

DOI: 10.20189/j.cnki.CN/61-1527/E.202503001

多源多视角飞行器认知导航数据集制备

陈璐¹, 李清格², 杨小冈¹, 卢瑞涛¹

(1. 火箭军工程大学, 陕西 西安 710025;

2. 西安石油大学 电子工程学院, 陕西 西安 710065)

摘要: 针对现有大部分飞行器视觉导航数据集特征信息有限的问题, 构建了一个多源、多视角、多场景的飞行器认知导航数据集, 可有效支撑飞行器的精确地理定位。首先, 基于深度特征匹配算法在认知导航区卫星图像内选取导航适配区, 并基于 SAM (segment anything model) 语义分割算法获得特征显著且稳定的典型导航标志物及其语义信息; 然后, 利用真实无人机航拍和 Google Earth 提供的 3D 仿真模型采集充足的导航标志物飞行器视角图像; 最后, 基于 SuperPoint 特征点提取算法自动对图像进行标注, 并设计基于 WordNet 体系的认知导航数据集多层次结构, 从而最终获得具有地理位置信息和深度语义特征标注的飞行器认知导航数据集。分析评估表明, 所构建的数据集覆盖场景广泛、特征信息丰富、数据来源多样, 为飞行器认知导航算法的训练和验证提供了高质量的数据样本。

关键词: 视觉认知导航; 飞行器匹配定位; 数据集制备; 导航标志物; 自动标注

中图分类号: TP391

文献标志码: A

文章编号: 2097-2741(2025)03-001-09

Multi-Source and Multi-View Aircraft Cognitive Navigation Dataset Construction

CHEN Lu¹, LI Qingge², YANG Xiaogang¹, LU Ruitao¹

(1. Rocket Force University of Engineering, Xi'an 710025, China;

2. School of Electronic Engineering, Xi'an Shiyou University, Xi'an 710065, China)

ABSTRACT: To overcome limited feature information in most existing aircraft visual navigation datasets, a multi-source, multi-view, and multi-scene aircraft cognitive navigation dataset was constructed to support accurate aircraft geolocation. First, navigation adaptation zones were selected from satellite images of cognitive navigation regions using a deep feature matching algorithm. Distinctive and stable navigation landmarks, along with their semantic information, were then extracted using the segment anything model (SAM) semantic segmentation algorithm. Subsequently, sufficient aircraft-view images of the navigation landmarks were collected using real drone aerial photography and 3D simulation models provided by Google Earth. Finally, the SuperPoint feature point extraction algorithm was applied to automatically annotate the images. A

收稿日期: 2025-04-23

修回日期: 2025-05-20

录用日期: 2025-05-25

基金项目: 国家自然科学基金 (62276274); 陕西省重点研发计划 (2024CY2-GJHX-42)

第一作者: 陈璐 (1991—), 女, 讲师, 主要从事图像识别与智能控制研究。E-mail: chenlu_dut@163.com

引用格式: 陈璐, 李清格, 杨小冈, 等. 多源多视角飞行器认知导航数据集制备[J]. 火箭军工程大学学报, 2025, 39 (3): 1-9.
CHEN Lu, LI Qingge, YANG Xiaogang, et al. Multi-source and multi-view aircraft cognitive navigation dataset construction[J]. Journal of Rocket Force University of Engineering, 2025, 39 (3): 1-9.

multi-level structure of cognitive navigation datasets based on the WordNet system was designed, producing an aircraft cognitive navigation dataset annotated with geolocation information and deep semantic features. The evaluation and analysis demonstrate that the constructed dataset covers a wide range of scenarios, provides rich feature information, and incorporates diverse image sources, offering high-quality data samples for the training and validation of aircraft cognitive navigation algorithms.

KEYWORDS: visual cognitive navigation; aircraft matching localization; dataset preparation; navigation landmark; automatic annotation

随着无人机、自动驾驶飞行器等低空智能装备的快速发展,视觉认知导航技术逐渐成为飞行器自主导航的核心研究方向之一。在复杂动态环境中,常见的卫星导航易出现拒止失效的情况,而视觉认知导航凭借其被动感知、高精度定位和强环境适应性等优势,成为解决这一问题的重要手段^[1]。视觉认知导航通过机载光学传感装置实时采集环境数据,采用深度神经网络对飞行场景进行关键特征解析,而后对这些特征信息进行知识推理,实现飞行器环境智能感知、自主定位定向和在线规划决策等功能^[2]。传统视觉导航依赖几何特征匹配,而视觉认知导航更加强调对场景的语义理解与智能决策,使飞行器具备自主理解外界环境的能力,能更好地适应环境的动态变化和不确定性^[3]。

