

徐智涵, 朱延江, 赵汝鑫. 基于混合遗传算法的军事图像分类模型超参数优化研究[J]. 智能计算机与应用, 2026, 16(2): 162-170. DOI:10.20169/j.issn.2095-2163.24042202

基于混合遗传算法的军事图像分类模型超参数优化研究

徐智涵, 朱延江, 赵汝鑫

(扬州大学 信息工程学院, 江苏 扬州 225127)

摘要: 针对面向军事图像分类的卷积神经网络模型训练超参数设置问题, 提出了一种混合遗传算法, 对自建的军事图像数据集进行实验分析。算法通过设计适应度函数增强个体辨识度, 并相继引入精英策略、自适应思想和多种群并行的思想, 对算法的选择、交叉、变异与整体迭代过程进行改进, 提升收敛速度、收敛精度与全局收敛能力。针对传统遗传算法存在个体趋同现象, 提出一种多样性保持策略, 降低陷入局部收敛的风险。实验结果表明, 在1000次迭代中, 相较于7种经典的智能优化算法, 所提算法的识别错误率平均降低1.41%, 全局收敛概率平均提升107.59%, 验证了所提算法解决面向军事图像分类的卷积神经网络模型训练超参数设置问题的有效性。

关键词: 混合遗传算法; 超参数优化; 自适应思想; 多种群并行思想; 多样性保持策略; 军事图像分类

中图分类号: TP391.4

文献标志码: A

文章编号: 2095-2163(2026)02-0162-09

Research on hyperparameter optimization of military image classification model based on hybrid genetic algorithm

XU Zhihan, ZHU Yanjiang, ZHAO Ruxin

(College of Information Engineering, Yangzhou University, Yangzhou 225127, Jiangsu, China)

Abstract: This paper proposes a hybrid genetic algorithm for the hyperparameter setting problem in convolutional neural network models aimed at military image classification, and conducts experimental analysis on a self-constructed military image dataset. The algorithm improves individual distinguishability through a designed fitness function and incorporates elite strategy, adaptive concept, and multi-population parallelism to enhance the selection, crossover, mutation, and overall iterative processes, which increases convergence speed, accuracy, and global convergence capability. A diversity preservation strategy is proposed to address the phenomenon of individuals converging to the same individual in traditional genetic algorithms, in order to reduce the risk of falling into local convergence. The experimental results have shown that over 1000 iterations, compared to 7 classic intelligent optimization algorithms, the proposed algorithm reduces the recognition error rate by an average of 1.41% and improves the global convergence probability by an average of 107.59%, verifying the effectiveness of the proposed algorithm in solving the hyperparameter setting problem of convolutional neural network model training for military image classification.

Key words: hybrid genetic algorithm; hyperparameter optimization; adaptive concept; multi-population parallelism; diversity maintenance strategy; military image classification

0 引言

在现代军事应用中, 基于卷积神经网络(Convolutional Neural Network, CNN)的图像识别技术扮演着至关重要的角色^[1-3]。CNN的性能在很大程度上依赖于其超参数设置, 包括学习率的设置、优

化器的选择、训练数据的批量大小等, 这些超参数的设置往往需要专业知识和繁琐的试验过程^[4-5]。针对超参数设置问题, 目前的主流方法之一是遗传算法。

作为智能优化算法中的代表, 遗传算法(Genetic Algorithm, GA)^[6]通过模拟生物界自然选择求解问题, 其实现门槛低, 不依赖于问题的具体特

基金项目: 扬州大学大学生创新创业训练计划(XCX20230402)。

作者简介: 徐智涵(2003—), 男, 本科生, 主要研究方向: 智能优化及应用; 朱延江(2002—), 男, 本科生, 主要研究方向: 聚类算法, 智能优化及应用。

通信作者: 赵汝鑫(1990—), 男, 博士, 讲师, 主要研究方向: 智能优化及应用。Email: nnzhaoms@126.com。

收稿日期: 2024-04-22

性,并且具有健壮性、普适性和强收敛性^[7],因此广泛应用于函数优化问题、组合优化问题、调度问题等。

针对传统遗传算法的改进与应用一直是研究的热点^[8],特别是在超参数优化问题上的应用。余维等学者^[9]提出一种改进实数编码遗传算法,采用实数编码扩大超参数搜索空间,引入分层比例选择算子增加解集多样性,设计改进的单点交叉和变异算子以更全面地探索超参数空间,有效提高了算法的效率和求解质量。邓翔等学者^[10]提出了一种利用量子遗传算法优化超参数的新型 Prophet 模型,大幅简化超参数优化,能够更好地求解时间序列预测问题。杜沅昊等学者^[11]提出了一种混合遗传蚁群算法,利用遗传算法优化蚁群算法的初始信息素分布,将蚁群算法寻优结果作为反向传播神经网络(Back Propagation Neural Network, BPNN)的最优权值和阈值,提升 BPNN 预测空气质量指数的精度。Nikbakht 等学者^[12]利用遗传算法结合超参数组合的禁忌列表,有效优化了深度能量方法(Deep Energy Method, DEM)的超参数,证明了超参数优化对 DEM 精度的影响。

综上,对传统遗传算法进行改进,本文提出了一种混合遗传算法(Hybrid Genetic Algorithm, HGA)。首先,设计适应度函数,增强个体辨识度,加速种群进化;其次,引入精英策略^[13]、自适应思想^[14]和多种群并行的思想^[15],提升算法收敛速度、求解精度和全局寻优能力;最后,提出一种多样性保持策略,使得算法在任意状态下能够搜索问题候选解空间,降低陷入局部收敛的风险。实验结果表明,相较于传统遗传算法和其他经典智能优化算法,改进后的算法具有更好的收敛性能。

