http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0420

面向低空起降场的FOD检测与识别

郑皓文[1]，胡海苗[1,2]，徐壮[3,*]，何壮[2]，胡昊鑫[2]

(1. 北京航空航天大学 计算机学院，北京 100191；2. 北京航空航天大学杭州创新研究院，杭州 310000；
3. 浙江省机场集团有限公司，杭州 311207)

摘 要：低空起降场跑道外来物(FOD)的检测与识别对于飞机飞行性能与机场运行安全具有重要意义。针对FOD目标尺寸小、无表观特征的特点及公开数据少的问题，构建FOD数据集，并提出一种基于物理特性的FOD检测与识别方法。采集常见FOD材质的多光谱图像，构建FOD的多光谱数据集。基于该数据集，对常见FOD(包括金属的锈迹与强反射现象)进行属性分析。基于沥青背景反射率，得到材质相对沥青的伪反射率。通过多种统计量，初步验证了不同材质的可分性，进而提出了使用密集嵌套连接的编码器-解码器结构神经网络，将伪反射率作为输入进行材质识别，在所构建数据集的测试集中，准确率为0.999 3，Macro-F1分数为0.774 9，优于同模型大小等级的其他材质识别和分割模型。

关 键 词：材质识别；机场跑道外来物；自编码器；多光谱图像；伪反射率
中图分类号：TP399
文献标志码：A **文章编号**：1001-5965(2026)09-3153-10

低空起降场的跑道外来物(foreign object debris, FOD)是指散落在跑道上或被风吹至跑道的非跑道本身物体，包括飞机零部件、工作人员遗落物品或跑道表面剥落的物质，如金属碎片、橡胶、碎石等[1]。这些FOD在飞机起降时可能对飞机各部位造成损害，轻则降低飞行性能，重则发生严重飞行事故。因此，增强FOD的检测与预警能力，精确识别并及时清除道面上的FOD，对控制飞机维护成本、确保飞行性能及运行安全具有重要意义。

然而，FOD类型多样、尺寸极小且无固定表观特征的特点导致FOD的检测与识别存在一定困难。不同材质的FOD空间特性区分度较低，基于单一的传感设备难以对其特征进行充分表达。目前，大部分起降场的FOD检测仍采用人工巡检，少数大型机场采用雷达技术、自动巡检机器人技术及图像处理技术。基于图像处理技术的FOD检测普遍基于可见光图像进行处理，将FOD检测转化为特定场景下的目标检测。深度学习技术为该思路提供了新的解决方案，主要集中于对两阶段网络(如更快速的区域卷积神经网络(faster region-based convolutional neural network, Faster R-CNN)[2])和单阶段网络(如YOLO系列[3])的优化与改进。这些方法通过优化网络架构、引入多尺度特征融合、结合自监督学习或异常检测策略，显著提升了FOD检测的精度和鲁棒性。两阶段目标检测网络以Faster R-CNN为代表，通过区域提议网络(region proposal network, RPN)和后续的分类与边界框回归实现目标检测。针对FOD检测的特殊需求，研究者对其进行了多项改进。例如，文献[4]提出一种专门针对FOD检测的框架，通过改进RPN以适应跑道场景中小目标的特点，并结合基于空间变换网络的卷积分类器，进一步增强了对FOD的定位和分类能

收稿日期：2025-07-01；**录用日期**：2025-08-08；**网络出版时间**：2025-08-20 17:36
网络出版地址：link.cnki.net/urlid/11.2625.V.20250820.1500.001
基金项目：浙江省“尖兵”“领雁”研发攻关计划(2025C01037)
*通信作者. E-mail：frank0139@163.com

引用格式：郑皓文，胡海苗，徐壮，等. 面向低空起降场的FOD检测与识别[J]. 北京航空航天大学学报，2026，52(9)：3153-3162.
Zheng H W，Hu H M，Xu Z，et al. FOD detection and recognition for low-altitude takeoff and landing sites[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(9)：3153-3162 (in Chinese).

力，文献 [5] 在 Faster R-CNN 的基础上引入了多尺度特征融合机制，通过融合不同层次的特征图，捕捉 FOD 在不同尺度下的视觉特征，同时，还结合了在线难样本挖掘策略，通过动态调整训练样本的权重，提升模型对复杂背景中 FOD 的检测性能。然而，这类方法依赖目标的二维表观特征，不能完全使用实际起降场 FOD 的特点。除了传统的监督学习方法，近年来，基于自监督学习和异常检测的 FOD 检测方法逐渐受到关注。这些方法利用跑道图像的内在规律，通过无监督或自监督方式学习正常跑道场景的特征，将 FOD 视为异常进行检测。文献 [6] 提出一种基于自监督学习的 FOD 检测框架，通过训练模型预测跑道图像的正常纹理和结构特征，检测偏离正常模式的区域作为潜在 FOD。在此基础上，文献 [7] 进一步提出一种基于重构自编码器的 FOD 检测方法，结合高斯混合模型对跑道图像的重构误差进行建模，将 FOD 视为异常点进行检测。这类方法适用于 FOD 的定位，但无法对 FOD 材质进行识别。

材质识别领域因深度学习技术的进步而快速发展，尤其是在大模型应用于材质识别任务后，显著提升了识别精度和跨场景泛化能力。传统方法依赖手工设计的纹理特征，而大模型通过大规模预训练，结合自注意力机制或多模态学习，提取高层次语义与低层次纹理特征，适用于复杂场景下的材质分类。文献 [8] 提出的情境材料数据集（materials in context database, MINC）为材质识别提供了基准，推动了基于卷积神经网络的研究，如文献 [9] 的深度纹理编码网络（deep texture encoding network, Deep TEN）通过纹理编码提升了性能。Vision Transformer[10] 通过自注意力机制捕捉全局与局部特征，在 MINC 和 Flickr 材料数据集（Flickr material database, FMD）上展现出优越性能。对比语言-图像预训练（contrastive language-image pre-training, CLIP）[11] 的多模态能力支持零样本材质识别，通过文本引导实现无需额外标注的分类。文献 [9] 的 DINOv2 通过自监督学习提取鲁棒特征，适合小样本场景。生成模型（如扩散模型[12]）通过材质样本生成增强了训练数据。然而，大模型的参数量庞大，导致计算资源需求高、推理时间长，限制了其在资源受限设备上的部署。

