代理预测模型。图 6 为四种不同代理模型的性能对比结果,对比模型包括 Kriging 模型、支持向量机(Support Vector Machine, SVM)以及未引入源域数据的 FCNN。性能对比基于类型 1 电机的优化数据开展,所有代理模型均采用前 $i-1$ 代优化数据作为训练集,预测第 i 代的目标函数值。

由图 6 可知,相较于其他代理模型,TSM 能够更快收敛至稳定、高精度的预测状态;在训练数据

充足的条件下,其预测精度与 Kriging 模型基本持平。更为关键的是,优化过程中存在的数据分布偏移问题会导致其他代理模型性能大幅下降,而 TSM 依托源域数据的知识迁移,有效提升了模型鲁棒性,削弱了数据偏移带来的负面影响。在实际工程优化问题中,TSM 的高鲁棒性与快速收敛特性,能够大幅提升优化算法的搜索寻优能力。

4 结语

本文提出一种基于 TSM 的优化算法,通过从源域中筛选有效样本,辅助搭建目标域代理模型,提高优化效率,能够在有限的计算预算内显著提升电机优化效率。相较于传统优化算法,基于 TSM 的优化算法的收敛速度更快。通过将 TSM 与其他代理模型对比,证实 TSM 不仅具备更优的鲁棒性,还能在优化过程中更早获得更低的预测误差。尽管受限于计算资源与迭代时长,本文未提供详尽的消融试验,但从算法设计的底层逻辑上,三个模块具有明确的分工与不可替代性。自编码器重构筛选的目标是解决输入空间(变量维度)的异构分布差异,通过可学习的重构权重剔除源域

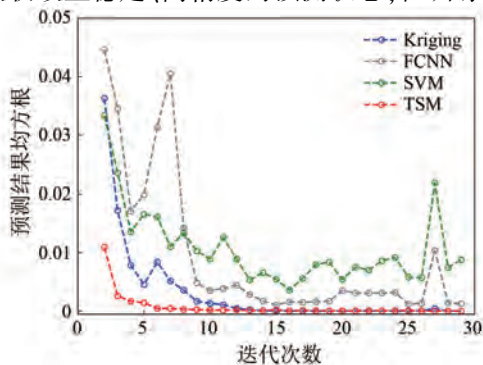


图6 不同代理模型的性能对比结果

Fig. 6 Performance comparison results of different surrogate models

中与目标域物理结构无关的干扰样本。而余弦相似度损失的目标是解决输出空间(标签性能)的非凸映射差异,确保特征模型能够正确筛选样本。这些模块的协同是 TSM 在少样本条件下仍能优于 Kriging 及普通神经网络的根本原因。

利益冲突声明

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

作者贡献

王志强进行了方案设计、试验验证、内容总结与论文撰写。乔振阳进行了方案设计、试验验证。张云鹏、傅为农进行了论文的审核与修改。所有作者均阅读并同意了最终稿件的提交。

The scheme design, experimental verification, content summary and paper writing were carried out by Wang Zhiqiang. The scheme design and experimental verification were conducted by Qiao Zhenyang. The manuscript was reviewed and revised by Zhang Yunpeng, Fu Weinong. All authors have read the last version of the manuscript and consented for submission.

参考文献

- [1] 邵克勇, 陈威杰, 郝晓伟. 改进金枪鱼群优化算法的永磁同步电机多参数辨识[J]. 电机与控制应用, 2025, 52(10): 1097-1107.
Shao K Y, Chen W J, Hao X W. Multi parameter identification of permanent magnet synchronous motor with improved tuna swarm optimization algorithm [J]. Electric Machines & Control Application, 2025, 52(10): 1097-1107.
- [2] 张宏, 周大伟, 陆丽, 等. 基于改进的樽海鞘群算法的永磁同步电机多目标优化设计[J]. 电机与控制应用, 2025, 52(2): 221-230.
Zhang H, Zhou D W, Lu L, et al. Multi-objective optimization design of permanent magnet synchronous motor based on improved salp swarm algorithm [J]. Electric Machines & Control Application, 2025, 52(2): 221-230.
- [3] Omidvar M N, Li X D, Yao X. A review of