1 遗传算法及其问题分析

遗传算法是一种模拟生物界自然选择和遗传机制的智能优化算法,该算法受以自然选择理论^[16]为代表的生物进化论启发,将问题的解空间视为自然界,算法的解集视为生物种群,种群中的每一个个体是问题的一个解。算法初始依某种规则产生一个种群,通过模拟生物进化过程中的选择、交叉和变异等基本操作,产生一群更适合环境的个体。通过不断地繁衍进化,从而得到问题的优质解。遗传算法的基本流程如图 1 所示。

在传统遗传算法中,存在如下问题:

(1) 对生物进化过程的模拟依赖于适应度函

数,其设计不当易导致算法性能下降。

(2) 选择操作使用适应度比例方法,其无法保障优质个体的必然传承与劣质个体的高效淘汰,从而致使优质解丢失与无效搜索,抑制了算法的收敛速度和收敛精度。

(3) 交叉操作和变异操作使用固定的交叉率和变异率设置,其个体间基因交流缺乏导向性,新鲜基因的加入缺乏针对性,不利于保留优质个体和打破局部收敛。

(4) 随着算法迭代次数的增加,趋同现象愈发明显,个体进化方向受限,易使种群进化陷入停滞。

(5) 对单一种群进化过程进行模拟,易出现早熟收敛现象,在多峰问题求解中尤为明显。

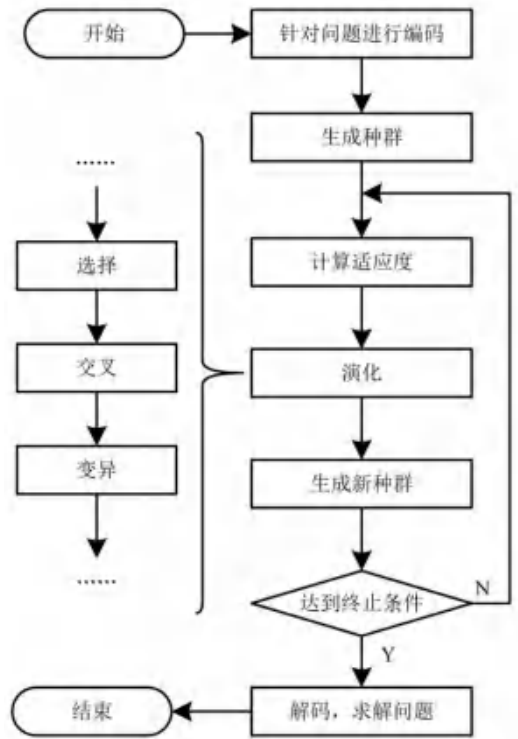


图 1 遗传算法的流程图

Fig. 1 Flowchart of genetic algorithm

2 混合遗传算法

针对传统遗传算法存在的上述问题,本文提出了一种混合遗传算法,算法基本流程如图 2 所示。

算法初始随机生成多个种群,各种群并行地执行迭代操作,以全部种群完成一次迭代作为算法完成一次迭代的标志,种群间依规则交流。在具体的迭代过程中,基于算法对适应度函数的依赖性,设计适当的目标函数映射关系,以更好地对个体进行优劣性评估;在选择操作中,引入精英策略,解决传统

轮盘赌算法中优质个体小概率丢失问题,并且引入自适应思想提升交叉和变异对种群进化的导向性作用;针对趋同现象,引入一种多样性保持策略,对种群依据一定规则进行干预,保障算法在任意状态下能够搜索问题候选解空间,提升算法的全局搜索能力。

(Fitness Proportionate Selection)^[6]的方法进行选择操作。适应度比例选择方法会单独计算并累计每个个体的适应度,以此计算每个个体的算术平均适应度,并将算术平均适应度作为个体在选择操作中被选中保留的概率,表示为:

$$p_i = \frac{F_i}{\sum F} \quad (2)$$

其中, p_i 表示第 i 个个体被选中保留的概率。

尽管适应度高的个体更容易被选中,但仍以小概率出现最差个体被选中的情况,从局部搜索能力的角度出发,位于最差个体邻域的解质量较差,容易导致无效搜索,最差个体参与迭代对种群进化贡献较小。此外,已有算法的适应度比例选择方法不能正确处理个体优质程度与保留概率的关系。对此,本文对模型错误率进行映射产生算法的适应度,适应度值越低,则个体越优质,表示为:

$$p_i = \frac{F_{\max} - F_i + \text{bias}}{\sum (F_{\max} - F + \text{bias})} \quad (3)$$

其中, F_{\max} 表示当前种群中所有个体的适应度的最大值。

经过改进后,算法能够正确处理个体优质程度与保留概率的映射关系,已有算法中的最差个体,经过适应度比例选择后以极大概率被淘汰;同时适应度最高的个体被保留。避免算法进行无效搜索,同时优质个体被保留的概率增加。

不过,改进的适应度比例的方法仍无法保障优质解的必然传承,致使优质解与优质基因的小概率丢失,或致使算法难以收敛至全局最优。针对这一问题,本文采用精英策略^[13]进一步改进选择操作。

具体来说,在种群初始化和每次迭代的结束时,算法将当前种群中最优的个体保留为历史最优个体;在每次迭代的选择操作中,针对通过适应度比例选择得到的种群中的最差个体用算法的历史最优个体进行替代,如图3所示。

