



文献 CSTR:
32001.14.11-6035.csd.2025.0031.zh



文献 DOI:
10.11922/11-6035.csd.2025.0031.zh
数据 DOI:
10.57760/sciencedb.j00001.01276

文献分类: 地球科学

收稿日期: 2025-03-15
录用日期: 2025-10-24
发表日期: 2025-12-22

可见光-SAR 遥感多模态机场检测数据集

赵旭阳^{1,2}, 赵理君^{1*}, 李宏益¹, 张正¹, 唐婷¹

1. 中国科学院空天信息创新研究院, 北京 100094
2. 中国科学院大学电子电气与通信工程学院, 北京 100049

摘要: 遥感影像目标检测是地球观测任务中的重要领域, 通过融合来自不同传感器的遥感数据, 利用多模态互补信息, 可以提高目标检测模型的性能。针对当前遥感多模态目标检测领域存在的数据集数量不足、模态组合欠缺、目标类型有限的问题, 本研究基于 Google Earth 可见光影像以及高分三号 (Gaofen-3, GF-3) 合成孔径雷达 (synthetic aperture radar, SAR) 影像, 构建了一个聚焦机场的可见光-SAR 遥感多模态目标检测数据集 MVSDA。在数据集构建过程中, 遵循规范的数据收集、处理、标注方法, 涵盖数据下载、空间对齐、无值区域处理、样本标注、质量控制与评价等一系列步骤。本数据集具备像素级对齐、空间跨度大、尺度跨度大、场景多样性丰富等特点, 可满足主流深度学习模型训练需求。本数据集能够为相关领域的科研人员提供重要的基础数据资源, 对推动遥感多模态目标检测领域研究具有实际应用价值。

关键词: 多模态; 目标检测; 遥感; 机场; 可见光; SAR

数据库 (集) 基本信息简介

数据库 (集) 名称	可见光-SAR 遥感多模态机场检测数据集
数据通信作者	赵理君 (zhaolj201934@aircas.ac.cn)
数据作者	赵旭阳, 赵理君, 李宏益, 张正, 唐婷
数据时间范围	2018–2020年
地理区域	全国34个省级行政区
空间分辨率	8–69 m
数据量	15.18 GB
数据格式	*.png, *.txt, *.xml, *.zip
数据服务系统网址	https://doi.org/10.57760/sciencedb.j00001.01276
基金项目	民用航天技术预先研究项目 (D040404)
数据库 (集) 组成	数据集主要由影像数据、标注文件数据两部分组成, 共包含4个zip压缩文件, 其中: (1) Annotations_VOC.zip是旋转框VOC格式的标注文件数据压缩包, 数据量为1915, 以xml格式存储; (2) Annotations_YOLO.zip是水平框YOLO格式的标注文件数据压缩包, 数据量为1915, 以txt格式存储; (3) Visible.zip是可见光模态的遥感影像数据压缩包, 数据量为1915, 以png格式存储; (4) SAR.zip是SAR模态的遥感影像数据压缩包, 数据量为1915, 以png格式存储。

* 论文通信作者
赵理君: zhaolj201934@aircas.ac.cn

引言

遥感影像目标检测是目前遥感地球观测领域中一个备受关注的研究方向，在环境监测、精准农业和城市规划等应用领域发挥着重要作用^[1]。在遥感技术领域中，多模态数据指通过不同传感器或不同成像方式获取的关于同一场景和目标的多种成像结果。基于单模态数据的检测技术无法充分利用多模态数据之间由于异构特性而产生的互补和关联信息，导致其智能解译能力不足，应用场景也较为狭窄^[2]。不同类型的遥感数据由于成像机理的差异，各自具备独特的特性与优势，通常能够提供更丰富的地物联合表达信息^[3]。例如，可见光影像能够呈现精细的颜色和纹理细节；合成孔径雷达（synthetic aperture radar, SAR）影像凭借其主动成像机制，在获取地物几何结构和辐射特性方面不易受到气候、云雾等外部因素干扰；红外影像则反映了地物目标的热辐射特性，具有昼夜成像能力强、抗干扰性高等优势；多光谱和高光谱影像则能够提供丰富的光谱特征信息，可区分地物之间细微的光谱差异。这些信息的互补性使得遥感多模态目标检测在复杂场景中展现出巨大的应用潜力。

通过融合来自不同传感器的遥感多模态数据，研究人员能够深入挖掘并充分利用多源数据的互补信息，从而提升目标检测模型的准确性与鲁棒性^[4]。近年来，得益于深度学习算法强大的特征提取能力以及一些公开可用的遥感影像多模态数据集，许多用于处理多源遥感数据的实时目标检测方法应运而生，例如 YOLOrs^[5]、SuperYOLO^[6]、BFMYOLO^[7]、BFDN_YOLOv8^[8]等。

然而，当前的遥感多模态目标检测数据集存在显著不足，制约了该领域的进一步发展。首先，现有遥感多模态目标检测数据集数量有限，目前可公开下载的数据集有多模态航空影像车辆检测数据集 VEDAI（Vehicle Detection in Aerial Imagery）^[9]、无人机海洋环境人船目标检测与跟踪数据集 SeaDrones^[10]和无人机跨模态车辆检测数据集 DroneVehicle^[11]。数据集的稀缺性极大地限制了研究者的选择和应用范围。尤其是在针对特定应用场景时，可用的数据集资源更是稀缺。其次，目前广泛使用的遥感多模态目标检测数据集大多局限于可见光与近红外两种模态，对于其他重要的模态组合方式的数据集鲜有报道，如可见光与 SAR 的结合等。这种单一的模态组合极大地限制了遥感多模态目标检测技术在多样化环境条件下的应用潜力。同时，现有的遥感多模态目标检测数据集在样本种类和场景多样性方面也存在较大局限，难以全面覆盖各种复杂场景和目标特征。这些限制对基于深度学习的遥感多模态目标检测算法的开发和研究产生了不利影响。表 1 给出了遥感领域单模态目标检测数据集 DOTA^[12]、多模态目标检测数据集与本研究所构建数据集 MVSDA 的对比情况。

