

# 多类在线 Boosting 的小样本细粒度图像目标识别算法

闵小翠<sup>1</sup>

MIN Xiaocui

## 摘要

在图像识别领域,传统方法处理小样本细粒度图像时局限于单模态,忽略了图像内部联系,导致识别效果差。为此,文章提出了一种针对小样本细粒度图像的多类在线 Boosting 目标识别算法。该算法先对图像进行预处理和分割,用双随机矩阵打乱碎片后重建图像。在重建图上提取细粒度特征,通过深度学习模型将这些特征映射至高维空间形成有效特征表示。这些特征随后被输入多类在线 Boosting 算法,结合多个弱学习器并更新模型,挖掘特征关联性,构建跨模态语义关联,实现精确识别。实验结果显示,对 200 组细粒度图像识别时,本算法在不同测试数据集上识别效果稳定,TP 识别数均在 192 以上,最高达 194。与对照组相比,本算法在识别可靠性和有效性方面优势显著。

## 关键词

多类在线 Boosting; 小样本细粒度图像; 目标识别; 拼图排列解算; 双随机矩阵

doi: 10.3969/j.issn.1672-9528.2025.01.031

## 0 引言

随着计算机视觉技术的飞速发展,小样本细粒度图像目标识别因其独特的挑战性和广泛的应用前景,吸引了越来越多研究者的关注<sup>[1]</sup>。细粒度图像识别要求算法不仅能够区分不同大类之间的物体,更需具备在高度相似的细粒度类别间识别出微小差异的能力。然而,在实际应用中,即便属于同一类别的物体,也可能因拍摄时光照条件、视角变化、遮挡情况等因素而展现出显著的视觉差异。这种类内变异性极大地考验了识别算法的鲁棒性和适应性<sup>[2-3]</sup>。

目前,针对小样本细粒度图像分类的研究已经取得了一定进展。文献[4]提出一种以特征融合及混合注意力为基础的小目标识别方法,通过多尺度特征融合解决特征信息不足,引入混合注意力机制增强有效信息,建立远距离

依赖,提高识别准确性。使用 Resnet50 优化网络结构,提高检测精度并保持速度。但该方法未考虑图像内部各分区相互影响,注意力机制过于关注局部特征,导致复杂或遮挡情况下误判或漏检。文献[5]提出递归标记的弱点目标识别法,通过递归细化目标,提高了识别敏感性,且能在弱监督下用少量标注信息实现识别。但该方法忽视目标间空间关系,导致模型缺乏全局视角,难以应对复杂场景。文献[6]提出基于张量 Tucker 分解的高光谱图像异常目标识别。该方法将图像转为张量,利用 Tucker 分解识别异常。但在小样本细粒度图像中,异常多为类别间细微差异,仅依赖 Tucker 分解因忽略这些差异而性能受限。

为了解决上述方法存在的问题,提出一种多类在线 Boosting 的小样本细粒度图像目标识别算法。

## 1 小样本细粒度图像目标识别算法设计

### 1.1 基于双随机矩阵扰动的图像分割与重组算法

细粒度图像拼图排列解算技术,通过巧妙地重排图像元素,能够生成多样化的排列组合,从而有效扩增“虚拟样本”库。这一过程中,图像首先被精细地分解,随后基于双随机矩阵扰动原理进行重新组合。双随机矩阵的引入,不仅增加了图像重组的随机性和多样性,还使得重组后的图像能够揭示出更多潜在的特征关联。

在重组后的图像集合中,通过对这些图像进行深入分析,可以构建出具有强表征能力的特征向量。这些特征向量能够更全面地捕捉图像中的关键信息,为实现精准重组提供有力

1. 广州华立科技职业学院 广东广州 511325

[基金项目] 广东省教育厅 2024 年度普通高校重点科研平台和项目(2024ZDZX4104); 广东省教育厅 2022 年度广东省普通高校特色创新类项目“基于互联网+共享农场系统设计与实现”(2022KTSCX372); 广州市科学技术局 2024 年度农村科技特派员专题“铁皮石斛附生乌榄树生态种植技术的应用推广”(2024E04J1274); 广州市增城区科技创新发展扶持资金项目(经济和社会科技专题项目)(2024ZCKJ13); 广东省教育厅第一批省高职院校高水平专业群:软件技术专业群(GSPZYQ2020082)”; 2022 年中国高校产学研创新基金—新一代信息技术创新项目(2022IT235)

