

## 多尺度 - 多方向 Transformer 下模糊遥感图像小目标识别方法

刘 婷<sup>1</sup>  
LIU Ting

## 摘 要

模糊遥感场景中, 大气散射与传感器噪声耦合干扰导致图像非均匀退化, 呈现局部多尺度模糊纹理交叠和目标边缘频域信息混叠现象。此复杂退化模式使传统单尺度特征提取方法难以有效分离目标与背景频谱成分, 且多方向边缘特征退化畸变加剧目标语义信息断裂, 影响识别精度与鲁棒性。为此, 文章提出了一种多尺度 - 多方向 Transformer 模糊遥感图像小目标识别方法。该方法通过传感器标定与平滑去噪实现大气校正, 利用非线性隶属度映射结合直方图均衡化增强图像对比度, 缓解频谱混叠现象; 采用卷积神经网络堆叠提取目标特征, 初步分离目标与背景的粗粒度信息; 引入多尺度 - 多方向 Transformer 编码器, 通过多头自注意力机制与方向可学习卷积核并行处理多尺度特征, 实现特征自适应增强与多方向边缘建模, 有效抑制语义信息断裂。基于交叉熵损失函数输出类别概率分布, 完成小目标识别。在 DOTA-F 数据集上的实验结果表明, 该方法 Kappa 系数达到 0.9, 显著提升了模糊遥感图像中小目标的识别精度。

## 关键词

多尺度 - 多方向 Transformer; 模糊遥感图像; 小目标识别; 识别精度

doi: 10.3969/j.issn.1672-9528.2026.03.016

## 0 引言

目标识别是将特定对象和其他对象区分开的过程, 目前目标识别在航天侦察和交通监测、城市规划以及海洋管理等领域都得到了广泛的应用, 但由于遥感图像拍摄距离较远、成像条件复杂等原因, 使得图像分辨率不高, 存在模糊现象, 另外, 目标在图像中的像素占比较小、特征不明显, 进一步增加了识别难度。因此, 如何有效降低复杂环境下模糊遥感图像小目标识别的错误率, 成为当下需要解决的问题。当前基于遥感图像目标识别的方法主要利用人工设计特征再结合分类器进行检测, 但人工设计特征难以全面捕捉小目标的复杂特征, 且图像模糊的鲁棒性较差, 在模糊情况下, 特征提取的准确性大幅下降, 导致识别精度不高。

针对遥感图像目标识别需求, 现有研究提出了多种技术路线。周末<sup>[1]</sup>采用多尺度图像分块策略, 结合特征金字塔网络生成旋转边界框进行目标识别, 但在处理模糊遥感图像时, 该方法难以有效应对局部多尺度模糊纹理交叠问题, 不同分块区域的模糊干扰相互影响, 导致目标与背景频谱成分混杂, 影响识别精度。史宝岱等<sup>[2]</sup>通过构建标注

训练集, 利用卷积神经网络提取特征并完成目标识别, 然而其卷积和池化操作在应对模糊图像中的纹理交叠和边缘频域混叠时存在局限, 卷积核同时捕获目标与背景的模糊特征, 造成频谱混淆, 影响目标定位与分类准确性。为此, 本文引入多尺度 - 多方向 Transformer 模型, 旨在提升模糊遥感图像中小目标的识别精度, 为相关应用提供更可靠的遥感信息支撑。

## 1 模糊遥感图像小目标识别方法设计

## 1.1 基于非线性隶属度映射的模糊图像分辨率增强

在模糊遥感图像中, 微小目标因像素信息少, 边缘和纹理特征易被背景噪声掩盖, 识别困难。有效提升目标与背景对比度, 得到高分辨率图像, 为后续多层级卷积特征融合提供优质数据, 大幅提升微小目标识别精度。

首先利用传感器定标方法对采集的模糊遥感图像进行大气校正, 去除辐射畸变, 使图像数据更准确地反映地物真实的辐射信息, 满足后续处理的要求, 进而对图像进行平滑处理, 消除噪声<sup>[3]</sup>。计算公式为:

$$\begin{cases} I = \frac{\pi p d^2}{Q \cos \theta} \\ N = f(x, y) \otimes \left( -\frac{lw^2}{X} \right) \end{cases} \quad (1)$$