表 1 遥感领域目标检测数据集对比

Table 1 Comparative overview of remote sensing object detection datasets

名称	影像来源	图像（对）数量	类别数量	模态	标注形式
DOTA	多遥感平台	2806	15	可见光	旋转框
SeaDrones	无人机	432	2	可见光+红外	水平框
VEDAI	卫星	1248	9	可见光+红外	水平框
DroneVehicle	无人机	28439	5	可见光+红外	旋转框
MVSDA	卫星	1915	1	可见光+SAR	水平框+旋转框

针对上述挑战，本研究结合了 Google Earth 可见光影像和国产高分三号（Gaofen-3, GF-3）SAR 影像，构建了一个面向机场类型、结合可见光与 SAR 模态数据的遥感多模态目标检测数据集 MVSDA

(Multimodal Visible-SAR Dataset for Airport detection)。本数据集涵盖了多样化的地理场景，可为复杂条件下遥感异源数据融合等关键技术突破提供数据支撑，推动遥感多模态目标检测相关算法的发展。

1 数据采集和处理方法

本研究中的数据采集和处理方法流程如图 1 所示，包括数据收集、数据处理、数据标注三部分，具体步骤包括数据下载、空间对齐、数据清洗、无值区域处理、人工标注、质量控制与评价等。

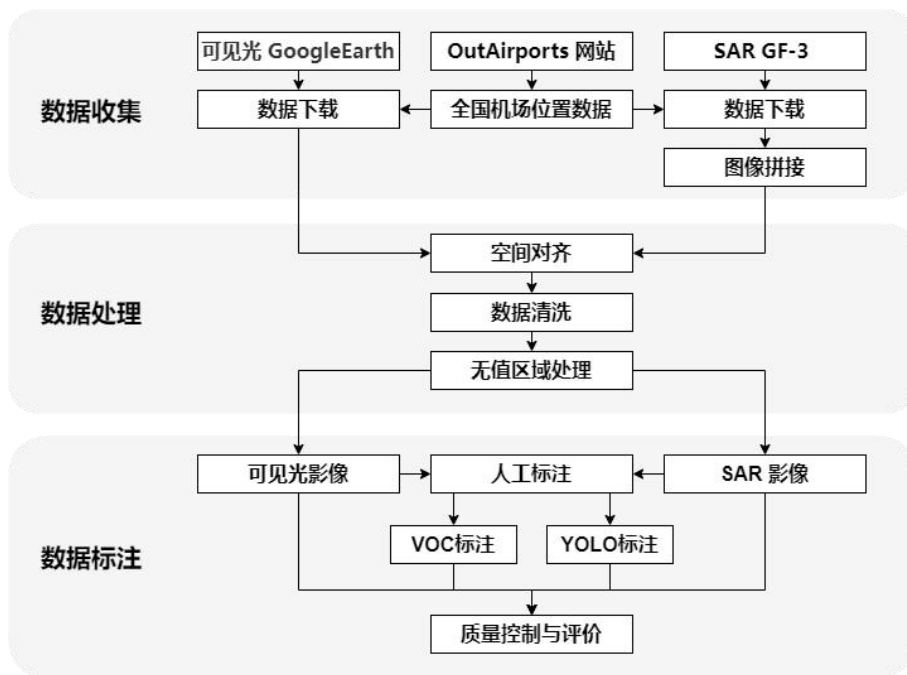


图 1 数据采集与处理流程

Figure 1 Main workflow of data acquisition and processing

1.1 数据收集

本数据集首先通过 OurAirports 网站 (<https://ourairports.com/>) 下载、整理了全国（含港澳台地区）的机场位置数据。在处理机场位置数据时，对收集到的相关数据进行汇总，并剔除其中的无效位置和无关信息。之后，以机场点位为中心，划定合适的空间范围进行影像数据收集与下载。

数据集包含可见光和 SAR 两种模态的遥感影像数据。可见光影像数据源自 Google Earth 卫星影像，选取 L12、L13、L14、L15 四个层级的产品。这些影像包含 R、G、B 三波段，其空间分辨率分别为 69 米、35 米、17 米、8 米。数据通过 91 卫图助手 (<https://www.91weitu.com/>) 等渠道下载，原始数据已经完成了辐射校正、几何纠正、影像镶嵌与拼接以及去云和阴影等预处理步骤。

SAR 影像数据主要使用了国产 GF-3 雷达遥感卫星数据，通过中国资源卫星应用中心 (<https://www.cresda.com/>)、前期购买等方式获取。该数据为 C 波段多极化类型，空间分辨率达到 10 米级，原始数据为经过基础去噪、几何纠正、地理投影处理后的 2 级产品。在获取单景数据后，利用 ENVI 遥感影像处理软件对影像进行拼接和镶嵌处理，最终得到全国的 SAR 影像数据。

可见光与 SAR 影像均采用 GCS_WGS_1984 地理坐标系统，数据格式为 Geotiff，文件扩展名为

tif。在空间覆盖范围上，影像覆盖全国区域，包括港澳台地区，影像的获取年份为 2018 至 2020 年。收集到的影像数据情况如表 2 所示。

表 2 不同模态影像数据获取情况

Table 2 Acquisition of different modality image data

模态	数据源	空间分辨率	波段	等级
可见光	Google Earth	8-69m	R,G,B	L12,L13,L14,L15 层级
SAR	GF-3	10m	C	2 级产品