支持。此外，该方法在目标识别领域具有显著优势，因为它能够显著提高特征刻画的准确性，进而提升识别的精度和鲁棒性。

对细粒度图像进行拼图排列解算时，核心是用解算模块将图像分割成  $n \times n$  碎片，用  $n^2 \times n^2$  双随机矩阵打乱顺序，再通过解算模块恢复排列，重建图像<sup>[7-9]</sup>。具体流程见图 1。

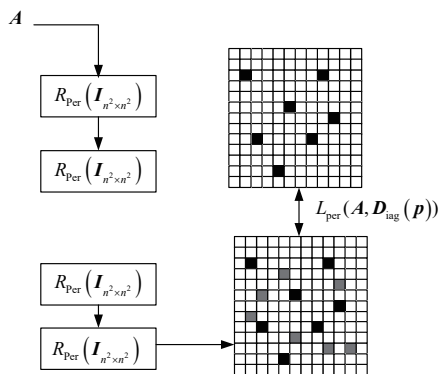


图 1 细粒度图像拼图排列解算模结构

在双随机矩阵构建阶段，设计其行和列的元素之和均为 1，且所有元素非负。在拼图排列中，将其视为一个概率分布矩阵，指导碎片随机分配<sup>[9]</sup>。双随机矩阵为：

$$A = R_{\text{Per}}(I_{n^2 \times n^2}) \times D_{\text{diag}}(p) \quad (1)$$

式中： $I_{n^2 \times n^2}$  表示  $n^2 \times n^2$  的单位矩阵； $R_{\text{Per}}(I_{n^2 \times n^2})$  表示生成细粒度图像像素行（或列）重排单位矩阵的函数； $D_{\text{diag}}(p)$  表示由概率向量  $p$  生成的对角矩阵。

通过双随机矩阵的逆映射过程，缩小双随机矩阵和概率向量矩阵之间的分布差距<sup>[10]</sup>。逆映射损失函数为：

$$L_{\text{per}} = \sum (A, D_{\text{diag}}(p)) = - \sum_{i=1}^{n^2} \sum_{j=1}^{n^2} [p_{ij} \log(A_{ij}) + (1 - p_{ij}) \log(1 - A_{ij})] \quad (2)$$

式中： $L_{\text{per}}$  表示细粒度图像双随机矩阵的逆映射损失函数； $p_{ij}$  表示细粒度图像双随机矩阵中元素  $(i, j)$  的概率向量； $A_{ij}$  表示细粒度图像双随机矩阵中元素  $(i, j)$  的全局分配结果。

## 1.2 重建图像上的细粒度特征高效表示

原始图像经过预处理、分割、打乱恢复后，内部结构关系重新调整，细微特征更明显。本研究采用特征提取函数  $F$  来表示这一

过程，得到的特征  $f$  表示为：

$$f = F(L) \quad (3)$$

利用深度学习模型特征  $f$  进行高效表示学习，形成紧凑有效的特征表示  $r'$ 。这个过程可以通过一个深度网络  $G$  来描述，其中包含了多层非线性变换：

$$r' = G(f) \quad (4)$$

## 1.3 多类在线 Boosting 算法的特征关联挖掘与精确识别

将提取的特征输入多类在线 Boosting 算法，该算法具备处理复杂数据结构的强大能力，适用于加载双模态数据集。在算法启动之初，首先初始化各项关键参数，为后续的迭代优化奠定坚实基础。在此基础上，算法深入挖掘单模态内部以及跨模态之间的语义关联，这些关联信息对于精准捕捉图像中的关键特征至关重要，是实现高效目标识别的核心所在。

结合细粒度图像处理技术，多类在线 Boosting 算法通过在线迭代的方式不断优化其内部的 Boosting 模型。每一次迭代都旨在提升模型的识别性能，使其能够更加准确地识别目标对象。在流程上，如图 2 所示，迭代过程始于对图像与元素的 Boosting 初始化。随后，进入核心迭代环节，算法每次都会集中优化一个 Boosting 模型，并根据当前的数据分布以及算法自身的反馈，动态地调整各项参数，以确保模型的持续增强。