式中:  $p$  表示激光光谱传感器的记录值;  $d$  表示传感器增益系

1. 商丘职业技术学院计算机工程学院 河南商丘 476000

[基金项目] 河南省高等学校重点科研项目“智慧教育背景下基于学习者画像的个性化学习路径研究”(24B520031)

数;  $Q$  表示日地距离修正系数;  $\theta$  表示太阳天顶角;  $l$  表示地表反射率;  $f(x, y)$  表示辐射校正后图像中任意像素点的位置;  $w$  表示图像中噪声的标准差;  $X$  表示高斯核标准差, 控制平滑程度;  $N$  表示去噪后图像。

遥感图像受光照、地形、大气等因素影响, 线性特征往往呈现非线性分布, 因此, 可将处理后的遥感图像映射到线性可分空间中, 利用非线性变换选取图像的线性隶属度<sup>[4]</sup>。即:

$$\beta_{i,j} = \frac{u_{i,j} - u_{\min}}{N} \int_0^v \left(1 + \frac{\varepsilon}{e}\right) |k + g| dx' \quad (2)$$

式中:  $u_{i,j}$  表示图像  $(i, j)$  位置处的像素值;  $u_{\min}$  表示高频子带系数最小值;  $v$  表示线性可分空间尺度;  $\varepsilon$  表示缩放系数;  $e$  表示边缘强度的最小值;  $k$  表示控制隶属度函数斜率的参数, 值越大, 函数曲线越陡, 隶属度变化越快, 值越小, 函数曲线越平缓, 隶属度变化越慢;  $g$  表示函数的中心参数;  $\beta_{i,j}$  表示图像  $(i, j)$  位置处像素点的隶属度。

通过计算所有像素的非线性隶属度, 并依据隶属度对图像进行变换, 以明确其模糊对比度, 在此基础上, 利用直方图均衡化处理将模糊区域的像素均匀地扩散到整个图像, 实现对图像的非线性延伸, 获取高分辨率图像。描述公式为:

$$\begin{cases} s = (L-1) \sum_{q=1}^L \frac{\beta_{i,j}}{\chi} \\ F = \text{round}(s - \mu) \cdot \lambda \end{cases} \quad (3)$$

式中:  $L$  表示图像中灰度级数量;  $\chi$  表示图像高频子带数量;  $s$  表示模糊对比度;  $\text{round}(\cdot)$  表示四舍五入, 保证灰度值始终为整数;  $\mu$  表示图像的最大灰度级数;  $\lambda$  表示灰度级出现的概率;  $F$  表示分辨率增强后的遥感图像。

## 1.2 基于多层级卷积特征融合的小目标特征提取

虽然对遥感图像进行了分辨率增强处理, 提高了图像的可用性, 但图像目标尺寸小使得目标与背景的区分度较差, 因此, 需要对目标的特征进行提取。以卷积神经网络为基础, 通过卷积层堆叠的方式提取目标特征<sup>[5]</sup>。堆叠过程可表示为:

$$h = \frac{1}{m} F + \varphi I \quad (4)$$

式中:  $m$  表示卷积层个数;  $F$  表示增强后的遥感图像;  $\varphi$  表示特征图尺寸;  $I$  表示池化窗口;  $h$  表示最终输出的特征图。

将图像分割为若干大小的单元格, 依据每个单元格像素值大小将灰度级投影到卷积层中, 求取图像的垂直梯度。公式为:

$$Y_z = \frac{1}{h} \ln \left[ \frac{v^2 + g^2}{2v\omega} \right] \quad (5)$$

式中:  $v$  表示傅里叶描述子;  $g$  表示图像的行程;  $v$  表示低阶几何矩;  $\omega$  表示目标的偏心率;  $Y_z$  表示单元格  $z$  的垂直梯度。

将图像单元格的垂直梯度进行拼接, 并采用堆叠卷积层计算多个层级的特征图, 由此获取图像小目标特征<sup>[6]</sup>。公式为:

$$W = \sum R \times (Y_z + b + T) \times \sqrt{\phi} \quad (6)$$

式中:  $R$  表示输出通道数;  $b$  表示池化层维度;  $T$  表示输入特征向量;  $\phi$  表示展平后的特征图维度;  $W$  表示小目标特征参数。