不同材质物体因分子结构差异，对不同波段电磁波的反射吸收能力各异。通过多光谱探测可获取物体的光谱曲线，进而根据材质对应的光谱特征识别其类别，适用于无表观特征的材质识别。文献 [13] 利用高光谱成像发现沥青地面在特定波段的稳定性，可作为其特征波段。文献 [14] 结合高光谱与机器学习分类器，实现无线充电装置内金属异物的检测，通过像素级分类模型识别金属光谱特性。这些研究为多光谱技术在 FOD 材质识别中的应用提供了基础。公开的 FOD 数据集较少，如机场外来物（FOD in airports, FOD-A）数据集[15] 和上海大学 FOD 数据集[16]，这些数据集均为无人机或可见光相机采集的可见光数据，FOD 尺寸较大且表观特征丰富。目前，公开的 FOD 数据集较少，多光谱信息在 FOD 检测识别领域的应用还未引起广泛关注。

基于以上分析，本文通过采集并构建多光谱 FOD 数据集，开发了一种基于光谱特性嵌入的 FOD 检测与识别方法，降低对表观信息的依赖，目标只以碎片化或零件化形式存在，仍能实现准确检测，同时，物理特性具有良好的区分性，不需要很多的训练数据，就能实现高精度的类别区分及材质识别，因此，能够很好地解决低空起降场道路异物检测困境，对于机场安全具有重要意义。

1 FOD 属性分析

1.1 FOD 数据集构建

目前，关于 FOD 的多光谱数据集较少，因此，本文使用波长范围为 385～1 009 nm 的 256 波段高光谱相机，采集并最终保留了 50 张相对高质量的数据，并使用水平翻转、垂直翻转、旋转不同角度的方式扩展为 200 张尺寸为 696 像素×775 像素、波段数为 256（即每张图像的尺寸为 696 像素×775 像素×256 像素）的数据构成数据集。数据集中包括但不局限于上述重点关注的材料；拍摄距离为 5 m、10 m、20 m 不等；拍摄环境包含向光（23 张）、背光（20 张）、阴天（7 张）时的场景。对于数据集的标注掩码，本文使用 ENVI 软件进行像素级的数据标注。

从光谱的角度，神经网络模型可以从少量的像素中充分学习不同材质的光谱特征，因此，本文数据集将多种 FOD 放在同一张图中，作为主要的训练来源。而在测试集中，在背景主要为沥青的场景中放置极少的 FOD，检验模型异常检测与材质识别的效果。图 1 展示了训练集的主要形式，其中包括了重点关注的 FOD 材料光谱特性。

FOD 的种类方面，根据文献 [1] 总结的常见 FOD 物品，从材质角度分析，重点需要关注金属（40 张）、硬质塑料（20 张）、石头（30 张）、木棍（20 张）、纸板（20 张）、树叶（30 张），一张图像一般包含多个材质。其中，除了木棍、纸板与树叶，其余材料为中危或高危材质。对于危险系数最高的金属，本文收集了锌、铝、钢等常见的金属材质进行



图 1　多光谱数据集中训练集示例(图片为使用红、绿、蓝 3 波段合成的伪 RGB 图像)

Fig. 1　Training set examples in the multispectral dataset (images are pseudo-RGB images synthesized using red, green, and blue wavelengths)

光谱数据采集, 同时, 针对金属会出现的强反射和锈迹现象, 本文数据集也专门进行了采集和单独标注, 如图 2 所示。对于强反射现象来说, 本文数据集认为在调整过曝光系数后, 像素的光谱值在 520～587 nm 之间超出相机响应值阈值(60 000), 应被视作发生了强反射; 对于锈迹, 本文数据集选用了整面全部为锈迹的材料作为样本, 便于标记的同时, 认定该区域全部为锈迹。

1.2　基于数据集的 FOD 属性分析

结合实验室以往工作中采集的多光谱数据集

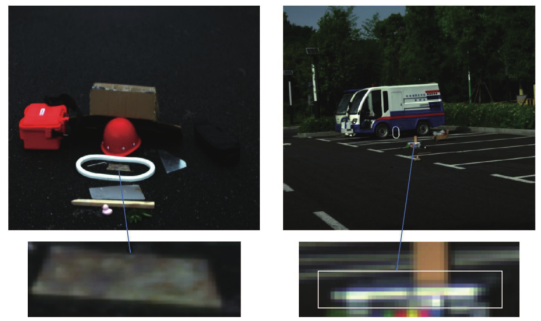


图 2　数据集中金属出现的锈迹与强反射现象

Fig. 2　Rust stains and strong reflection phenomena appearing on metals in the dataset

及本文采集得到的数据可以观察得到, 使用同一种相机拍摄得到的数据中, 沥青地面的光谱响应值趋势保持高度的稳定性及相对固定的特征, 仅在太阳光照强度不同、地面干燥程度不同时存在响应幅值的差异。这种特性使在传统的白板校正方法中使用沥青光谱信息替代白板具有合理性, 如图 3 所示。

反射率提取领域中, 经典的本征分解[17]模型为

$$\boldsymbol{I}=\boldsymbol{R}\circ\boldsymbol{S} \tag{1}$$

式中: $\boldsymbol{I}$为相机采集得到的响应值; $\boldsymbol{R}$为材质本征反射率; $\boldsymbol{S}$为光照信息; "∘" 表示哈达玛积。

带有强反射的本征分解模型为

$$\boldsymbol{I}=\boldsymbol{R}\circ\boldsymbol{S}+\boldsymbol{R}_{\mathrm{H}} \tag{2}$$

式中: $\boldsymbol{R}_{\mathrm{H}}$为强反射部分, 当$\boldsymbol{R}_{\mathrm{H}}$不为 $\boldsymbol{0}$ 时, $\boldsymbol{S}$为 $\boldsymbol{0}$。