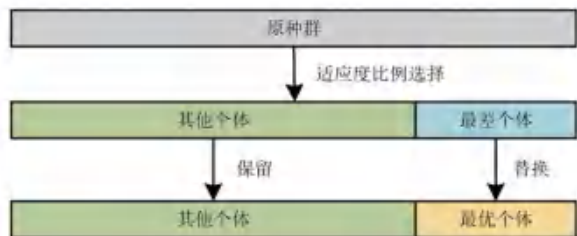


图3 选择操作的过程

Fig. 3 The process of selecting operation

通过适应度比例选择进行精英种群替换,避免

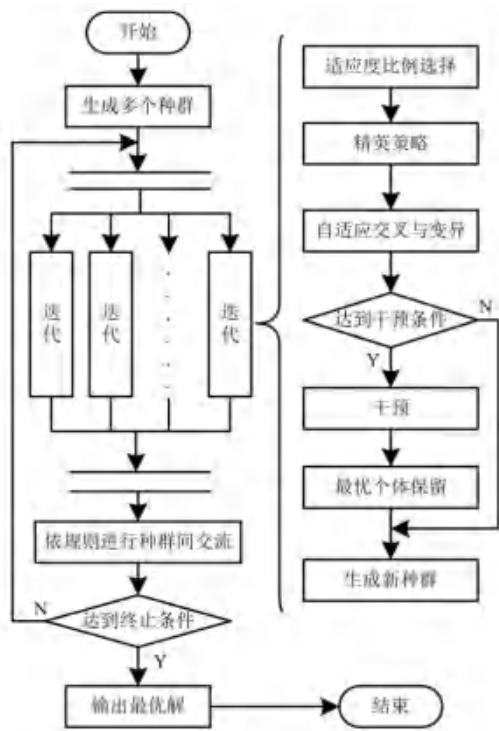


图2 混合遗传算法的基本流程

Fig. 2 The basic process of hybrid genetic algorithm

2.1 适应度函数设计

适应度函数是遗传算法中个体优劣判别的主要依据,对算法的运行发挥着决定性作用。在某些特定问题中,对于问题解空间对应的目标函数值空间,各目标函数值之间可能彼此差异较小而缺乏辨识度,或差异较大而致使数据失真,易使算法出现收敛速度慢、早熟收敛等问题。因此,针对这一问题,本文结合对数变换对目标函数值进行如下映射以得到合适的适应度,表示为:

$$F_i = \log_2 \left(\left| \frac{O_i - O_{\min} + \text{bias}}{O_{\max} - O_{\min} + \text{bias}} \right| + 1 \right) \quad (1)$$

其中, F_i 表示当前种群中第 i 个个体的适应度; O_i 表示当前种群中第 i 个个体的目标函数值; O_{\max} 表示当前种群中所有个体的目标函数值的最大值; O_{\min} 表示当前种群中所有个体的目标函数值的最小值。在本文中, bias 是一个数值上相对极小的偏置项,避免运算过程中除零错误的产生。

2.2 精英策略改进适应度函数

传统遗传算法通常采用适应度比例选择

了当前最优个体的小概率丢失对后续进行交叉操作的不利影响,保障了其他个体能够和当前最优个体进行交叉的可能,有利于交叉操作中产生出更优质的个体。

2.3 自适应思想

传统的遗传算法使用固定的交叉率与变异率,固定的交叉率与变异率缺失对个体差异的评估,不利于保留优质个体和打破局部收敛。针对这一情况,本文引入自适应思想^[14]对交叉率与变异率进行动态调节。

具体来说,在交叉行为中,对于位于种群前 $m\%$ 最优的个体,其与更优个体的交叉概率为 $m\%$,与更劣个体的交叉概率为 $(1 - m)\%$ 。当个体越优质,其与更劣个体交叉的概率就越大,则该个体劣化概率就越大,故对于在种群中排名靠前的个体,其交叉率应动态缩小,以减小种群在交叉行为中劣化的概率,增强交叉行为对种群高质量发展的导向作用。交叉率的调整方法表示为:

$$cr_i = \begin{cases} cr_{\max} \cdot \left(1 - \frac{F_i - F_{\text{avg}} + \text{bias}}{F_{\min} - F_{\text{avg}} + \text{bias}}\right), & F_i < F_{\text{avg}} \\ cr_{\max}, & F_i \geq F_{\text{avg}} \end{cases} \quad (4)$$

其中, cr_i 表示当前种群中第 i 个个体的交叉率; cr_{\max} 表示种群个体的交叉率可以取到的最大值; F_{\min} 表示当前种群中所有个体的适应度的最小值; F_{avg} 表示当前种群中所有个体的适应度的平均值。

在传统的自适应策略中,变异率也应动态缩小,以更好地保护更优个体。然而,与交叉行为不同,变异行为中个体不参与种群内基因交换,优质个体劣化概率取决于问题候选解空间,动态缩小的策略不利于算法打破局部收敛的情形。因此,基于精英策略对历史最优解的保留,应动态增大变异率,提升优质个体获得全新基因的概率,增强优质个体的多样性与分布不确定性,拓展种群进化方向,降低陷入局部收敛的风险。变异率的调整方法表示为:

$$mr_i = \begin{cases} mr_{\min} \cdot \exp\left(\frac{F_i - F_{\text{avg}} + \text{bias}}{F_{\min} - F_{\text{avg}} + \text{bias}}\right), & F_i < F_{\text{avg}} \\ mr_{\min}, & F_i \geq F_{\text{avg}} \end{cases} \quad (5)$$