1.2 数据处理

数据处理主要包括空间对齐、数据清洗以及无值区域处理三部分。

(1) 空间对齐：可见光与 SAR 影像均已完成几何校正，且地图投影坐标系一致。为使 GF-3 SAR 影像与 Google Earth 可见光影像的 L15、L14、L13、L12 层级在空间分辨率上相匹配，需对 SAR 影像进行重采样处理，从而生成空间对齐的同一地理范围内的可见光与 SAR 初始影像对。为了确保像元值的平滑过渡、降低计算时间，重采样选用计算复杂度低的双线性插值法。为了验证 GF-3 SAR 影像与 Google Earth 可见光影像空间对齐效果，随机抽取了 100 对样本，通过选取同名控制点（如道路交叉口、建筑物顶点），计算可见光与 SAR 影像的均方根误差（Root Mean Square Error, RMSE），最终 RMSE 均值为 1.39 像素，总体满足多模态机场检测的样本要求。RMSE 具体计算公式如下所示。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{SAR,i} - x_{Visible,i})^2 + (y_{SAR,i} - y_{Visible,i})^2} \quad (1)$$

其中， $x_{SAR,i}$ 和 $y_{SAR,i}$ 为 SAR 影像中第 i 个控制点的坐标； $x_{Visible,i}$ 和 $y_{Visible,i}$ 为可见光影像中对应控制点的坐标。 n 为选取的控制点数量。

(2) 数据清洗：在处理过程中，需考虑所收集机场位置数据的真实性和机场在遥感影像中的可辨识度。对初始影像对进行数据清洗，剔除影像中机场目标缺失或无法辨识的影像数据。

(3) 无值区域处理：由于 SAR 影像在影像拼接过程中存在因数据缺失导致的无值区域，为确保可见光与 SAR 模态在空间上实现像素级对齐，需对数据集中的可见光与 SAR 影像对进行无值区域的检测、剔除以及裁切处理。

1.3 数据标注

本研究分别使用非商业标签工具 roLabelImg (<https://github.com/cgvict/roLabelImg>) 和 LabelImg (<https://github.com/HumanSignal/labelImg>)，对经过预处理后的可见光-SAR 遥感影像对进行标注，并添加了“airport”类别。在具体标注过程中，根据全国机场的形状和方向特点，分别按照 VOC 旋转框格式和 YOLO 水平框格式，对影像数据中机场的位置与大小进行人工标注。其中，VOC 格式数据集结构丰富，包含了影像中的位置信息、类别信息以及其他相关元数据，支持复杂场景的标注，广泛应用于计算机视觉领域的目标检测、图像分割和图像分类任务，是业界公认的标准数据集格式^[13]。YOLO 格式更为简洁明了，易于读取和处理，尤其适合实时检测和大规模数据处理，是当前主流遥感多模态目标检测算法中常用的标注格式。

为确保数据集的高质量，在其制备与标注过程中，实施了严格的质量控制和评估措施，以提升数据集的可用性和准确性。在确认所有内容无误后，将可见光影像、SAR 影像、VOC 标注、YOLO 标注进行了系统化的整理与汇总。

2 数据样本描述

本数据集的文件组织形式如表 3 所示，主要包含影像数据、标注数据两部分。所有数据均采用统一的标准格式进行记录和存储。在表 3 中，<class>、<number>、<layer>均遵循统一的命名规则。其中，<class>表示地物类型，当前仅限定为 airport；<number>对应机场位置数据的编号，例如 cn1、cn2 等；<layer>则表示采集的 Google Earth 影像层级，具体包括 L12、L13、L14、L15。下面将对这两类数据进行详细描述和分析。

表 3 数据集文件组织形式

Table 3 File organization of the dataset

文件夹	文件名称明细	内容说明
Annotations-VOC	<class>_<number>_<layer>.xml	VOC 格式标注文件
Annotations-YOLO	<class>_<number>_<layer>.txt	YOLO 格式标注文件
Visible	<class>_<number>_<layer>.png	可见光影像数据
SAR	<class>_<number>_<layer>.png	SAR 影像数据

2.1 影像数据

数据集地理空间范围涵盖中国大陆及港澳台地区共 34 个省级行政区。机场地理位置及周边环境主要涉及城市中心及近郊、城市远郊、海边或岛屿、山区或丘陵地带、荒漠或半荒漠地区、高原等场景。因此，本数据集具有空间跨度大、类型与场景多样性丰富的特点。

影像数据包括可见光与 SAR 两种模态的遥感影像数据，所有影像数据均已删除地理位置信息，最终以.png 格式存储，位深为 8 位。依据空间分辨率将影像划分为 L12、L13、L14、L15 四类。图像尺寸大小范围为 $699 \times 623 \sim 2150 \times 2150$ 像素。每个模态数据包含 1915 张图片，影像数据示例如图 2 所示。

2.2 标注数据

标注数据包括 VOC 扩展格式标注数据和 YOLO 格式标注数据。VOC 扩展格式标注数据是旋转框格式，使用 roLabelImg 工具进行人工标注，数据量为 1915。结果以.xml 格式存储，其中 cx、cy、w、h、angle 分别表示标注框中心点的 x 坐标、标注框中心点的 y 坐标、标注框的宽度、标注框的高度以及标注框相对于水平线的旋转角度，正值表示顺时针旋转，负值表示逆时针旋转。YOLO 格式标注数据是水平框格式，使用 LabelImg 工具进行人工标注，数据量为 1915。结果以.txt 格式存储。具体内容表示为：<class_id> <x_center> <y_center> <width> <height>，每个字段的值均为相对于影像宽度或影像高度的比例，取值范围在 0-1 之间。标注数据示例如图 3 所示。