该像素点的反射率信息无法复原, 在光谱信息中, 强反射在可见光范围内常见为过曝, 即响应值

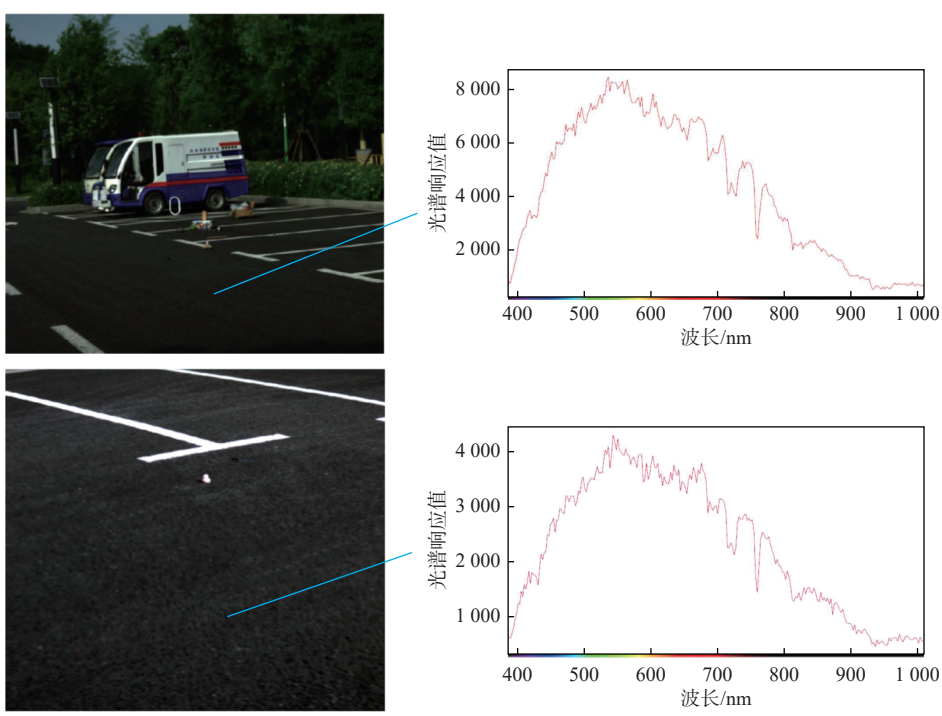


图 3　沥青地面光谱的稳定性

Fig. 3　Stability of asphalt pavement spectra

超出相机最大响应值。本文使用单独标注该种现象的方法，使强反射可以视作一种特殊的材质，进而，在后续计算反射率时，将带有强反射的模型泛化为经典的本征分解模型。进而，有

$$\boldsymbol{I}_{\mathrm{x}}=\boldsymbol{R}_{\mathrm{x}}\circ\boldsymbol{S} \tag{3}$$

$$\boldsymbol{I}_{\mathrm{y}}=\boldsymbol{R}_{\mathrm{y}}\circ\boldsymbol{S} \tag{4}$$

式中：下标 x 表示 FOD 材料；下标 y 表示沥青。

式(3)和式(4)相除，可得

$$\frac{\boldsymbol{R}_{\mathrm{x}}}{\boldsymbol{R}_{\mathrm{y}}}=\frac{\boldsymbol{I}_{\mathrm{x}}}{\boldsymbol{I}_{\mathrm{y}}} \tag{5}$$

基于$\boldsymbol{R}_{\mathrm{y}}$的稳定性，可以认为式(5)左边的值为基于沥青反射率稳定性的伪反射率。利用伪反射率具有 2 个方面的优势：①其面向当前视图，消除了光照信息对反射率的影响，具体来说，其极大地消除了向光、背光、阴天对 FOD 光谱特性的影响；②在实际应用中，可以省去传统白板校准的过程，如图 4 所示。

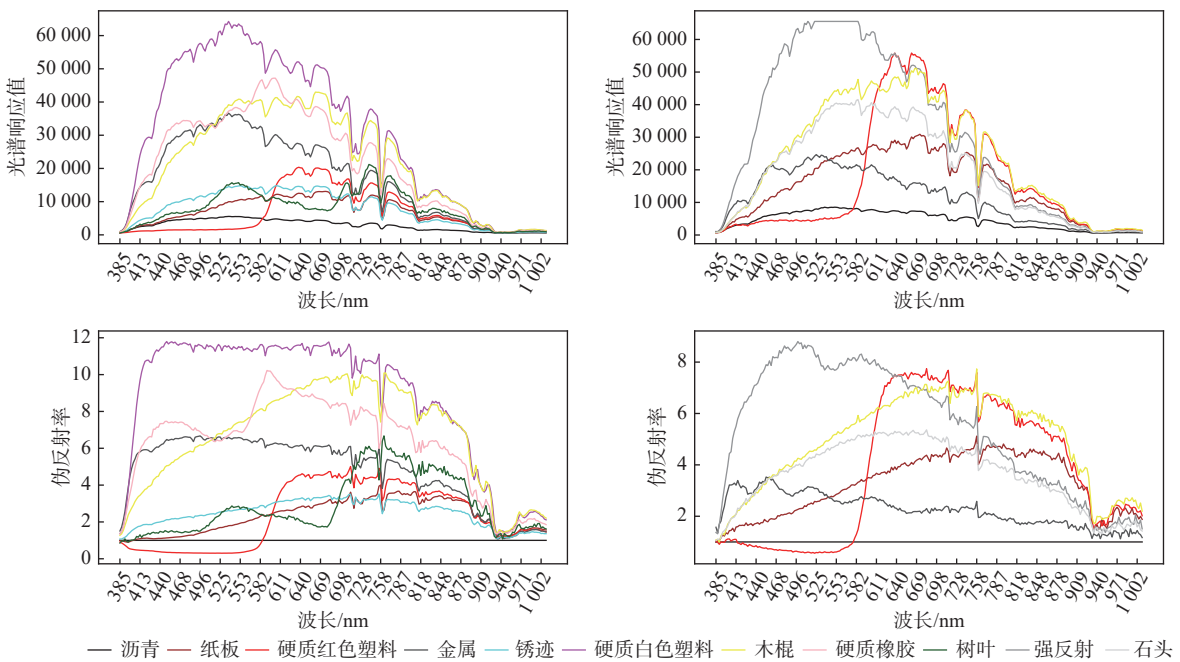


图 4 光谱响应值与伪反射率

Fig. 4 Spectral response values and pseudo-reflectance

限于实际需求与计算性能，本文从 256 波段中选择了 6 个波段用于后续检测。该 6 波段结合了手动观察特征与特征选择网络，最终确定为波长为 480, 550, 649, 740, 850, 908 nm 的 6 个波段数据，范围涵盖了可见光区域和红外区域。该 6 波段的可分性将在实验部分进行各种指标的验证。