- population-based metaheuristics for large-scale black-box global optimization—Part I [J]. IEEE Transactions on Evolutionary Computation, 2022, 26(5): 802-822.
- [4] Zhu X Y, Shu Z M, Quan L, et al. Multi-objective optimization of an outer-rotor V-shaped permanent magnet flux switching motor based on multi-level design method [J]. IEEE Transactions on Magnetism, 2016, 52(10): 8205508.
- [5] Liu S B, Lin Q Z, Li J Q, et al. A survey on learnable evolutionary algorithms for scalable multiobjective optimization [J]. IEEE Transactions on Evolutionary Computation, 2023, 27(6): 1941-1961.
- [6] Jin Y C, Wang H D, Chugh T, et al. Data-driven evolutionary optimization: An overview and case studies [J]. IEEE Transactions on Evolutionary Computation, 2019, 23(3): 442-458.
- [7] Ye Q C, Cai H X, Bian Y Q. Ensemble-based offline data-driven evolutionary optimization for operation parameter of blast furnace [C]//International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, Guangzhou, China, 2024: 1-7.
- [8] Ling X N, Jia X N. An effective kriging-assisted water flow optimization for ultra-wideband monopole antenna design [C]//2024 IEEE 7th International Conference on Electronic Information and Communication Technology, Xi'an, China, 2024: 540-544.
- [9] Guo D, Wang X L, Gao K L, 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: Systems, 2022, 52(4): 2084-2097.
- [10] Bai F S, Zhou Z. A modified stochastic response surface algorithm for expensive black-box global optimization [C]//2017 22nd International Conference on Digital Signal Processing, London, United Kingdom, 2017:1-5.
- [11] Ye P C, Pan G, Dong Z M. Ensemble of surrogate based global optimization methods using hierarchical design space reduction [J]. Structural and Multidisciplinary Optimization, 2018, 58(2): 537-554.
- [12] Cheng R, Jin Y C, Narukawa K, et al. A

- multiobjective evolutionary algorithm using gaussian process-based inverse modeling [J]. IEEE Transactions on Evolutionary Computation, 2015, 19 (6): 838-856.
- [13] He C, Huang S H, Cheng R, et al. Evolutionary multiobjective optimization driven by generative adversarial networks (GANs) [J]. IEEE Transactions on Cybernetics, 2021, 51(6): 3129-3142.
- [14] Wang Z Z, Hong H K, Ye K, et al. Manifold interpolation for large-scale multiobjective optimization via generative adversarial networks [J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(8): 4631-4645.
- [15] Regis R G. Evolutionary programming for high-dimensional constrained expensive black-box optimization using radial basis functions [J]. IEEE Transactions on Evolutionary Computation, 2014, 18 (3): 326-347.
- [16] Li K, Chen R Z, Yao X. A data-driven evolutionary transfer optimization for expensive problems in dynamic environments [J]. IEEE Transactions on Evolutionary Computation, 2024, 28 (5): 1396-1411.
- [17] Tan K C, Feng L, Jiang M. Evolutionary transfer optimization—A new frontier in evolutionary computation research [J]. IEEE Computational Intelligence Magazine, 2021, 16(1): 22-33.
- [18] Wu S H, Li J Y, Zhan Z H, et al. Multi-view transfer with marginal and conditional alignment for many-task optimization [C]//2023 IEEE Congress on Evolutionary Computation, Chicago, USA, 2023: 1-8.
- [19] Chen H, Liu H L, Gu F, et al. A multiobjective multitask optimization algorithm using transfer rank [J]. IEEE Computational Intelligence Magazine, 2023, 27(2): 237-250.
- [20] Liang H, Fu W D, Yi F J. A survey of recent advances in transfer learning [C]//2019 IEEE 19th International Conference on Communication Technology, Xi'an, China, 2019: 1516-1523.
- [21] Zhang Y, Yang Q. A survey on multi-task learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(12): 5586-5609.
- [22] Pan S J, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis [J]. IEEE Transactions on Neural Networks, 2011, 22 (2): 199-210.
- [23] Xiao Q, Zhang Y, Yang Q. Selective random walk for transfer learning in heterogeneous label spaces [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(6): 4476-4488.
- [24] Pan S J, Tsang I W, Kwok J T, et al. Domain adaptation via transfer component analysis [J]. IEEE Transactions on Neural Networks, 2011, 22 (2): 199-210.
- [25] Deng Z, Choi K S, Jiang Y, et al. Generalized hidden-mapping ridge regression, knowledge-leveraged inductive transfer learning for neural networks, fuzzy systems and kernel methods [J]. IEEE Transactions on Cybernetics, 2014, 44 (12): 2585-2599.
- [26] Xiao Q, Zhang Y, Yang Q. Selective random walk for transfer learning in heterogeneous label spaces [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(6): 4476-4488.
- [27] Tan B, Zhang Y, Pan S, et al. Distant domain transfer learning [C]//31st AAAI Conference on Artificial Intelligence, California, USA, 2017: 2604-2610.
- [28] Minoofam S A H, Bastanfard A, Keyvanpour M R. TRCLA: A transfer learning approach to reduce negative transfer for cellular learning automata [J]. IEEE Transactions on Neural Networks and Learning Systems, 2023, 34(5): 2480-2489.
- [29] 刘志坚, 黄建, 骆军. 基于 STFT 图像和迁移学习的次同步振荡源定位方法[J]. 电机与控制应用, 2024, 51(7): 119-131.
- Liu Z J, Huang J, Luo J. Localization method for sub-synchronous oscillation sources based on STFT images and transfer learning [J]. Electric Machines & Control Application, 2024, 51(7): 119-131.

收稿日期:2026-05-23

收到修改稿日期:2026-06-24

作者简介:

王志强(2001—),男,硕士研究生,研究方向为人工智能在电机设计与控制中的应用,zq.wang1@siat.ac.cn;

*通信作者:张云鹏(1990—),男,博士,副研究员,研究方向为电机优化设计与电磁场数值分析,yp.zhang2@siat.ac.cn。