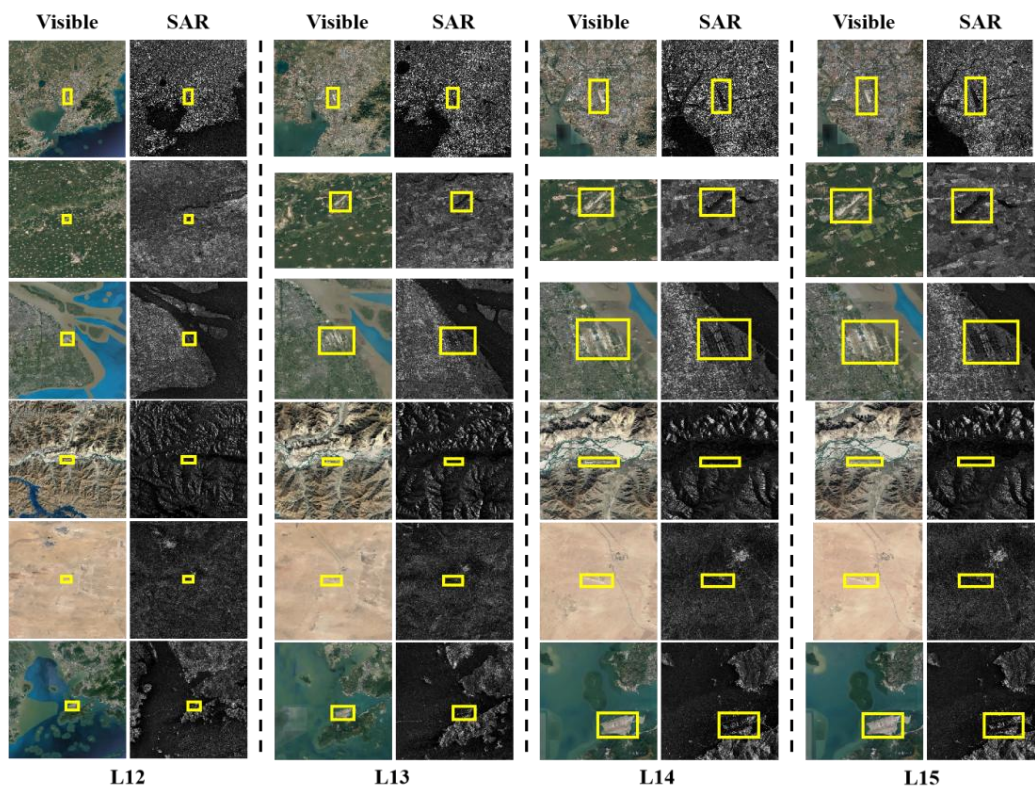


图2 影像数据示例

Figure 2 Example of image data

```
<size>
  <width>1820</width>
  <height>2150</height>
  <depth>3</depth>
</size>
<segmented>0</segmented>
<object>
  <type>robndbox</type>
  <name>airport</name>
  <pose>Unspecified</pose>
  <truncated>0</truncated>
  <difficult>0</difficult>
  <robndbox>
    <cx>852.3316</cx>
    <cy>1106.178</cy>
    <w>554.1914</w>
    <h>806.1586</h>
    <angle>2.981593</angle>
  </robndbox>
</object>
```

```
0 0.47226798462383307 0.5102325581395348 0.3651839648544756 0.4055813953488372
```

图3 标注数据示例 (上: VOC 旋转框格式; 下: YOLO 水平框格式)

Figure 3 Example of annotation data (Top: VOC rotated bounding box format; Bottom: YOLO horizontal bounding box format)

为了便于后续选择合适的目标检测模型架构以及评估模型对不同尺度目标的检测能力，本研究对数据集进行了统计分析，如图 4 所示。从数据特征来看，本数据集呈现出小目标数量多的特性，这与遥感影像的特点相契合，能够满足研究人员对遥感多模态目标检测的需求。