2 密集嵌套连接的编码器-解码器 FOD 识别网络

本文的 FOD 识别主要分为 2 个部分：①将采集的光谱数据转为伪反射率；②将其输入在编码器-解码器结构基于密集连接的神经网络模型中，得到像素级的分割结果。

2.1 FOD 识别网络整体架构

2.1.1 伪反射率提取

本节主要得到沥青的反射率。基于数据集背景基本由沥青地面组成的先验，在每张图像随机采集了 50 000 个点，约占全图像总像素数的 10%。该 50 000 个点中可能存在 FOD 本身或沥青的一些偏离本身光谱特性的杂质点，因此，本文采用孤立森林[18]算法，将异常点比例设置为 5%。利用孤立森林算法无需定义相似性阈值、时间复杂度低于聚类等方法的优势，得到一个更为可靠的相似点集，如图 5 所示，其中，$\boldsymbol{R}_{\mathrm{fake}}$ 为伪反射率。

孤立森林算法通过随机分割数据构建多棵孤立树，利用异常点通常具有较短路径长度的特性进行识别。路径长度定义为

$$h(x)=e+c(m) \tag{6}$$

式中：$x\in\mathbf{R}^6$为 6 波段的数据点集；e为根节点到叶节点的边数；m为叶节点中的数据点数量；$c(m)$为调整项，近似于随机二叉搜索树的平均路径长度。

平均路径长度指包含 n 个数据点的子树预期路径长度，定义为

$$c(n)=2H(n-1)-\frac{2(n-1)}{n} \tag{7}$$

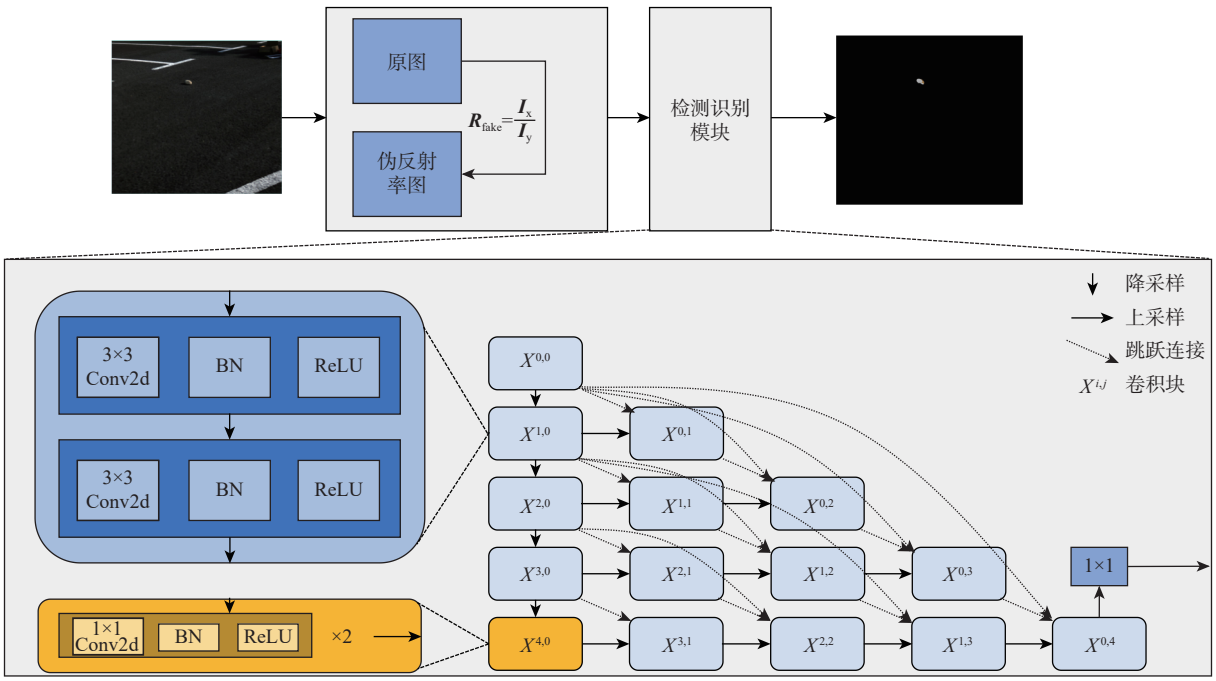


图 5　FOD 检测与识别整体结构

Fig. 5　Overall structure of FOD detection and identification

式中: $H(n-1)$ 为调和数, $H(n-1)\approx\ln(n-1)+\gamma$, γ为欧拉-马斯克罗尼常数。

数据点x的异常分数基于其在 T 棵孤立树中的平均路径长度$E[h(x)]$, 定义为

$$s(x,n)=2^{-\frac{E[h(x)]}{c(n)}} \tag{8}$$

$$E[h(x)]=\frac{1}{T}\sum_{t=1}^{T}h_t(x) \tag{9}$$

式中: 异常分数 $s(x,n)\in[0,1]$, 其越接近 1, 表示 x更可能是异常点; 本文中, T取 100。

选择得分在前 5% 的点作为异常点进行剔除, 得到的相似点集均值作为沥青的反射率。将全图响应值除以该反射率, 即可得到伪反射率图。

2.1.2　FOD 识别网络细节

FOD 缺乏表观特征, 有效信息只有相对固定的光谱响应值; 此外, 数据量相对于计算机视觉领域的常见数据集而言极少。因此, 参考文献 [19] 的思路, 使用密集嵌套连接的编码器-解码器结构, 从而充分利用每一层的形谱信息, 对于 FOD 检测与识别的精度具有较好效应。

本文网络结构由编码器、解码器、跳跃连接、瓶颈层和输出层组成。本文对数据集进行编码部分由 4 层卷积块组成, 每个下采样块包括 2 组 3×3 卷积层、标准化层和激活函数。每层卷积块之间用 2×2 最大池化层进行下采样。图 5 中, 编码路径节点标注为$X^{i,0}$, i为下采样层级。瓶颈层位于网络底部, 标记为$X^{4,0}$, 其包含 2 组 1×1 卷积层、标准化层和激活函数。解码器包含转置卷积上采样层和卷积层。解码路径节点标注为$X^{i,j}$, i为上采样层级, j为该层级的嵌套节点标号。嵌套跳跃连接部分不局限于传统 U-Net[20], 只对对称的编码、解码层进行跳跃连接, 该部分公式表达为

$$X^{i,j}=\mathrm{Conv}(\oplus(\mathrm{Up}(X^{i+1,j-1}),X^{i,0},X^{i,1},\cdots,X^{i,j-1})) \tag{10}$$