设置高斯核标准差  $X$  从 0.5 至 4.0 共 8 个等级,  $X$  值越大表示模糊程度越严重。在每个模糊等级下, 分别采用四种方法提取特征并计算其特征一致性指数 (FCI): 方法 A 为本文多层级卷积特征融合方法; 方法 B 为单层卷积特征提取方法; 方法 C 为基于 SIFT 与 HOG 的传统手工特征方法; 方法 D 为基于 PCA 降维的全局特征方法。FCI 定义为当前模糊等级下特征向量与清晰图像 ( $X=0$ ) 特征向量的余弦相似度, FCI 越高表明特征稳定性越好。不同方法的特征一致性随模糊程度的变化趋势如图 1 所示。

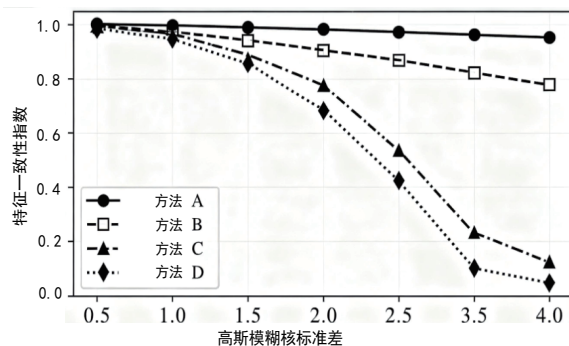


图 1 不同方法的特征一致性随模糊程度的变化趋势图

由图 1 可知, 所有方法的特征稳定性均随模糊程度加剧而下降, 表现为四条曲线整体呈下降趋势。本文方法 (方法 A) 的稳定性显著优于其他方法, 在轻度模糊 ( $X < 2.0$ ) 时 FCI 始终维持在 0.9 以上, 曲线最为平缓, 下降幅度极小; 在重度模糊 ( $X > 2.0$ ) 时其曲线仍位于最上方, 即使在  $X=4.0$  条件下 FCI 仍高于 0.90; 方法 B (单层卷积) 稳定性次之, 曲线下降斜率更大, 表明其感受野有限, 对模糊更敏感; 方法 C (SIFT+HOG) 在  $X > 1.5$  后 FCI 急剧下降, 因手工特征依赖的边缘与角点信息在模糊中快速退化; 方法 D (PCA 特征) 稳定性最差, 曲线始终处于最低位置且迅速下滑, 说明全局特征对局部结构变化缺乏鲁棒性。

## 1.3 基于多尺度 - 多方向 Transformer 特征增强的小目标识别

传统单尺度特征提取方法在处理这种复杂情况时, 难以

有效分离目标与背景频谱成分，多方向边缘特征退化畸变还会进一步加剧目标语义信息断裂。而多尺度 - 多方向 Transformer 特征增强基于注意力机制，能够联合多尺度和多方向的特征表示，更适应这种复杂退化模式，从而更好地实现小目标识别<sup>[7]</sup>。其表达式为：

$$D' = \text{Attention}(N, K, E) = \text{soft max} \left( WU / \sqrt{\zeta} \right) \quad (7)$$

式中：Attention(·) 表示自注意力计算操作；N 表示网络结构中包含的注意力头数量；K 表示键矩阵中的子矩阵数量；E 表示键矩阵的维度；softmax(·) 表示输出层的激活函数；W 表示小目标特征参数；U 表示增强系数；ζ 表示方向数量；D' 表示增强后的小目标特征参数。

特征增强后，通过方向滤波器组对应的多方向特征编码器将特征切分为多个方向的子特征<sup>[8]</sup>。公式为：

$$S_{x_0} = D' + \gamma_{x_0} \sum_{\zeta=1}^M \psi \cdot P \quad (8)$$

式中：γ<sub>x<sub>0</sub></sub> 表示第 x<sub>0</sub> 个方向的 Gabor 滤波器；M 表示特征切分数量；ψ 表示将多个注意力头部的输出进行融合的权重矩阵；S<sub>x<sub>0</sub></sub> 表示第 x<sub>0</sub> 个方向的子特征。

进一步构建多尺度多方向 Transformer 编码器对不同方向的子特征进行卷积操作<sup>[9]</sup>，通过对特征进行建模以获取二维图像的特征向量。公式为：

$$\omega = \frac{1}{S_{x_0}} + l_{a_0, b_0} \oplus E_0 \quad (9)$$

式中：l<sub>a<sub>0</sub>, b<sub>0</sub></sub> 表示第 a<sub>0</sub> 个尺度、第 b<sub>0</sub> 个方向的滤波器系数；⊕ 表示卷积操作；E<sub>0</sub> 表示卷积核的权重；ω 表示特征向量。