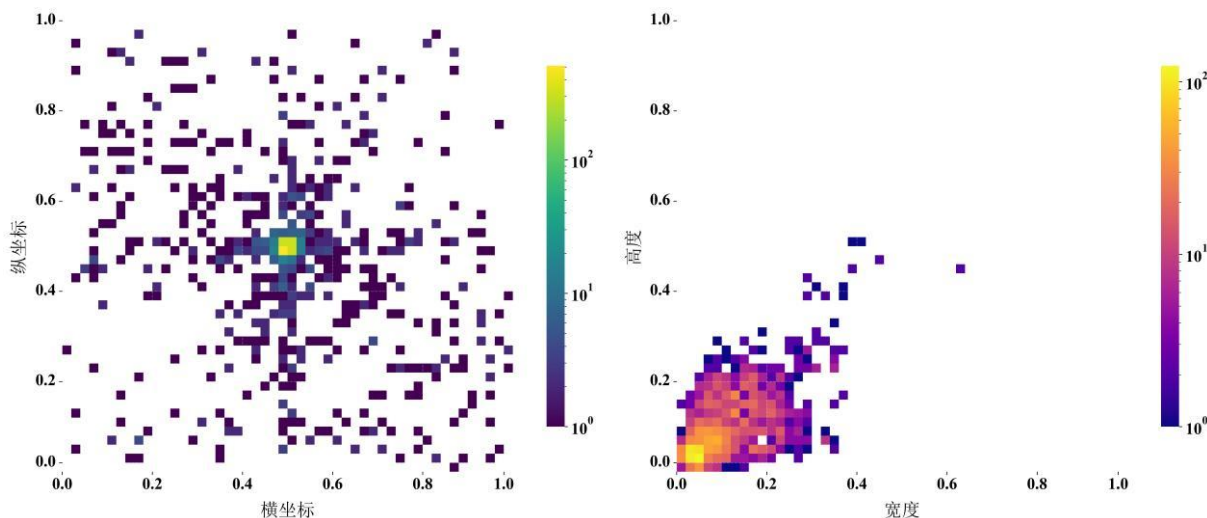


图 4 数据集统计分析（左：影像中心点相对位置；右下：目标相对高宽；数值均已归一化）

Figure 4 Statistical analysis of the dataset (Left: relative positions of image center points; Bottom-right: relative target heights and widths of targets; all values normalized)

3 数据质量控制和评估

3.1 数据质量控制

数据质量控制主要包括遥感影像样本数据质量控制与标注文件样本数据质量控制两大方面。

针对遥感影像样本数据，数据质量控制主要体现在以下几个方面：（1）数据完整性检查：对影像的投影信息、波段数量、存储完整性及存储格式进行全面检查，确保数据不存在明显问题。（2）数据多样性保障：确保影像数据采集涵盖多样化的场景和机场类型，同时具备较大的空间跨度和尺度跨度，以提升数据集的泛化能力。（3）数据处理流程优化：制定科学合理的数据清洗流程，并有效去除无值区域，从而提升数据整体质量。

针对标注文件样本数据，数据质量控制主要包括以下内容：（1）文件完备性审查：对文件命名、标注格式、存储格式以及文件字段的完整性进行检查，及时发现并解决数据问题。（2）标注标准制定：明确影像标注的标准，包括机场边界框的标注形式（水平框或旋转框）和机场类别的命名。（3）标注格式审查：对 VOC 和 YOLO 格式的标注进行格式检查，对无法按照通用标准格式解析并可视化的样本数据进行修正，确保数据集标注的可用性和准确性。（4）标注一致性检验：主要涉及目标是否存在和边界框是否准确的核实，确保影像数据中的实际内容与标注文件的描述内容一致。由于数据集中目标类型为机场，面积较大，通过目视判读即可辨别，因此，为了确保标注的一致性，本研究采用多轮质检并逐一核对的方式，对不一致的标注文件进行修正，从而保证样本标注的准确性和可靠性。

3.2 数据评估方法

为了科学地评估所构建数据集的质量和可用性，以水平框标注形式为例进行了数据评估。

在模型架构选择上，选取以下代表模型：双阶段模型 Faster RCNN^[14]，基于 CNN 的单阶段模型 YOLOv5n^[15]、YOLOv8n^[16]、YOLOv11n^[17]，基于 Transformer 的单阶段模型 RT-DETR^[18]，以及遥感多模态目标检测模型 SuperYOLO^[6]。其中，SuperYOLO 是 Zhang 等人基于 YOLOv5 框架改进的用于遥感可见光-红外目标检测领域的先进模型。

实验设置方面，数据集按 8:1:1 比例随机划分为训练集(train)、验证集(val)、测试集(test)。实验硬件为 NVIDIA GeForce RTX 2080Ti（显存 12 GB）。软件环境为 CUDA 12.3、Pytorch 2.4 深度学习框架、Python 3.10 编程语言。统一训练参数为随机梯度下降（SGD）优化器、100 训练轮数、0.001 初始学习率。

在评估指标选取上，采用准确率（Precision）、召回率（Recall）、F1 分数（F1-Score）、交并比（IoU）阈值为 0.5 时的平均精度均值（mAP@0.5）、IoU 阈值从 0.5 到 0.95 时的平均精度均值（mAP@0.5:0.95）。各指标具体计算公式如（2）-（5）所示。式中，TP（True Positive）是真正例，即模型正确地将正例预测为正例的数量。FP（False Positive）是假正例，表示模型错误地将负例预测为正例的数量。FN（False Negative）是假负例，表示模型错误地将正例预测为负例的数量。C 表示类别数量， $AP = \int_0^1 P(r)dr$ 是对某一特定类别的 Precision-Recall 曲线下的面积。不同目标检测模型在所构建数据集上的实验结果详见表 4。

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = 2 \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (5)$$

实验数据表明，在相同数据规模、不同模型架构下，基于 YOLO 架构的单模态模型表现突出，可见光模态下 YOLOv8n 达到 95.8% 的 Precision、85.4% 的 F1-Score 和 86.2% 的 mAP@0.5，SAR 模态下 YOLOv5n 取得 88.0% 的 Precision 与 81.7% 的 mAP@0.5，验证了所构建数据集的可靠性与模态特性差异。在相同数据规模、相同模型架构下，可见光模态的目标检测性能评估指标普遍优于 SAR 模态。多模态模型的评估指标接近于单模态模型的结果，但并未全面超越。值得注意的是，多模态模型 SuperYOLO（84.7% mAP@0.5）虽优于传统 Faster RCNN（67.2% mAP@0.5）在可见光的性能，但相较于最优单模态模型仍存在 1.5% 的 mAP 差距。这种性能衰减现象与 Belmouhcine 等人^[19]的研究结论一致，即多模态目标检测模型在遥感数据上的有效性仍有待验证。针对本数据集上述实验结果，多模态目标检测模型性能评估指标稍逊于单模态模型的可能影响因素包括：1）SuperYOLO 最