式中: Conv为卷积操作; “⊕” 表示拼接; Up为上采样操作。

每个解码器节点$X^{i,j}$接收上采样特征$X^{i+1,j-1}$和跳接特征。跳接特征包括编码器节点$X^{i,0}$和同一层级的较低节点$X^{i,k}(k<j)$。

输出层对$X^{0,4}$使用 1×1 卷积层, 生成类别数通道的分割掩码。该结构通过多层次的特征融合, 更好地利用各通道的光谱信息, 提升对不同光谱特征的区分能力; 其通过嵌套和密集跳跃连接, 将不同感受野的特征充分融合, 使模型既能捕捉全局结构, 又能保留细节和边界信息; 跳跃连接使浅层和深层信息在网络中高效流动, 提升对目标可能存在的形状、纹理、边界信息的感知能力, 从而实现形谱信息融合, 提高识别的准确性。

2.2　损失函数

在损失函数的使用上, 将 Tversky 指数[21] 与 Focal Loss[22] 以相乘的形式结合, 计算如下:

$$L=\sum_{c'=1}^{C'}\left[(1-\tau_{c'})^{\gamma'}\left(-\alpha_{c'}\sum_{i\in\Omega}y_{i,c'}\ln p_{i,c'}\right)\right] \tag{11}$$

式中: 多分类 Tversky 指数τ为每个类别 c'进行单独

计算：

$$\tau_{c'}=\frac{\sum_{l} y_{l,c'}p_{l,c'}}{\sum_{l} y_{l,c'}p_{l,c'}+\alpha_{c'}\sum_{l}\left(1-y_{l,c'}\right)p_{l,c'}+\beta_{c'}\sum_{l} y_{l,c'}\left(1-p_{l,c'}\right)} \tag{12}$$

其中：C'为类别总数；$y_{l,c'}\in\{0,1\}$为第i个像素在类别c'的真实标签，为独热编码；$p_{l,c'}\in[0,1]$为模型预测第l个像素属于类别c'的概率；Ω为所有像素集合，即全图；$\alpha_{c'},\beta_{c'}\in[0,1]$为类别$c'$的权重，分别控制假阳性和假阴性，本文中分别设置为0.7、0.3；$\gamma'\geqslant 0$为聚焦参数，控制对难分割区域的关注程度；$\sum_{l} y_{l,c'}p_{l,c'}$为类别$c'$的真阳性近似，即预测概率和真实标签的乘积和；$\sum_{l}\left(1-y_{l,c'}\right)p_{l,c'}$为类别$c'$的假阴性近似；$\tau_{c'}$为类别$c'$的Tversky指数；$(1-\tau_{c'})^{\gamma'}$为类别$c'$的区域重叠误差聚焦因子。

聚焦因子、多种权重的加入使模型更加关注小目标，从而降低漏检率。聚焦因子使用乘法的形式，更加强调模型对难分割区域的关注。其相较于Focal Loss的问题是通过优化全局区域，增强模型对形状和连通性的认知，并通过平衡局部和全局优化，生成更加平滑、连贯的分割掩码，这使模型在实际判断过程中，分割区域比真值面积大。但由于该问题在危害程度上弱于漏检，其仍具有应用价值。

3 方法验证

3.1 FOD材质可分性的初步验证

本文通过多个指标（马氏距离和离散度）对所选择的6个波段FOD数据可分性进行衡量。此外，利用多元方差分析(multivariate analysis of variance, MANOVA)的角度验证所选波段的显著差异。进一步，利用*t*-分布随机近邻嵌入降维方法将本文所选的谱段组合进行数据降维并投影，得到更加直观的可区分性结果。

马氏距离不同于常用的欧氏距离和曼哈顿距离，通过协方差矩阵消除了特征之间的相关性影响，在观察不同材质之间的差异性上更优；离散度作为反映2个类别之间可分性的统计距离指标，能够在2个类别均值相等或相近时有效区分2个类别。

MANOVA作为一种统计方法，用于分析多个因变量如何受到一个或多个自变量的影响。在本文中，不同波段的响应值视为不同的因变量，材质视为自变量，MANOVA通过检验因变量的联合分布，评估材质对于多个波段响应值的总体效应是否存在显著差异。表1展示了马氏距离与离散度数据，同时展示了MANOVA分析的结果，对FOD材质的可分性进行了初步验证。MANOVA的4个统计量的p值均约为0，体现了不同材质在这6个波段下的显著差异。

表1 多种衡量尺度下FOD数据的可分性验证

Table 1 Verification of FOD data separability under multiple measurement scales

指标类别	指标	p值
衡量尺度	马氏距离	4.950
	离散度	4.983
MANOVA统计量	Wilks' Lambda	0
	Pillai's Trace	0
	Hotelling's Trace	0
	Roy's Largest Root	0

对所选波段进行像素点采样，并进行*t*-分布随机近邻嵌入降维投影，得到结果如图6所示。

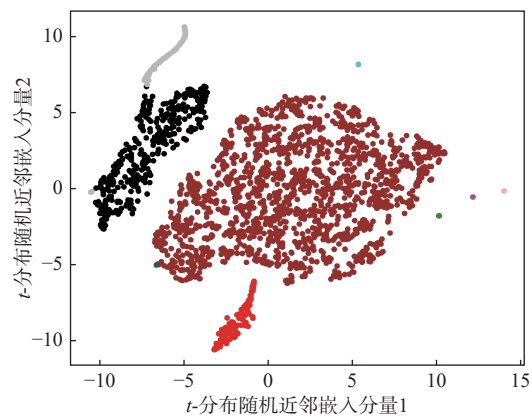


图6 *t*-分布随机近邻嵌入可视化结果

Fig. 6 *t*-distributed stochastic neighbor embedding visualization results

3.2 网络训练环境与参数设置

本文实验运行于服务器环境。硬件方面，实验使用Nvidia Tesla V100显卡，CPU为Intel Xeon Gold 5218R @ 2.10 Hz。软件方面，实验以Python为编程语言，以Pytorch为神经网络模型的框架工具。

参数设置方面，在模型训练过程中，使用Adam优化器，并设置初始学习率为0.000 3，训练10个批次。在训练过程中，根据验证集的损失对学习率进行动态调整。由于不同材质在选取波段上具有显著可分性，因此，10个批次即可实现快速收敛。图像作为输入时，将图像进行零填充，使图像长与宽可以被64整除，将图像分割为64像素×64像素大小的块后，才进入神经网络。