其中, mr_i 表示当前种群中第 i 个个体的变异率, mr_{\min} 表示种群个体的变异率可以取到的最小值。

在具体的交叉与变异操作中,基于历史最优解

的保留对求解质量的保障,算法采用多点交叉策略^[17]与均匀变异策略^[18],以增强算法的全局搜索能力。

2.4 多样性保持策略—干预

传统的遗传算法中对新鲜基因的引入仅依赖于小概率的变异操作,致使种群个体随迭代次数的增加而大概率趋同,种群进化停滞,易出现早熟收敛现象。通过对变异操作进行改进,增强了算法引入新鲜基因的能力,但为保持较好的收敛速度和局部搜索能力,需要平衡优质解的开发与对新空间的探索,以确保算法的稳定收敛。因此,对变异率的动态调整被限制在一个较小范围内,算法避免早熟收敛和跳出局部最优的能力仍然受限。针对这一情况,本文提出一种干预的措施,如图4所示。

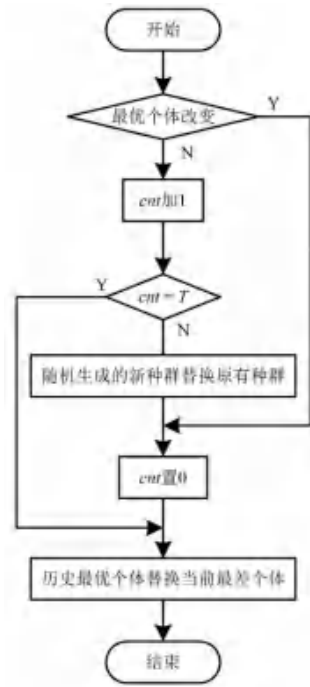


图4 干预的过程

Fig. 4 The process of intervention

如果算法达到的是局部收敛,则意味着对新空间的探索难以获得更优解,对原有优质解的开发也难以跳出局部最优。通过对种群的干预能够以大概率在整体候选解空间中引入新鲜基因,拓宽探索范围,并通过最优个体的保留保障了当前开发空间不会劣化,有利于避免种群进化陷入停滞,早熟收敛问题得以解决,增强算法跳出局部最优的能力和全局搜索能力。

2.5 多种群并行协同进化

传统的遗传算法采用单一的种群进行迭代,所有个体在同一环境下进化,不能充分探索问题的候

选解空间,多种群并行的方法被相应提出^[15]。然而,传统的多种群并行策略采用种群间迁徙的方法,较大规模迁徙存在破坏算法收敛的稳定性及打破开发优质解与探索新空间的平衡的风险,较小规模迁徙对充分探索问题候选解空间的贡献不大,方法需基于先验知识与反复试验以在算法稳定性与种群多样性间取得平衡。因此,本文提出一种基于精英种群替换的多种群并行策略。

具体来说,算法初始生成1个精英种群与 $n-1$ 个普通种群,这 n 个种群并行地迭代,每个种群的大小均为 N 。每进行 k 次迭代,将从 $n-1$ 个普通种群中各选取其前 N/n 个最优个体,以替换精英种群中的后 $(n-1) \times N/n$ 个最劣个体。种群交流的过程如图5所示。

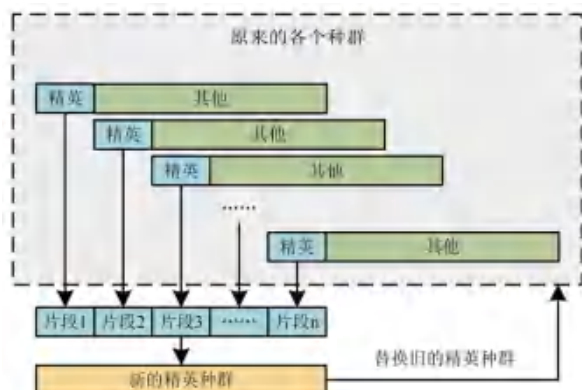


图5 种群交流的过程

Fig. 5 Process of communication between populations

通过多种群并行迭代,算法可以同时问题解空间的多个区域内进行探索,高效地寻找优质个体,提升算法的收敛速度。对精英种群与普通种群的划分,一方面由占多数的普通种群保障了算法的稳定性;另一方面,精英种群中各个体分别为各局部候选解空间中的优质解,种群可以更好地探索可能存在于多个局部候选解空间之间的其他局部候选解空间,进一步增强了算法的全局搜索能力。

3 实验验证

为验证所提算法优化超参数的能力,本文选取其他经典的智能优化算法进行对比实验。在本实验中,编程环境为 Python 3.10.12,操作系统为 Windows 10 64 bit, CPU 为 AMD Ryzen 7 6800H,内存为 32 GB 6 400 MHz。对比算法的选取为 GA^[6]、EGA^[13]、GWO^[19]、PSO^[20]、SA^[21]、CSA^[22]、DE^[23]。对于本文提出的算法,选取5个种群并行运行的 HGA-5,以探究 HGA 在一般情况下的性能表现;选

取一个种群单独运行的 HGA-1,以探究 HGA 在不支持多种群并行环境下的性能表现。

3.1 数据集的搭建

本实验通过 Python 爬虫程序和后期的人工筛选从网络上获取共 5 059 张军事目标的图像,包含飞机、装甲单位、民用车辆、直升机和步兵 5 个类别,每个类别约有 1 000 张对应的图像。数据集按 8:2 的比例划分为训练集与测试集。所有图像均通过缩放处理成 32×32 的标准尺寸,如图6所示。