初是针对可见光和红外两种模态下目标检测而设计的，与本数据集在模态类型上存在明显差异，且同一地物类型的 SAR 影像电磁散射特性与红外影像热辐射特性存在本质区别；2) 本数据集所涉及的目标类型相比车辆尺度更大，影像空间分辨率相对更粗，在模型结构设计过程中需要采取针对性的优化策略。综上，本研究构建的数据集可为后续遥感多模态目标检测相关算法的进一步提升研究提供统一的基准数据。

表 4 实验结果

Table 4 Experimental results

模型	模态	Precision	Recall	F1-Score	mAP@0.5	mAP@0.5:0.95
Faster RCNN	可见光	70.5	65.0	67.6	67.2	30.0
Faster RCNN	SAR	58.0	50.0	53.7	54.4	23.4
RT-DETR	可见光	85.2	78.8	81.9	85.6	55.4
RT-DETR	SAR	78.1	67.5	72.4	74.8	47.6
YOLOv5n	可见光	90.0	77.4	83.2	84.6	51.2
YOLOv5n	SAR	88.0	72.4	79.4	81.7	50.0
YOLOv8n	可见光	95.8	77.0	85.4	86.2	53.9
YOLOv8n	SAR	87.5	66.8	75.8	78.8	49.3
YOLOv11n	可见光	87.9	78.3	82.8	85.5	51.1
YOLOv11n	SAR	83.9	69.5	76.0	77.9	45.6
SuperYOLO	可见光+SAR	87.7	82.0	84.8	84.7	49.8

4 数据使用方法和建议

4.1 数据使用方法

本数据集是首个公开发布的可见光与 SAR 遥感多模态机场检测数据集，有效弥补了遥感多模态目标检测领域在数据集数量不足、模态组合单一以及目标类型有限等方面的不足。本数据集的应用流程如图 5 所示。

本数据集能够为遥感多模态目标检测算法的研究提供有力支持。具体而言，可将其作为可见光与 SAR 遥感多模态目标检测算法的基准数据集，通过编程实现影像与标注文件的批量解析和读取，并运用遥感多模态目标检测算法进行训练。通过对训练结果的对比分析，可以进一步优化遥感多模态检测算法。

此外，本数据集在多种目标检测模型的实验中均表现出较高的性能。因此，可根据实际应用需求，将本数据集作为初始数据样本，通过补充特定应用场景下的机场数据样本，对样本数据集进行重新构建与划分。随后，通过训练得到机场目标检测模型，并将其部署于深度学习环境中。在此过程中，用户只需输入可见光或 SAR 待检测机场目标影像数据，即可自动获得可靠的检测结果。

4.2 数据局限性

本数据集在影像噪声和影像分辨率跨度方面对目标检测的影响分析与应对建议如下：

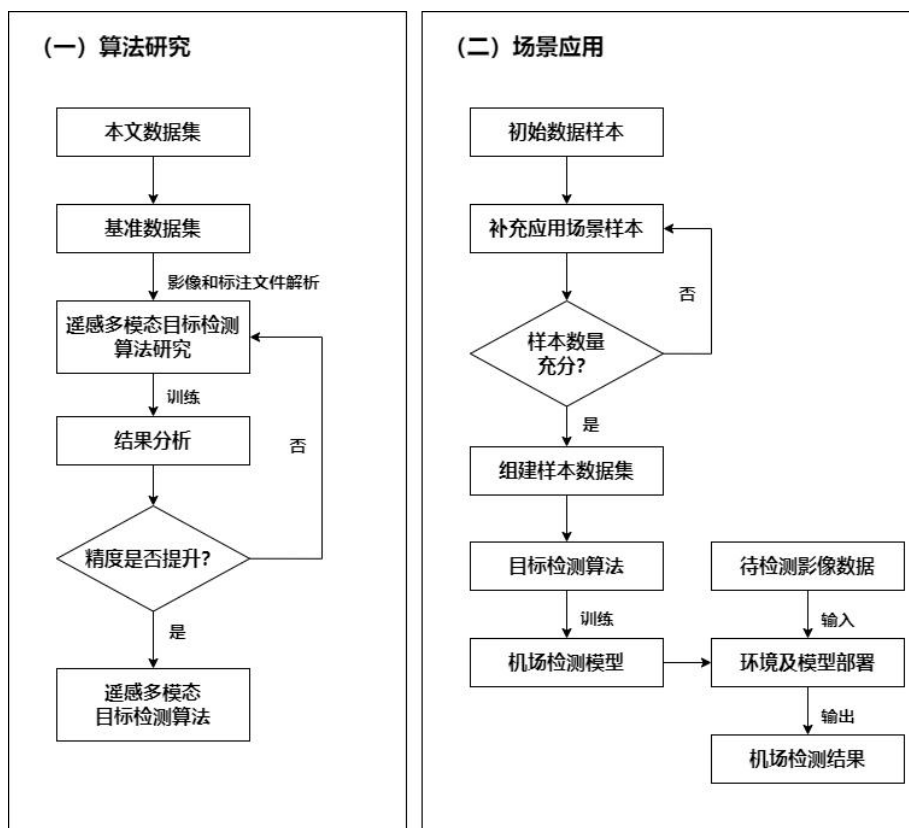


图 5 数据集用于目标检测的应用流程

Figure 5 Application workflow of the dataset for object detection

(1) SAR 影像噪声

GF-3 卫星是我国首颗 C 频段多极化 SAR 卫星,其通过主动发射微波并接收回波信号进行成像,在成像过程中会引入散斑噪声 (speckle noise),从而影响图像的清晰度和可解释性。在数据集构建过程中,所采集的 SAR 数据为 L2 级产品,本身已经经过了预处理,对大部分噪声进行了抑制,但受去噪算法限制,噪声的抑制可能并不彻底。因此,用户在后续数据集使用过程中,可根据研究的目的采取不同的噪声应对策略。如需开展最佳成像质量情况下多模态目标检测算法研究,可进一步通过滤波去噪、深度学习方法抑制 SAR 图像中的噪声,以提高图像的清晰度,其中深度学习去噪方法可以是基于频域特征动态去噪的算法 DenoDet^[20]、结合卷积神经网络的去噪模型 EchoCNN-Denoiser^[21]等近年来最新提出的模型方法;如专门研究实际成像条件下不同模态联合解译过程中各模态的重要性,则可不再进行 SAR 数据去噪处理。