3.3 对比方法与评价手段

本文将所采用的神经网络模型与其他图像分割模型一同在本文数据集的测试集中进行测试，通

过对比指标的手法进行模型识别能力的评价。测试指标为准确率和 Macro-F1 分数。

准确率 A 指预测正确的样本数与总样本数的比值, 是对分类器整体正确率的一种评价, 计算公式为

$$A=\frac{n_{\mathrm{TP}}+n_{\mathrm{TN}}}{n_{\mathrm{TP}}+n_{\mathrm{FP}}+n_{\mathrm{TN}}+n_{\mathrm{FN}}} \tag{13}$$

式中: n_{TP} 为预测正确的正样本数; n_{FP} 为预测错误的正样本数; n_{TN} 为预测正确的负样本数; n_{FN} 为预测错误的负样本数。

Macro-F1 分数 F_{Macro} 相较于准确率和 F_1 分数, 关注不同类别的分布情况, 尤其适合评价少数类的性能评价, 避免大类别主导结果。本文中, 大类别为沥青背景, FOD 均为少数类。Macro-F1 分数位于 0～1 之间, 1 表示完美分类, 0 表示完全错误, 计算公式为

$$F_{\mathrm{Macro}}=\frac{1}{C'}\sum F_q \tag{14}$$

式中: F_q 为类别 q 的 F_1 分数, F_1 分数基于精确率与召回率得到, 取值在 0～1 之间, 1 表示完美分类。

精确率 P_{re} 和召回率 R_{ec} 的计算公式分别为

$$P_{\mathrm{re}}=\frac{n_{\mathrm{TP}}}{n_{\mathrm{TP}}+n_{\mathrm{FP}}} \tag{15}$$

$$R_{\mathrm{ec}}=\frac{n_{\mathrm{TP}}}{n_{\mathrm{TP}}+n_{\mathrm{FN}}} \tag{16}$$

此外, 模型大小也是本文基于实际需求进行参考的重要指标。

本文在数据集的类别标注中, 将锈迹、金属与强反射分为 3 类, 但在实际危险评级中, 这 3 类都按照金属来判断, 因此, 本文在测试结果中, 着重观察了这 3 类错检的类别情况, 如果错检结果为另 2 类, 并不影响实际判断结果。

3.4　对比实验

表 2 展示了其他较为先进的图像分割方法与本文方法在本文数据集上的表现, 图 7 为分割结果。所有方法的整体准确率高, 是因为数据集标记了相当一部分的沥青地面, 该类数据的数据量较大, 且普遍易被正确识别。FCN-8s[23] 是经典全卷积分割网络, 融合多层特征, 输出分辨率较高, 适合基础语义分割任务, 作为对比的基准; MST-DeepLabv3+[24] 针对高分辨率遥感图像分割, 加入 SENet[25] 注意力机制和迁移学习, 但模型较大, 且在此场景中误检率偏高; SR-SegNet[26] 模型较小、适合实时分割, 但由于 FOD 场景缺乏表观信息, SR-SegNet 的简单结构难以分析不同材质之间的信息。U-Net 是本文的基础架构, 是编码器-解码器结构与同层跳跃连接的结合, 但对各层信息提取仍不够充分, 在准确率上不如本文方法。MSNet[27] 采用 ResNet-50 提取特征, 通过级联上采样和特征金字塔结构融合多尺度特征, 在本文数据集中的金属和强反射识别能力较弱。SSTN[28] 结合空间注意力和光谱关联模块, 捕捉高光谱图像的长距离依赖和光谱-空间相关性, 但由于 FOD 的空间特征弱, SSTN 的效果并不突出。UNetFormer[29] 结合 U-Net 和 Transformer 架构, 效果较好, 但硬质塑料的识别准度相对较低, 且对强反射特性易识别错, 没有较好地捕捉金属的特殊特性。

表 2　各方法的测试结果与模型大小

Table 2　Test results and model sizes for each method

方法	准确率	Macro-F1	模型大小/KB
FCN-8s[23]	0.787 7	0.199 2	57 877.19
MST-DeepLabv3+[24]	0.998 9	0.545 3	104 620.54
SR-SegNet[26]	0.999 0	0.511 1	3 090.85
U-Net[20]	0.998 9	0.642 0	30 391.00
MSNet[27]	0.998 9	0.694 3	141 193.00
SSTN[28]	0.999 3	0.768 8	26 776.00
UNetFormer[29]	0.999 1	0.774 5	28 980.00
本文	0.999 3	0.774 9	23 151.00

3.5　消融分析

在消融实验部分, 本文着重验证 6 波段相较于 3 波段(RGB)FOD 区分的必要性、在损失函数中引入 Tversky 指数的必要性及网络结构模块的合理性。

3.5.1　不同波段数下 FOD 的易分性

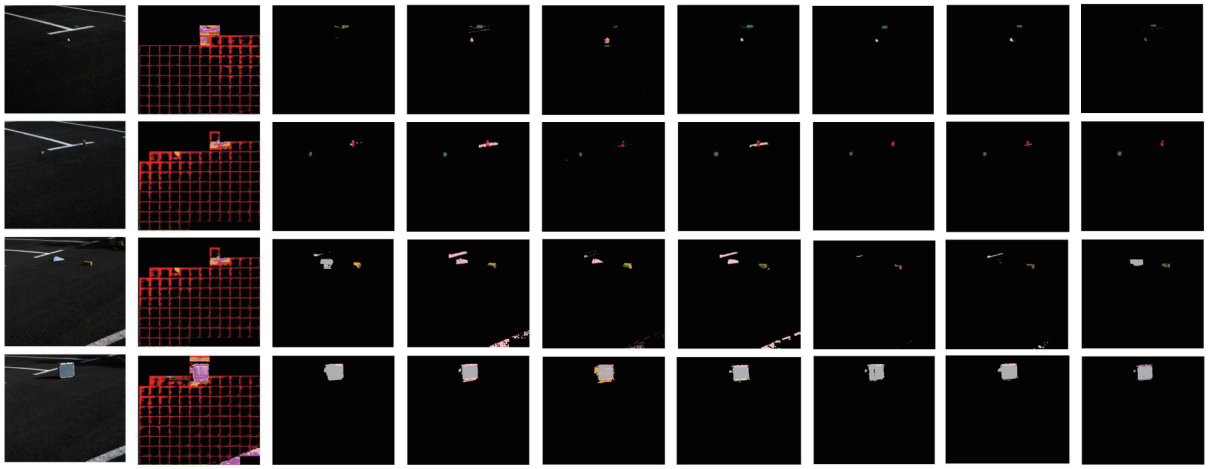
为验证多波段进行 FOD 检测与识别的重要性, 将原数据集中选择 470 nm、550 nm、634 nm 这 3 个波段(分别对应蓝色、绿色、红色)的数据生成 RGB 数据集进行实验, 结果如表 3 所示。RGB 的图像识别强依赖于形状与纹理信息, 而 FOD 的应用场景中缺乏该类信息, 因此, 在光谱维度增加信息量具有显著的优势。