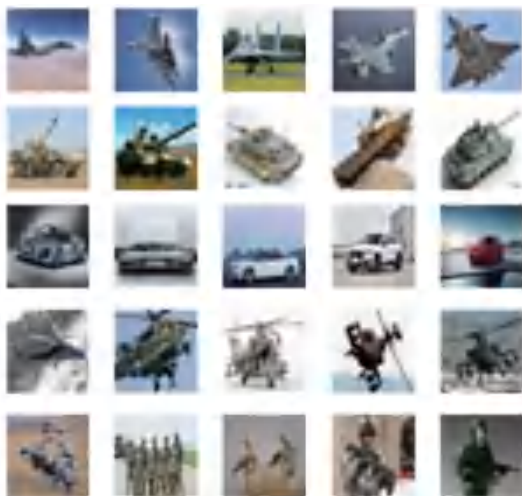


图6 部分自建军事图像数据集图片

Fig. 6 Partial images of self-constructed military image dataset

3.2 模型架构选择与图像分类过程

相较于其他图像分类任务,军事图像分类通常用于实时战场感知和目标打击决策,且所处理的信息具有较高的复杂度,对模型的复杂信息处理能力和识别速度提出较高要求。本实验使用 ResNet18^[24]作为 CNN 架构。相较于其他 CNN 架构,ResNet 创新性地引入了残差连接,表示为:

$$y_l = f(R(x_l, W_l) + h(x_l)) \quad (6)$$

其中, x_l 表示第 l 个残差块的输入; y_l 表示第 l 个残差块的输出; W_l 表示第 l 个残差块的权重参数集合; R 表示残差学习单元; h 表示恒等映射(Identity Mapping); f 表示一个可以将 $h(x_l)$ 与 $R(x_l, W_l)$ 进行整合的函数关系,实现方式包括直接叠加、使用激活函数映射等。

残差连接有效解决了深度网络训练中存在的梯度消失与梯度爆炸问题,模型能够充分利用深层架构进行训练,增强模型对复杂信息的处理能力。此外,残差连接隐含一种端到端传播机制,允许模型直接从原始像素映射到目标类别,提升模型的图像处理速度,满足军事图像分类的实时性要求。

ResNet18 进行图像分类的过程如图7所示。

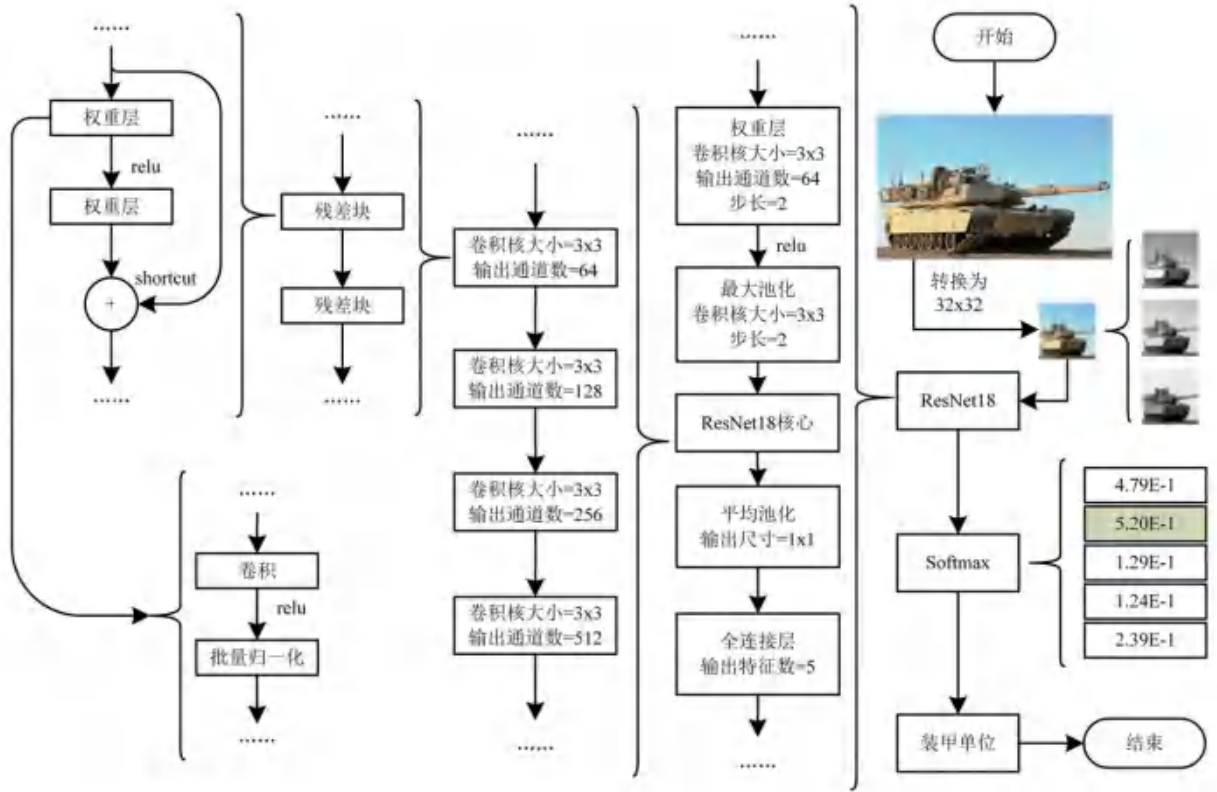


图 7 图像分类的过程

Fig. 7 The process of image classification

首先,在图像输入 CNN 模型前,图像将被处理成 32×32 的分辨率,在保留必要图像信息的同时,减少图像矩阵的信息量,缩短 CNN 模型训练过程所消耗的时间;其次,图像经过处理后将输入 CNN 模型中,经过各残差块及其他网络层由 32×32×3 的图像矩阵转换为 5 个可以象征置信度的浮点数;最后,可以通过 Softmax 函数将输出的 5 个浮点数转化为各个类别对应的概率,表示为:

$$\sigma(z_i) = \frac{\exp(z_i)}{\sum \exp(z)} \quad (7)$$

其中, z_i 表示第 i 个浮点数; $\sum \exp(z)$ 表示所有浮点数的算数和; $\sigma(z_i)$ 表示第 i 个类别对应的概率。

在图 7 的分类过程中,输入图像是一张坦克的图像,在经过识别后,得到第 2 个类别的概率最大。因此,该图像被成功地分类为装甲单位。

3.3 问题编码与目标函数构建

为了确保算法能够有效地搜索解空间并找到高质量的解,在使用智能优化算法对军事图像分类模型的超参数进行优化前,需要对问题进行合适的编码和目标函数的构建。

本实验选取了 4 类数值型超参数进行优化,分别是学习率、学习率调整策略、调整幅度与批量大

小。其中,学习率决定了模型在训练过程中权重更新的步长,本文选取 0.010、0.007、0.004 和 0.001 这 4 类学习率进行详测。批量大小决定了每次训练迭代中使用的样本数量,本文选取 16、32、64 和 128 这 4 类批量大小进行评测。以随机梯度下降 (Stochastic Gradient Descent,SGD)^[25] 为代表的非自适应优化器在机器学习中被广泛使用,这类优化器的效果直接受学习率调整策略的影响,本实验将学习率调整策略中的调整幅度分为激进与保守两类,见表 1。

表 1 学习率调整策略与调整幅度

Table 1 Learning rate adjustment strategy and adjustment range

学习率调整策略	调整幅度 (gamma)	
	保守	激进
StepLR	0.5	0.25
MultiStepLR	0.5	0.25
ExponentialLR	0.98	0.96
LinearLR	0.01	0.001
CyclicLR	0.5	0.25
OneCycleLR	5	10
CosineAnnealingLR	0.5	0.25
CosineAnnealingWarmRestarts	0.5	0.25

这4类超参数共可以组成 $4 \times 8 \times 2 \times 4 = 256$ 种组合,依次分配序号0至255。各组合分别独立地应用到 ResNet18 上运行5次,每次进行120个 epoch 的训练,任意组合在其每个 epoch 对应的模型精度取其5次独立运行的平均值,研究得到的部分组合对应错误率变化如图8所示。

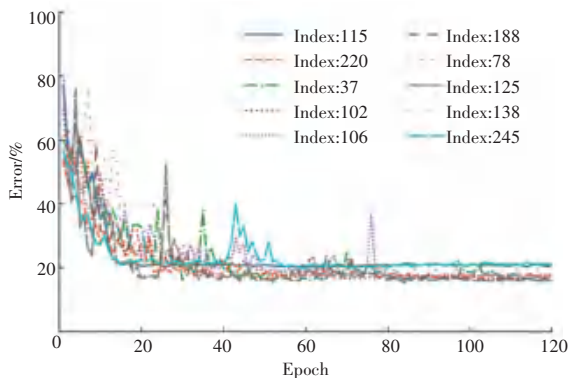


图8 部分超参数组合的收敛曲线

Fig. 8 The convergence curve of partial hyperparameter combinations

通过观察图8可知,对于几乎所有的组合,模型在后20个 epoch 时错误率已基本稳定。故取每个超参数组合对应的后20个 epoch 的模型错误率的平均值作为评价超参数优劣的指标,表示为:

$$\text{error}_i = \frac{1}{20} \sum_{j=101}^{120} (1 - \text{accuracy}_{i,j}) \quad (8)$$

其中, error_i 表示第 i 个组合在后20个 epoch 的训练中平均的模型识别的错误率, $\text{accuracy}_{i,j}$ 表示第 i 个组合在第 j 个 epoch 的训练后的模型精度。

所有 error_i 分布情况如图9所示。图9中,最小值为15.05%,平均值为18.31%,标准差为1.99E-2。

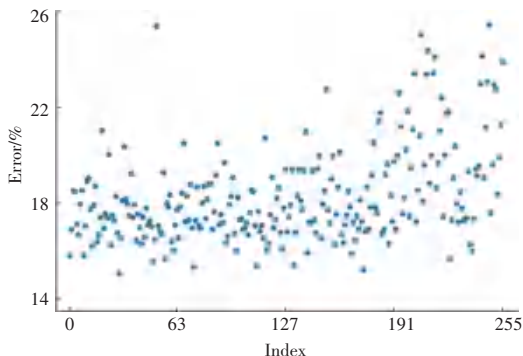


图9 各组合对应模型错误率分布图

Fig. 9 Distribution of model error rates for each combination

为使每个算法都可以编码求解问题,设定算法的编码方式为浮点数编码,解空间的上界 $\text{ub} = [3, 7, 2, 4]$,下界 $\text{lb} = [0, 0, 0, 0]$,维度 $D = 4$ 。则有:

$$\text{mul}_j = \begin{cases} 1, & j = D - 1 \\ \prod_{k=j+1}^{D-1} (\text{ub}_k + 1), & j < D - 1 \end{cases} \quad (9)$$

$$\text{seq}_i = \sum_{j=0}^{D-1} (\lfloor P_{i,j} + 0.5 \rfloor \cdot \text{mul}_j) \quad (10)$$