(2) 影像分辨率跨度

数据集中影像分辨率跨度大,部分低分辨率影像中机场目标仅占少量像素,易导致目标漏检和尺度兼容性问题。对用户的应对建议包括:1) 在检测模型中采用自适应多尺度特征金字塔结构,融合 L12-L15 层级的跨分辨率特征;2) 在训练阶段引入超分辨率子网络,对低分辨率影像进行增强;3) 引入跨层级注意力模块,强化低分辨率特征的空间响应能力;4) 当实际应用场景的数据分辨率相对固定时,也可筛选出与之相匹配的样本子集,构建尺度相当的训练集,并据此进行模型训练。

作者分工职责

赵旭阳 (2001—),男,山东省临沂市人,硕士生,研究方向为遥感影像处理。主要承担工作:机

场点位数据收集、影像数据整理（数据清洗、无值区域处理）、数据标注、相关程序的编写、数据质量控制、数据质量评估、论文撰写等。

赵理君（1986—），男，河南省洛阳市人，博士，副研究员，研究方向为遥感影像智能解译。主要承担工作：遥感影像收集、方案设计、数据采集、数据审核、论文修改。

李宏益（1986—），男，湖南省澧县人，博士，高级工程师，研究方向为遥感影像组织管理。主要承担工作：数据分幅、样本标注。

张正（1989—），男，江苏省丰县人，博士，副研究员，研究方向为遥感影像智能处理。主要承担工作：SAR 影像数据预处理。

唐娉（1968—），女，宁夏回族自治区银川市人，博士，研究员，研究方向为遥感大数据处理与分析。主要承担工作：数据集制作方案可行性指导。

参考文献

- [1] SUN X, TIAN Y, LU W X, et al. From single- to multi-modal remote sensing imagery interpretation: a survey and taxonomy[J]. Science China Information Sciences, 2023, 66(4): 140301. DOI:10.1007/s11432-022-3588-0.
- [2] HUANG M Y, XU Y, QIAN L X, et al. A bridge neural network-based optical-SAR image joint intelligent interpretation framework[J]. Space: Science & Technology, 2021, 2021: 2021/9841456. DOI:10.34133/2021/9841456.
- [3] 付琨, 王佩瑾, 冯瑛超, 等. 遥感跨模态智能解译: 模型、数据与应用[J]. 中国科学(信息科学), 2023, 53(8): 1529 - 1559. DOI:10.1360/SSI-2023-0055.
FU K, WANG P J, FENG Y C, et al. Cross-modal remote sensing intelligent interpretation: method, data, and application[J]. Science in China (Information Sciences), 2023, 53(8): 1529 - 1559. DOI:10.1360/SSI-2023-0055.
- [4] BAHADURI B, MING Z H, FENG F C, et al. Multimodal transformer using cross-channel attention for object detection in remote sensing images[C]//2024 IEEE International Conference on Image Processing (ICIP). Abu Dhabi, United Arab Emirates. IEEE, 2024: 2620 - 2626. DOI:10.1109/ICIP51287.2024.10647683.
- [5] SHARMA M, DHANARAJ M, KARNAM S, et al. YOLOrs: object detection in multimodal remote sensing imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14: 1497 - 1508. DOI:10.1109/JSTARS.2020.3041316.
- [6] ZHANG J Q, LEI J, XIE W Y, et al. SuperYOLO: super resolution assisted object detection in multimodal remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5605415. DOI:10.1109/TGRS.2023.3258666.
- [7] 刘春霞, 孟吉星, 潘理虎, 等. 融合 RGB 与 IR 图像的遥感小目标检测方法[J]. 2025(7): 326 - 338. DOI:10.19678/j.issn.1000-3428.0069510. [LIU C X, MENG J X, PANG L H, et al. Fusion of RGB and IR for small target detection in remote sensing images[J/OL]. Computer Engineering, 2025(7): 326 - 338. DOI:10.19678/j.issn.1000-3428.0069510.]
- [8] 常凯旋, 黄建华, 孙希延, 等. 基于双模态图像融合的无人机光学小目标检测算法[J]. 激光与光电