3.5.2　损失函数消融

Focal Loss 是多分类问题域中样本分布不均时使用的主要损失函数。本文通过对比单一 Focal Loss 与本文采用的损失函数下训练模型得到的指标结果, 验证 Tversky 指数的必要性, 如表 4 所示。其中, Focal Loss 的定义为

$$\mathrm{FL}(p)=-\frac{1}{N}\sum_{q=1}^{C'}\alpha_q\left(1-p_q\right)^{\gamma}\ln p_q \tag{17}$$

式中: N 为批量大小; p 为概率; p_q 为模型预测样本属于第 q 类的概率; $\left(1-p_q\right)^{\gamma}$ 为聚焦因子, 降低易分类样本的损失贡献; $-\ln p_q$ 为标准交叉熵损失; α_q 为类别权重, 针对本文数据集关键材质数量远小于沥



(a) 原图 (b) FCN-8s[23] (c) MST-DeepLabv3+[24] (d) SR-SegNet[26] (e) U-Net[20] (f) MSNet[27] (g) SSTN[28] (h) UNetFormer[29] (i) 本文方法

图 7 各方法的分割结果

Fig. 7 Segmentation results for each method

表 3 6 波段与 RGB 下模型的测试指标值

Table 3 Test values for 6-band and RGB sub-models

波段数	准确率	Macro-F1
3(RGB)	0.787 7	0.488 4
6	0.999 3	0.774 9

表 4 Tversky 指数加入与否的测试指标值

Table 4 Test values with or without Tversky index

损失函数	准确率	Macro-F1
Focal Loss	0.999 1	0.616 0
本文	0.999 3	0.774 9

青等背景材质的情况，将α_q设置为该类在训练集中的像素数的倒数，从而减少数据量不平衡带来的问题。

Focal Loss 虽然缓解了类别分布不均衡的问题，但无法针对假阳性或假阴性的误分类情况进行调整，导致容易出现漏检问题，其对于 FOD 的检测是致命的。加入 Tversky 指数，通过增加对漏检的惩罚，提高 FOD 各种类的召回率，使模型在 FOD 的检测上更加鲁棒。

3.5.3 网络结构

本文采用的密集连接编码器-解码器结构与 U-Net、UNet++类似。当本文结构去除密集嵌套的跳跃连接结构时，网络结构退化为 U-Net；当本文瓶颈层的卷积核与编码路径相同(都为 3×3)时，网络结构进化为 UNet++。UNet++的模型大小达到 35 347 KB，是本文方法的近 150%。原模型在瓶颈层进行更细致的特征提取，对于 Macro-F1 分数的提升具有重要作用，具体体现在同危险等级下不同材质的识别能力更强。本文方法与 UNet++的区别在于瓶颈层的卷积核大小，这样的改动极大降低了模型大小，但没有明显影响不同危险等级下材质的识别能力，在检验危险等级的效果上也没有特别显著的下降，对于本文来说是可接受的，不同 U-Net 结构的测试结果与模型大小如表 5 所示。

表 5 不同 U-Net 结构的测试结果与模型大小

Table 5 Test results and model sizes for different U-Net structures

方法	准确率	Macro-F1	模型大小/KB
U-Net[20]	0.998 9	0.642 0	30 391
UNet++	0.999 2	0.870 6	35 347
本文	0.999 3	0.774 9	23 151

经过对误检像素点的数据统计，观察到本文方法在锈迹、金属和强反射上的误检除了部分被认为是硬质白色塑料，大部分被误检为这 3 类中的另外 2 类。而金属和硬质塑料在危险等级同属最高危物品，因此，这几类的误检是可容忍的。与此同时，本文方法对其他几类材质的检测准确率与 UNet++差距不大。

3.6 恶劣场景

在 3.4 节中的分割结果图中，本文方法将阴影部分误检为树叶。虽然常见的 FOD 材质没有出现严重的“异物同谱”现象，但阴影等现象会与部分 FOD 产生该种情况。轮胎碎片作为一种 FOD 材料，主要构成为深色橡胶，其与沥青或水泥地面会出现较稳定的“异物同谱”现象。在进一步的优化与完善中，将考虑温度等信息，缓解“异物同谱”带来的误检问题。

此外，伪反射率方法仅限于背景地面基本为沥青和水泥等单一材质的情况。如果实际应用过程中的背景地面杂质较多，该方法大概率将失效，届时，仍需要使用经典的白板校正方法获得像素点的反射率信息。

4 结　论

1) 本文收集常见 FOD, 拍摄并构建了 FOD 的高光谱数据集。

2) 从目标本征特性出发, 通过采集并构建常见 FOD 数据集, 构建了针对 FOD 的目标材质属性建模方法。

3) 针对 FOD 缺乏表观特征的问题, 设计了基于编码器-解码器结构的密集嵌套连接的神经网络, 充分利用像素点的形信息与谱信息, 提高了 FOD 材质识别的准确度, 对于机场安全具有重要意义。

参考文献（References）

[1] Federal Aviation Administration. Airport foreign object debris detection equipment: AC 150/5220-24—2009[S]. Washington, D. C. : USA Department of Transportation, 2010.

[2] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[3] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.

[4] Cao X G, Wang P, Meng C, et al. Region based CNN for foreign object debris detection on airfield pavement[J]. Sensors, 2018, 18(3): 737.

[5] 韩松臣, 张比浩, 李炜, 等. 基于改进 Faster-RCNN 的机场场面小目标物体检测算法[J]. 南京航空航天大学学报, 2019, 51(6): 735-741.
Han S C, Zhang B H, Li W, et al. Small target detection in airport scene via modified Faster-RCNN[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2019, 51(6): 735-741(in Chinese).