其中, seq_i 表示种群中第 i 个个体对应的组合的序号; $P_{i,j}$ 表示种群中第 i 个个体在第 j 个维度上的数值; mul_j 表示第 j 维的乘数; ub_k 表示第 k 个维度的上界。于是,可以得到问题的目标函数为:

$$O_i = \text{error}_s \quad \text{where } s = \text{seq}_i \quad (11)$$

当目标函数值 O_i 取到最小值,组合的序号为29,对应的超参数组合为 $[0, 3, 1, 1]$,即初始的学习率设置为0.1,学习率调整策略设置为 LinearLR,调整幅度设置为保守,训练数据的批量大小设置为64。

3.4 实验结果与分析

为保证实验的公平性与一般性,确保实验结果的可靠性与普适性,所有参与实验的算法种群大小均设置为10,即总超参数组合数的约4%,在相同的软硬件环境下独立运行1000次。此外,为了衡量算法的收敛能力,本实验中以达到问题最优解作为算法运行达到全局收敛的判别条件,通过对1000次独立运行结果进行统计来计算并比较各个算法的收敛概率,表示为:

$$\text{Probability}_i = \text{Num}_i / \text{Count}_{\text{num}} \quad (12)$$

其中, Probability_i 表示第 i 次迭代后算法的收敛率; Num_i 表示算法经过 i 次迭代后满足收敛判别条件的独立运行结果的个数; $\text{Count}_{\text{num}}$ 表示算法的独立运行总次数,本实验中设为1000。

实验结果见表2。

表2 实验结果

Table 2 Results of experiments

算法	错误率/%	收敛概率/%	耗时/s	
			单次迭代	达到收敛*
HGA-5	15.05	100	1.14E-3	8.09E-2
HGA-1	15.05	100	1.03E-3	2.10E-1
GA ^[6]	15.54	7.8	5.31E-4	-
EGA ^[13]	15.06	56.2	5.51E-4	-
GWO ^[19]	15.41	17.2	3.26E-4	-
PSO ^[20]	15.37	37.9	2.87E-4	-
SA ^[21]	15.05	100	5.00E-3	9.15E-1
CSA ^[22]	15.06	99.2	4.62E-4	2.52E-1
DE ^[23]	15.37	18.9	8.14E-4	-

* 以收敛概率达到95%作为算法达到收敛的判别条件

此外,为了更直观地展示实验结果,以更好理解各算法的性能表现,本实验还对各算法的收敛曲线和耗时情况进行了绘制,如图 10 和图 11 所示。

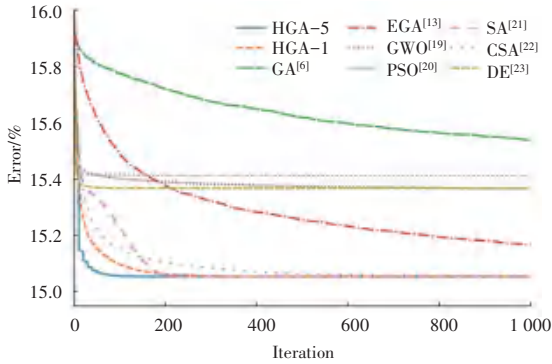


图 10 错误率收敛曲线

Fig. 10 Convergence curve of error rate

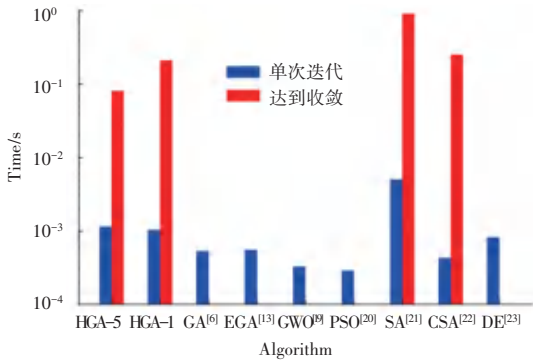


图 11 各算法耗时情况

Fig. 11 Time cost of algorithms

通过分析表 2 及图 10 与图 11,不难发现,本文所提出的 HGA 在解决卷积神经网络超参数优化的问题上优势明显,主要体现在以下 3 个方面:

(1) 在机器学习中,模型的错误率是衡量其预测精度和在未知数据上泛化能力的主要指标。根据表 2 结果可知,除 HGA 外的其他 7 类算法所得的平均错误率约为 15.27%,则 HGA 所得的超参数组合在应用到 CNN 模型上后的识别错误率平均降低了 $(15.27\% - 15.05\%) / 15.27\% \approx 1.41\%$ 。这表明, CNN 模型在通过 HGA 优化超参数组合后能够具有更高的精度和更强的泛化能力。

(2) 虽然多数算法在理论上均具有达到全局收敛能力,但在实际应用中往往要求能够在有限时间内被求解,而更高的收敛概率意味着算法在有限次迭代内求解问题的能力更强。根据表 2 可知,除 HGA 外的其他 7 类算法所得的平均收敛概率约为 48.2%,则 HGA 所得的全局收敛概率平均提升了 $(100\% - 48.2\%) / 48.2\% \approx 107.59\%$,可以较为稳定地在有限迭代内求解卷积神经网络数值型超参数

优化的问题,即在军事图像识别的特定应用中具有较好的实用价值。

(3) 虽然 HGA 每一次迭代所消耗的时间较长,但 HGA 的优秀搜索策略保障了高效的搜索效率,使得算法以 95% 的概率在 70 次迭代时达到全局收敛,耗时 $8.09E-2$ s,具有优秀的全局收敛速度。