值得注意的是，在每一次迭代过程中，预测器都会根据最新的模型状态进行更新。同时，迭代计数器也会相应地递增，以记录算法的迭代进度。这一过程将持续进行，直至满足预设的停止条件，如达到最大迭代次数、识别精度达到预设阈值等。

此外，为了进一步提升算法的识别性能和稳定性，还可以在迭代过程中引入正则化技术、早停策略等优化

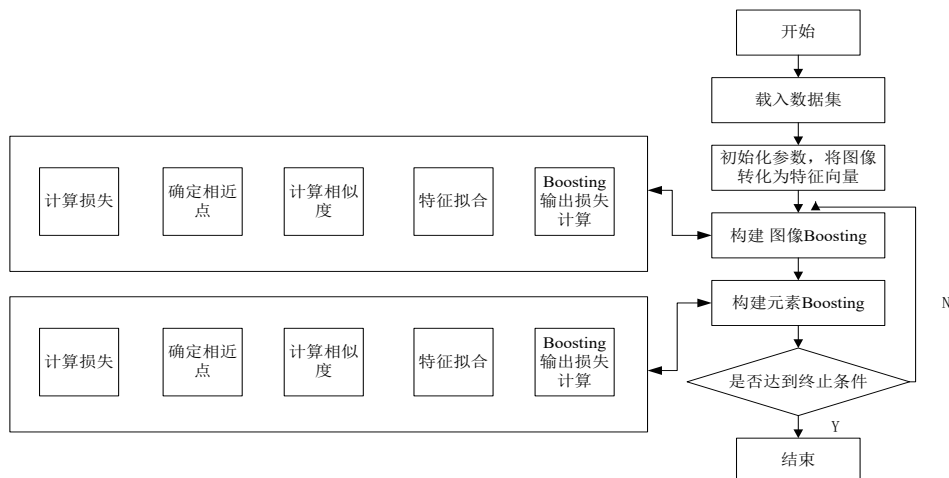


图 2 多类在线 Boosting 目标识别算法流程

手段。这些手段有助于防止模型过拟合，确保算法在迭代过程中能够逐步收敛到最优解。同时，通过引入多样化的特征选择和融合方法，算法还能够更好地利用双模态数据集中的互补信息，进一步提升目标识别的准确性和鲁棒性。

Boosting 训练阶段包括以下步骤：

(1) 数据预处理与特征降维。利用设计的损失函数指导学习，初始损失计算见式(2)。定义  $\mathbf{v}_n$  为对应于第  $n$  个元素值的特征向量，细粒度图像表示为输入元素的行向量函数，具体为：

$$F = L_{\text{per}} \left[ \langle f, \mathbf{v}_n \rangle \mathbf{v}_1 + \sum \langle f, \mathbf{v}_n \rangle \mathbf{v}_n \right] \quad (5)$$

式中： $F$  表示针对每个输入元素的行向量函数构成的细粒度图像， $\langle f, \mathbf{v}_n \rangle$  表示细粒度图像内在的结构与特征分布， $\sum \langle f, \mathbf{v}_n \rangle \mathbf{v}_n$  表示整体细粒度图像中各元素值的特征向量关联。

(2) 特征向量映射与类别标记。按照式(5)将细粒度图像表示为行向量函数，考虑图像结构与特征分布，以及元素间特征向量的关联，为图像中的每个点分配明确的类别标记。

(3) 局部相似度量优化。设定阈值并筛选出与特定点距离小于该阈值的所有点。通过比较不同局部相似度量下的损失值，选择最佳局部相似度量，迭代移除已处理点。

## 2 测试与分析

### 2.1 测试准备

本文选用了 4 个细粒度图像的公开数据集作为具体的测试数据来源。数据集的具体构成如表 1 所示。

表 1 测试数据集准备

| 构成         | 测试细粒度图像数据集      |                  |                   |                  |
|------------|-----------------|------------------|-------------------|------------------|
|            | CUB<br>200-2011 | Stanford<br>Cars | FGVC-A<br>ircraft | Stanford<br>Dogs |
| 总规模 / 张    | 11 895          | 16 243           | 10 216            | 20 164           |
| 种类数 / 类    | 206             | 197              | 125               | 224              |
| 训练数据规模 / 张 | 1500            | 1000             | 800               | 2000             |
| 测试数据规模 / 张 | 200             | 200              | 200               | 200              |