[6] Munyer T, Brinkman D, Zhong X, et al. Foreign object debris detection for airport pavement images based on self-supervised localization and vision Transformer[C]//Proceedings of the 2022 International Conference on Computational Science and Computational Intelligence. Piscataway: IEEE Press, 2023: 1388-1394.

[7] 荆颖, 郑红, 郑文韬. 基于高斯混合特征改进的自编码器机场跑道 FOD 精准检测[J]. 计算机辅助设计与图形学学报, 2025, 37(9): 1593-1606.
Jing Y, Zheng H, Zheng W T. Gaussian mixture feature improved autoencoder for FOD detection on airport runway[J]. Journal of Computer-Aided Design & Computer Graphics, 2025, 37(9): 1593-1606(in Chinese).

[8] Bell S, Upchurch P, Snavely N, et al. Material recognition in the wild with the materials in context database[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 3479-3487.

[9] Zhang H, Xue J, Dana K. Deep TEN: texture encoding network[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 2896-2905.

[10] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: Transformers for image recognition at scale [EB/OL]. (2021-06-03)[2025-06-20]. https://arxiv.org/abs/2010.11929.

[11] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision[EB/OL]. (2021-02-26)[2025-06-20]. https://arxiv.org/abs/2103.00020.

[12] Dhariwal P, Nichol A. Diffusion models beat GANs on image synthesis[EB/OL]. (2021-06-01)[2025-06-20]. http://arxiv.org/abs/2112.10752.

[13] Taupik J, Alamsyah T, Wulandari A, et al. Airport runway foreign object debris (FOD) detection based on YOLOX architecture [C]//Proceedings of the 2023 International Conference on Computer Science, Information Technology and Engineering. Piscataway: IEEE Press, 2023: 40-43.

[14] Tian Y, Li Z, Lin Y W, et al. Metal object detection for electric vehicle inductive power transfer systems based on hyperspectral imaging[J]. Measurement, 2021, 168: 108493.

[15] Munyer T, Huang P C, Huang C Y, et al. FOD-A: a dataset for foreign object debris in airports[EB/OL]. (2022-01-26)[2025-06-20]. https://arxiv.org/abs/2110.03072.

[16] Xu H Y, Han Z Q, Feng S L, et al. Foreign object debris material recognition based on convolutional neural networks[J]. EURASIP Journal on Image and Video Processing, 2018, 2018(1): 21.

[17] Barrow H G, Tenenbaum J M. Recovering intrinsic scene characteristics from images[J]. Computer Vision Systems, 1978, 2: 3-26.

[18] Liu F T, Ting K M, Zhou Z H. Isolation forest[C]//Proceedings of the 2008 Eighth IEEE International Conference on Data Mining. Piscataway: IEEE Press, 2009: 413-422.

[19] Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N, et al. UNet++: a nested U-Net architecture for medical image segmentation[C]//Proceedings of the Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Berlin: Springer, 2018: 3-11.

[20] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Proceedings of the Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. Berlin: Springer, 2015: 234-241.

[21] Salehi S S M, Erdogmus D, Gholipour A. Tversky loss function for image segmentation using 3D fully convolutional deep networks [C]//Proceedings of the Machine Learning in Medical Imaging. Berlin: Springer, 2017: 379-387.

[22] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2999-3007.

[23] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 3431-3440.

[24] Wang Y, Yang L, Liu X Z, et al. An improved semantic segmentation algorithm for high-resolution remote sensing images based on DeepLabv3+[J]. Scientific Reports, 2024, 14: 9716.

[25] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision

and Pattern Recognition. Piscataway: IEEE Press, 2018: 7132-7141.

[26] Weng W, Zhuo Z, Li H, et al. Semantic segmentation of remote sensing images based on dilated convolution and SegNet[J]. Journal of Applied Remote Sensing, 2020, 14(4): 046511.

[27] Tao C X, Meng Y Z, Li J J, et al. MSNet: multispectral semantic segmentation network for remote sensing images[J]. GIScience & Remote Sensing, 2022, 59(1): 1177-1198.

[28] Zhong Y F, Zhang L P, Huang B, et al. A spectral-spatial Transformer network for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1-14 .

[29] Wang L B, Li R, Zhang C, et al. UNetFormer: a UNet-like Transformer for efficient semantic segmentation of remote sensing urban scene imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 190: 196-214.

FOD detection and recognition for low-altitude takeoff and landing sites

ZHENG Haowen[1], HU Haimiao[1, 2], XU Zhuang[3, *], HE Zhuang[2], HU Haoxin[2]

(1. School of Computer Science and Engineering, Beihang University, Beijing 100191, China;
2. Hangzhou Innovation Institute of Beihang University, Hangzhou 310000, China;
3. Zhejiang Provincial Airport Group Co., Ltd., Hangzhou 311207, China)

Abstract: The detection and recognition of foreign object debris (FOD) on low-altitude takeoff and landing runways are of critical importance to aircraft performance and airport operational safety. A FOD dataset was created and a physics-based FOD detection and identification approach was suggested in order to address the issues caused by the small size, lack of distinctive visual traits, and restricted availability of public data for FOD. Initially, multispectral images of common FOD materials were collected to establish a multispectral FOD dataset. Based on this dataset, attribute analysis was conducted on common FOD types, including metal rust and strong reflection phenomena. Utilizing the reflectivity of asphalt backgrounds, a pseudo-reflectivity metric relative to asphalt was derived for different materials. Through various statistical measures, the separability of different materials was preliminarily validated. Subsequently, a neural network with an encoder-decoder architecture featuring densely nested connections was proposed, using pseudo-reflectivity as input for material identification. The suggested approach outperformed existing material identification and segmentation models with similar parameter ranges on the test set of this dataset, achieving an accuracy of 0.999 3 and a Macro-F1 score of 0.774 9.

Keywords: material recognition; airport runway foreign object debris; autoencoder; multispectral image; pseudo-reflectivity

Received: 2025-07-01; **Accepted**: 2025-08-08; **Published Online**: 2025-08-20 17:36

URL: link.cnki.net/urlid/11.2625.V.20250820.1500.001

Foundation item: Zhejiang Provincial "Pioneer" and "Leading Goose" R&D Program (2025C01037)

* **Corresponding author**. E-mail: frank0139@163.com