利用图像特征向量对模型进行训练，针对遥感图像小目标识别任务，选取交叉熵损失函数作为模型的损失函数，继而输出小目标识别标签<sup>[10]</sup>。即：

$$\hat{O}_\tau = -\omega \sum_{\tau=1}^I O_\tau \log(\text{Loss}) \quad (10)$$

式中：I 表示目标类别总数；O<sub>τ</sub> 表示第 τ 种类别真实标签的独热编码；Loss 表示损失函数；Ō<sub>τ</sub> 表示模型预测的类别 τ 概率分布。通过概率分布情况确定图像小目标的真实类型。

## 2 实例论证分析

实验数据集是航空影像数据集，包含若干多尺度、低对比度目标，数据源自全色分辨率 0.8 m、多光谱分辨率 3.2 m 以及全色分辨率 0.31 m、多光谱分辨率 1.24 m 的 WorldView-3 卫星。数据集共有 15 000 幅像素为 10 000 px×10 000 px 的遥感图像，选取不同模糊场景并剔除重叠区域后，最终选用 500 幅图像用于实验，图像示例如图 2 所示。

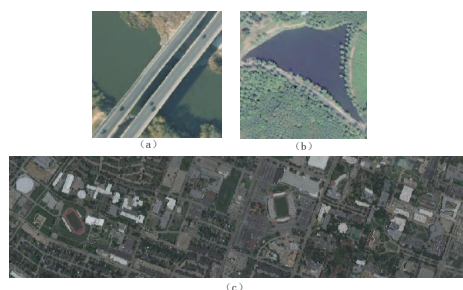


图 2 实验数据集示例

图 2 中，(a) 和 (b) 为单一尺度小目标样本，(c) 为多尺度堆叠的小目标样本，具体地物类别包括建筑、道路、小路、树木、草地和阴影 6 类，类别标记及样本数量如表 1 所示。

表 1 遥感图像类别标记及样本数量

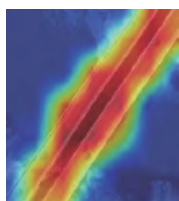
| 类别编号 | 类别名称 | 训练样本量 / 幅 | 测试样本量 / 幅 | 颜色标记 |
|------|------|-----------|-----------|------|
| 1    | 建筑   | 198       | 90        |      |
| 2    | 道路   | 220       | 89        |      |
| 3    | 小路   | 165       | 200       |      |
| 4    | 树木   | 184       | 124       |      |
| 5    | 草地   | 236       | 98        |      |
| 6    | 阴影   | 104       | 214       |      |

在服务器上配置 PyTorch 1.12.0 版本的深度学习框架，并安装 CUDA 11.6。计算软件和 cuDNN 8.4.0 编程模型，以开发和运行目标识别模型。多尺度 - 多方向 Transformer 算法的参数设置如表 2 所示。

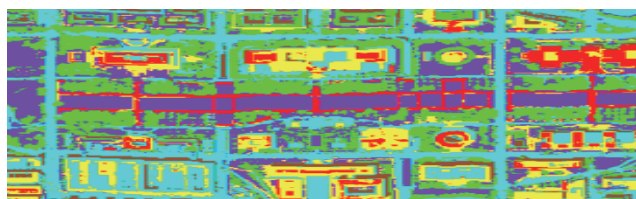
表 2 多尺度 - 多方向 Transformer 算法参数

| 参数名称      | 数值    |
|-----------|-------|
| 嵌入维度      | 256   |
| 层数        | 6     |
| 注意力头数     | 4     |
| 尺度数量      | 3     |
| 初始下采样率    | 4 倍   |
| 方向角度数     | 4     |
| 方向嵌入维度    | 32    |
| 前馈维度      | 1 024 |
| Dropout 率 | 0.1   |
| 权重衰减      | 0.1   |
| 初始学习率     | 0.2   |
| 学习率衰减因子   | 0.9   |
| 梯度裁剪阈值    | 1.0   |
| 批次大小      | 64    |