### 2.2 测试方案设计

引入基于特征融合与混合注意力算法，以及基于递归标记算法作为对照组。本次研究的识别任务聚焦于精确区分图像中的目标对象类别。样本目标图像如图 3。

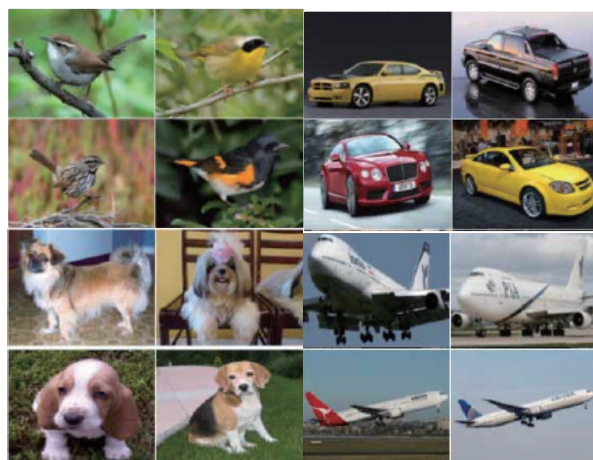


图 3 样本图像

### 2.3 图像重建性能分析

从图 3 中的样本图像集中选取一张，本算法对原始图像重建结果如图 4 所示。



图 4 图像重建结果

从图 4 中可以看出，基于小样本细粒度图像中目标对象的像素分布特性，本算法采用双随机矩阵成功重建。从视觉效果上评估，重建图像在保持结构完整性、细节丰富度以及色彩准确性方面均展现出优异性能。

### 2.4 对比测试结果与分析

对于测试结果的分析与评价，本文采用真实的正样本 (true positive, TP) 和虚假的正样本 (false positive, FP) 作为具体的评价指标。其中，TP 值越大且 FP 值越小，表示识别效果越好；反之，TP 值越小且 FP 值越大，表示识别效果越差。

测试结果如表 2 所示。结合表 2 对三种不同识别算法的性能进行分析。特征融合算法针对不同测试数据集的识别效果存在明显差异，TP 最大值在 FGVC-Aircraft 数据集达到 194，而最小值在 CUB 200-2011 数据集仅为 179，适用范围存在一定的局限性。递归标记算法具有较高的稳定性，但 TP 均处于 180~185 区间内，存在提升的空间。本文设计算法不



仅对不同测试数据集的识别效果稳定，且 TP 均在 192 以上，最大值达到 194，与对照组相比，在可靠性和有效性方面均具有显著优势。本文算法优势在于提取具有判别性的细粒度特征并进行高效表示学习。在面对不同的测试数据集时，这种高质量的特征表示能够更准确地反映出目标的本质特征，从而提高了识别的准确性。

表 2 不同算法测试结果对比表

| 识别算法   | 识别结果 | 测试细粒度图像数据集      |                  |                   |                  |
|--------|------|-----------------|------------------|-------------------|------------------|
|        |      | CUB<br>200-2011 | Stanford<br>Cars | FGVC-A<br>ircraft | Stanford<br>Dogs |
| 特征融合算法 | TP   | 179             | 188              | 194               | 186              |
|        | FP   | 21              | 12               | 6                 | 14               |
| 递归标记算法 | TP   | 183             | 184              | 182               | 182              |
|        | FP   | 17              | 16               | 18                | 18               |
| 本文设计算法 | TP   | 192             | 194              | 193               | 194              |
|        | FP   | 8               | 6                | 7                 | 6                |