4 结束语

针对军事图像分类中 CNN 超参数的优化问题,本文提出了一种改进的遗传算法。在传统遗传算法的基础上,对适应度函数进行了设计,引入了精英策略、自适应思想和多种群并行的思想,还提出针对种群的干预措施。其中,对适应度函数进行适当的映射转换能够使算法可以更好地评价个体的优劣,引导种群向更有前途的区域搜索;精英策略的引入保障了优秀解与优质基因的传承,有助于提高算法的收敛速度和解的质量;自适应思想抑制了固定参数可能带来的局限性,提高算法的灵活性与打破局部收敛情况;多种群并行策略增加了算法的多样性,减少其陷入局部收敛的概率,提高了算法的全局收敛能力;干预措施使得算法在任意状态下均能探索问题候选解空间,从而进一步增强算法跳出局部收敛的能力。

通过在军事图像分类问题上的实验结果表明,相较于传统遗传算法和其他一些经典智能优化算法,本文所提出的 HGA 在收敛速度和稳定性上有着明显的提升,为解决军事图像识别中的 CNN 超参数优化问题提供一种有效的解决方法。

参考文献

- [1] MAURÍCIO J, DOMINGUES I, BERNARDINO J. Comparing vision transformers and convolutional neural networks for image classification: A literature review [J]. Applied Sciences, 2023, 13(9): 5521.
- [2] ALZUBAIDI L, ZHANG J, HUMAIDI A J, et al. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions [J]. Journal of Big Data, 2021, 8:53.
- [3] DHILLON A, VERMA G K. Convolutional neural network: A review of models, methodologies and applications to object detection [J]. Progress in Artificial Intelligence, 2020, 9(2): 85-112.
- [4] REDDI S J, KALE S, KUMAR S. On the convergence of adam and beyond [J]. arXiv preprint arXiv,1904.09237, 2019.
- [5] SMITH L N, TOPIN N. Super-convergence: Very fast training of neural networks using large learning rates [C]// Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications. San Francisco, CA: SPIE, 2019, 11006: 369-386.

- [6] HOLLAND J H. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence* [M]. Cambridge, MA: MIT Press, 1992.
- [7] 徐宗本, 聂赞坎, 张文修. 遗传算法的几乎必然强收敛性; 鞅方法 [J]. 计算机学报, 2002, 25(8): 785-793.
- [8] KATOCH S, CHAUHAN SS, KUMAR V. A review on genetic algorithm; Past, present, and future [J]. *Multimedia Tools and Applications*, 2021, 80: 8091-8126.
- [9] 余维, 李阳, 钟李红, 等. 基于改进实数编码遗传算法的神经网络超参数优化 [J]. 计算机应用, 2024, 44(3): 671-676.
- [10] 邓翔, 彭杰, 吕一清. 基于量子遗传算法优化的新 Prophet 模型及其验证 [J]. 系统工程, 2022, 38(5): 141-150.
- [11] 杜沅昊, 刘媛华. 混合遗传蚁群算法优化 BP 神经网络预测空气质量 [J]. 计算机系统应用, 2023, 32(4): 223-230.
- [12] NIKBAKHT S, ANITESCU C, RABCZUK T. Optimizing the neural network hyperparameters utilizing genetic algorithm [J]. *Journal of Zhejiang University-SCIENCE A (Applied Physics & Engineering)*, 2021, 22(6): 407-426.
- [13] DEJONG K A. *An analysis of the behavior of a class of genetic adaptive systems* [M]. Ann Arbor: University of Michigan Press, 1975.
- [14] SRINIVAS M, PATNAIK L M. Adaptive probabilities of crossover and mutation in genetic algorithms [J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1994, 24(4): 656-667.
- [15] CANTÚ-PAZ E. A survey of parallel genetic algorithms [J]. *Calculateurs Paralleles, Reseaux et Systems Repartis*, 1998, 10(2): 141-171.
- [16] DARWIN C. *Origin of the species* [M]// *British Politics and the Environment in the Long Nineteenth Century*. London: Routledge, 2023: 47-55.
- [17] EIBEN A E, RAUE P E, RUTTKAY Z. Genetic algorithms with multi-parent recombination [C]// *Proceedings of International Conference on Parallel Problem Solving From Nature (PPSN)*. Cham: Springer, 1994: 78-87.
- [18] LIM S M, SULTAN A B M, SULAIMAN M N, et al. Crossover and mutation operators of genetic algorithms [J]. *International Journal of Machine Learning and Computing*, 2017, 7(1): 9-12.
- [19] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer [J]. *Advances in Engineering Software*, 2014, 69: 46-61.
- [20] KENNEDY J, EBERHART R. Particle swarm optimization [C]// *Proceedings of ICNN '95 - International Conference on Neural Networks*. Piscataway, NJ: IEEE, 1995: 1942-1948.
- [21] KIRKPATRICK S, GELATT JR C D, VECCHI M P. Optimization by simulated annealing [J]. *Science*, 1983, 220(4598): 671-680.
- [22] YANG Xinshe, DEB S. Cuckoo search via Lévy flights [C]// *World Congress on Nature & Biologically Inspired Computing (NaBIC)*. Piscataway, NJ: IEEE, 2009: 210-214.
- [23] STORN R, PRICE K. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces [J]. *Journal of Global Optimization*, 1997, 11: 341-359.
- [24] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Piscataway, NJ: IEEE, 2016: 770-778.
- [25] 李兴怡, 岳洋. 梯度下降算法研究综述 [J]. *软件工程*, 2020, 23(2): 1-4.