以图 2 所示 (a) 和 (c) 图像为待识别样本, 则本文方法的识别效果如图 3 所示。其热区分布呈现细长形态, 与图 2 (a) 所示的“桥梁”目标一致, 表明本文方法能够有效聚焦主体目标区域, 抑制背景干扰, 实现单一尺度目标的精准定位。图 3 (c) 对应复杂多目标场景, 可见不同类别的小目标均被有效区分, 各区域颜色标记与表 1 预设类别一致, 验证了方法在多重模糊干扰下仍具备稳定的多目标识别与分类能力。



(a)



(c)

图 3 基于本文方法的遥感小目标识别结果

与文献 [1] 的深度学习方法及文献 [2] 的卷积神经网络方法进行对比, 结果如图 4 所示。本文方法在不同特征数量下均保持 Kappa 系数在 0.9 以上。

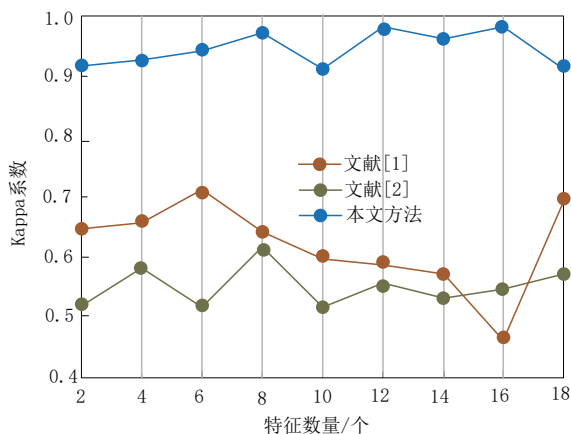


图 4 遥感图像小目标识别准确度对比

### 3 结语

针对模糊遥感图像中存在的大气散射、传感器噪声耦合干扰导致的非均匀退化、多尺度纹理交叠及边缘频域混叠等问题, 提出了一种基于多尺度 - 多方向 Transformer 的小目标识别方法。在 DOTA-F 数据集上对不同地物类别的小目标均表现出良好的识别性能, Kappa 系数稳定在 0.9 以上, 性能显著优于传统深度学习方法与卷积神经网络模型, 验证了其在复杂退化场景中具有较强的分类一致性与鲁棒性。

该研究为高分辨率遥感监测、灾害应急响应等场景下的小目标检测任务提供了有效的技术解决方案, 具有较好的工程应用前景。

### 参考文献:

- [1] 周末. 基于深度学习算法的遥感图像识别检测技术研究 [J]. 信息记录材料, 2024,25(4):162-164.
- [2] 史宝岱, 徐艳召, 崔俊杰, 等. 基于卷积神经网络的无人机图像自动识别算法 [J]. 信息工程大学学报, 2023, 24(5): 526-532.
- [3] 廖晓芳, 胡新乾. 基于改进 TSO 优化图像熵的图像识别与分割 [J]. 计算机工程与设计, 2024,45(11):3471-3478.
- [4] 杨佳丽. 基于无人机遥感影像的农业种植病虫害图像识别数智研究 [J]. 自动化与仪器仪表, 2025(7):160-163.
- [5] 郭志强, 吉顺文, 丁国章. 基于时空信息特征的大规模遥感影像目标物变化检测算法 [J]. 测绘与空间地理信息, 2025, 48(10): 119-122.
- [6] 李长冬, 龙晶晶, 刘勇, 等. 基于 EfficientNet 的滑坡遥感图像识别方法: 以贵州省毕节市为例 [J]. 华南地质, 2023, 39(3): 403-412.
- [7] 王科平, 左鑫浩, 杨艺, 等. 基于伪全局 Swin Transformer 的遥感图像识别算法 [J]. 模式识别与人工智能, 2023, 36(9): 818-831.
- [8] 何战. 宝鸡地区主干断裂的遥感图像识别方法研究 [J]. 科技与创新, 2023(1):169-172.
- [9] 谢立东, 刘静超, 文雪凤. 优化 YOLOv11 模型: 基于多尺度注意力机制的小目标检测性能提升研究 [J]. 人工智能与机器人研究, 2025, 14(1):229-236.
- [10] 王新蕾, 王硕, 翟嘉政, 等. 多任务联合学习下的复杂天气航拍图像目标检测算法 [J]. 计算机工程与应用, 2025, 61(2): 97-111.

### 【作者简介】

刘婷 (1988—), 女, 河南商丘人, 硕士, 讲师, 研究方向: 智能控制、模式识别。

(收稿日期: 2025-11-13 修回日期: 2026-03-13)