为了进一步验证本文提出的小样本细粒度图像目标识别算法的可靠性，选取了召回率作为关键实验指标，并将特征融合算法、递归标记算法以及本文设计的算法应用于测试集中的目标识别任务，结果如图 5 所示。通过对图 5 实验结果的分析可知，与传统算法相比，本文算法的识别召回率更趋近于 1。这一结果表明，在所有实际为正类的样本里，本文算法能够更为精准地识别出正类样本。具体分析可知，本文算法的高召回率主要在重建的原始图像基础上，提取的细粒度特征通过深度学习模型进行高效表示学习，并被映射至高维空间。这一过程极大地丰富了特征的维度和表达能力，形成了更加紧凑且有效的特征表示，为后续的分类识别提供了高质量的特征输入。

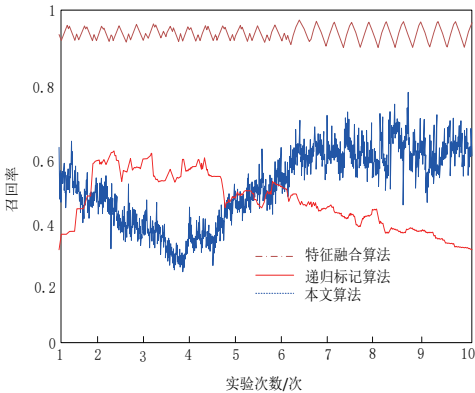


图 5 不同算法的目标识别召回率

3 结语

在小样本细粒度图像识别中，传统方法受限于单模态及忽略图像内部联系，难以精准识别。本研究提出的多类在线

Boosting 算法为解决此问题提供新思路。该算法通过独特的图像预处理（分解、打乱、恢复）、特征提取以及多类在线 Boosting 操作（组合弱学习器挖掘特征关联性与语义关联），实现了精确识别。实验中对 200 组图像的识别结果表明，本算法在不同数据集上稳定且 TP 数较高，可靠性和有效性优势明显。这表明集成多个分类器并在线动态调整的方式，能有效应对小样本细粒度图像的类间细微差异和类内多变性，从而提高识别精度。虽然本研究针对小样本细粒度图像识别取得了一定成果，但样本的多样性有限。未来研究将收集来自不同来源的小样本细粒度图像数据，通过融合这些多源数据，增加样本的多样性，提高算法的泛化能力。

参考文献：

[1] 刘光花, 杨发顶, 程亚伟, 等. 基于深度神经网络的火箭图像目标识别与跟踪 [J]. 沈阳航空航天大学学报, 2024, 41(4): 59-66.

[2] 沈晓兵. 基于机器视觉的防止电力操作走错间隔图像识别模型 [J]. 微型电脑应用, 2024, 40(8): 108-112.

[3] 王悠, 韩立祥, 付贵. 基于改进 YOLOv5s 的航拍红外图像目标识别方法 [J]. 红外技术, 2024, 46(7): 775-781+801.

[4] 严荣慧, 郭前, 雷鸣, 等. 基于特征融合及混合注意力的小目标船舶识别研究 [J/OL]. 中国舰船研究, 1-10[2024-04-12].<https://doi.org/10.19693/j.issn.1673-3185.034891>.

[5] 潘文, 周波, 曹志浩. 基于递归标记的红外图像弱小点目标自动识别方法 [J]. 遵义师范学院学报, 2024, 26(4): 85-89.

[6] 陈楚申, 唐国吉. 基于张量 Tucker 分解的高光谱图像异常目标识别 [J]. 现代电子技术, 2024, 47(13): 43-46.

[7] 黄东. 基于多模态特征融合的无人机测绘图像目标识别研究 [J]. 资源导刊, 2024(12): 27-30+34.

[8] 王芳, 李慧涛, 王凯, 等. 改进 PP-YOLOv2 的水下侧扫声呐图像多目标识别 [J]. 南开大学学报 (自然科学版), 2024, 57(3): 1-7.

[9] 任永梅, 蔡理昌, 龙亚, 等. 基于点特征直方图的道路场景点云目标识别方法 [J]. 电脑知识与技术, 2024, 20(17): 36-38.

[10] 曾祥书, 黄一飞, 蒋忠进. 基于改进 YOLOv5 卷积神经网络的 SAR 图像目标识别 [J]. 现代雷达, 2024, 46(2): 138-145.

【作者简介】

闵小翠（1984—），女，湖南益阳人，硕士，副教授，研究方向：计算机应用。

（收稿日期：2024-09-24）