- 子学进展, 2025, 62(4): 0428001. DOI:10.3788/LOP241283.
- CHANG K X, HUANG J H, SUN X Y, et al. Optical small target detection method by drone based on dual-modal image fusion[J]. *Laser & Optoelectronics Progress*, 2025, 62(4): 0428001. DOI:10.3788/LOP241283.
- [9] RAZAKARIVONY S, JURIE F. Vehicle detection in aerial imagery: a small target detection benchmark[J]. *Journal of Visual Communication and Image Representation*, 2016, 34: 187 - 203. DOI:10.1016/j.jvcir.2015.11.002.
- [10] VARGA L A, KIEFER B, MESSMER M, et al. SeaDronesSee: a maritime benchmark for detecting humans in open water[C]//2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, HI, USA. IEEE, 2022: 3686 - 3696. DOI:10.1109/WACV51458.2022.00374.
- [11] SUN Y M, CAO B, ZHU P F, et al. Drone-based RGB-infrared cross-modality vehicle detection *via* uncertainty-aware learning[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 32(10): 6700 - 6713. DOI:10.1109/TCSVT.2022.3168279.
- [12] XIA G S, BAI X, DING J, et al. DOTA: a large-scale dataset for object detection in aerial images[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. IEEE, 2018: 3974 - 3983. DOI:10.1109/CVPR.2018.00418.
- [13] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, et al. The pascal visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303 - 338. DOI:10.1007/s11263-009-0275-4.
- [14] 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. DOI:10.1109/TPAMI.2016.2577031.
- [15] Jocher G, et al. YOLOv5. GitHub repository (2020). <https://github.com/ultralytics/yolov5>.
- [16] SOHAN M, SAI RAM T, RAMI REDDY C V. A review on YOLOv8 and its advancements[C]//Data Intelligence and Cognitive Informatics. Singapore: Springer, 2024: 529 - 545. DOI:10.1007/978-981-99-7962-2_39.
- [17] KHANAM R, HUSSAIN M. YOLOv11: an overview of the key architectural enhancements[EB/OL]. 2024: arXiv: 2410.17725. <https://arxiv.org/abs/2410.17725>
- [18] ZHAO Y A, LV W Y, XU S L, et al. DETRs beat YOLOs on real-time object detection[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA. IEEE, 2024: 16965 - 16974. DOI:10.1109/CVPR52733.2024.01605.
- [19] BELMOUHCINE A, BURNEL J C, COURTRAI L, et al. Multimodal object detection in remote sensing[C]//IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium. Pasadena, CA, USA. IEEE, 2023: 1245 - 1248. DOI:10.1109/IGARSS52108.2023.10281486.
- [20] DAI Y M, ZOU M R, LI Y X, et al. DenoDet: attention as deformable multisubspace feature denoising for target detection in SAR images[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2025, 61(2): 4729 - 4743. DOI:10.1109/TAES.2024.3507786.
- [21] TWINKLE S A, KAMILYA S, MUKHERJEE J. EchoCNN-Denoiser: a reservoir computing inspired deep learning model for enhanced synthetic aperture radar image despeckling[J]. *Journal of Applied*

Remote Sensing, 2025, 19(2): 026501. DOI:10.1117/1.jrs.19.026501.

论文引用格式

赵旭阳, 赵理君, 李宏益, 等. 可见光-SAR 遥感多模态机场检测数据集[J/OL]. 中国科学数据, 2025, 10(4). (2025-12-22). DOI: 10.11922/11-6035.csd.2025.0031.zh.

数据引用格式

赵旭阳, 赵理君, 李宏益, 等. 可见光-SAR 遥感多模态机场检测数据集[DS/OL]. V1. Science Data Bank, 2025. (2025-04-08). DOI: 10.57760/sciencedb.j00001.01276.

A dataset of multimodal visible-SAR imagery for airport detection in remote sensing

ZHAO Xuyang^{1,2}, ZHAO Lijun^{1*}, LI Hongyi¹,
ZHANG Zheng¹, TANG Ping¹

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, P.R. China

2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, P.R. China

*Email: zhaolj201934@aircas.ac.cn

Abstract: Remote sensing image object detection is a significant area in Earth observation tasks. By integrating remote sensing data from various sensors and utilizing multimodal complementary information, the performance of object detection models can be significantly improved. To address the current limitations in the remote sensing multimodal object detection field, such as insufficient datasets, limited modality combinations, and restricted object categories, this study constructs a multimodal Visible-SAR dataset for airport detection in remote sensing imagery named MVSDA, which is based on Google Earth visible images and Gaofen-3 (GF-3) Synthetic Aperture Radar (SAR) imagery. During dataset construction, standardized protocols were followed for data collection, processing, and annotation, encompassing a series of procedures such as multi-source data acquisition, spatial alignment, non-informative region filtering, expert-guided annotation, quality control, and evaluation. This dataset is characterized by pixel-level alignment, large geographic coverage, multiple spatial resolutions, and diverse environmental conditions, which supports the training requirements of mainstream deep learning models. This dataset provides essential foundational resources for researchers in related fields and holds practical value for advancing research in remote sensing multimodal object detection.

Keywords: multimodal; object detection; remote sensing; airport; visible light; SAR

Dataset Profile

Title	A dataset of multimodal visible-SAR imagery for airport detection in remote sensing
Data corresponding author	ZHAO Lijun (zhaolj201934@aircas.ac.cn)
Data authors	ZHAO Xuyang, ZHAO Lijun, LI Hongyi, ZHANG Zheng, TANG Ping
Time range	2018–2020
Geographical scope	34 provincial administrative regions across China
Spatial resolution	8–69 m
Data volume	15.18 GB
Data format	*.png, *.txt, *.xml, *.zip
Data service system	https://doi.org/10.57760/sciencedb.j00001.01276
Source of funding	Civil Aerospace Technology Pre-research Project (D040404)
Dataset composition	The dataset mainly consists of two parts: image data and annotation file data, including a total of 4 compressed files in zip format as follows: (1) Annotations_VOC.zip is a compressed package of annotation files in the rotated bounding box VOC format. The data volume is 1915, stored in xml format. (2) Annotations_YOLO.zip is a compressed package of annotation files in the horizontal bounding box YOLO format, with The 1915 data entries, stored in txt format. (3) Visible.zip is a compressed package of remote sensing images in the visible light modality, with 1,915 data entries, stored in png format. (4) SAR.zip is a compressed package of remote sensing images in the SAR modality, with 1,915 entries, stored in png format.