大规模、高质量的数据集为提升认知导航算法在复杂场景下的鲁棒性与泛化能力提供了强有力的支撑。近年来,飞行器视觉导航领域涌现出多个具有代表性的数据集^[4]。例如, CVUSA^[5]与 CVACT^[6]是经典的跨视角地理定位数据集,提供了覆盖乡村和城市等多样化场景的地面全景图像与卫星图像,适用于无人机与地面的协同定位或跨视角图像检索。VIGOR^[7]数据集采自于美国 4 个主要城市的建筑场景,通过地面图像与卫星图像一对多的匹配实现了无人机视觉定位。University-1652^[8]以多模态跨视角为特色,整合了 72 所大学 1 652 个建筑的无人机航拍图、地面街景及卫星三视图,每栋建筑物平均有 1 幅卫星图像和 54 幅无人机图像,支持建筑级别的细粒度检索与定位。SUES-200^[9]数据集包含多高度多场景的无人机航拍图以及对应的卫星图像,覆盖了公园、学校、湖泊、公共建筑等多种城镇场景,可有效提高视觉导航在复杂多变场景下的适应能力。以上这些数据集提供了丰富的场景和图像,在推动飞行器跨视角检索定位及复杂环境导航等研究中发挥

了重要作用。

然而,上述数据集也存在一定的局限性。首先,其数据大多来自于谷歌地图、微软必应地图等,难以充分反映真实飞行场景,尽管 SUES-200 数据集中包含了真实航拍图,但仅覆盖上海工程科技大学及其周边,数据多样性不足;其次,数据均为可见光图像,来源单一,提供环境信息有限,无法支撑开展多模态融合感知与导航相关研究;最后,数据集大多将飞行器导航问题简单定义为跨视角图像检索任务,即通过无人机视角图像检索最相似的卫星图像来实现地理定位,但由于缺乏精细的标注信息,也不包含经纬度数据,导致无法获得飞行器的精确位置^[10-11]。伴随新型飞行器视觉认知导航技术的持续发展,对导航数据集的覆盖场景、图像模态以及时空信息等均提出了更高的要求,因此本文制备了一个多源、多视角、多场景飞行器认知导航数据集,可为飞行器高精度视觉认知导航算法的训练与验证提供数据基础。

1 飞行器认知导航数据集制备流程

飞行器认知导航数据集的制备流程如图 1 所示,主要是通过导航适配区选取、导航标志物选择、标志物航拍图像采集以及数据集自动标注等关键环节来完成飞行器认知导航数据集的制备。

1.1 认知导航适配区选取

高分辨率遥感卫星影像中的导航基准地物兼具高辨识度与时空稳定性特征,其鲜明几何特征与精确地理坐标可作为导航的可靠信标,极大地降低导航偏差^[12]。为准确选取导航标志物,首先需要在认知导航区域内计算出辨识度高、特征稳定且大小适中的导航适配区^[13]。传统的适配区选取方法大多依赖先验知识,缺乏通用性和可扩展

性^[14], 因此本文采用了基于深度特征匹配的适配区选取方法, 如图2所示。具体的选取流程如下。

1) 网格化剖分: 对导航区域卫星遥感影像进行等间距网格采样, 生成标准尺寸的参考子图集;

2) 模拟航拍: 对于每一张参考子图, 利用高斯模糊、椒盐噪声、仿射变换、透视变换等生成模拟航拍图;

3) 特征关联: 基于深度特征, 实现基准子图与模拟航拍图之间的跨模态特征匹配;

4) 量化评估: 计算匹配次数和匹配成功率;

5) 动态筛选: 建立双阈值约束机制, 由同时满足匹配次数和匹配成功率阈值条件的参考子图集形成导航适配区。

1.2 典型导航标志物选择

理想的导航标志物应同时满足3个核心特征, 即唯一性、显著性和稳定性, 以确保认知导航的精准性。当空间环境中存在特征相似的标志物时, 易引发飞行器的位置误判, 导致导航性能显著下降^[15]。具体而言, 合格的导航标志物需在几何构型(如特殊轮廓)、光学属性(如高对比色系)或物理尺寸等维度与周边背景形成显著差异。同时, 导航标志物应具备抵御环境干扰的能力, 如季节更替带来的植被变化、光照条件改变以及自然侵蚀等因素引发的形态变化。本文基于图像语义分

割模型SAM(segment anything model)^[16], 通过解析标志性建筑的深层语义特征, 获取高精度的像素级语义分割结果, 进而筛选出符合导航需求的标志物集合, 能够有效实现认知导航区域内导航标志物的智能识别与择优选取。

1.3 导航标志物航拍图像采集

飞行器认知导航数据集中, 每个导航标志物均包含卫星基准视图与飞行器航拍视图2类数据。其中, 卫星基准视图来源于Google Earth平台, 是由认知导航区卫星遥感影像依次经过适配区选取、导航标志物分割而生成的参考子图。飞行器航拍视图则由多源传感器协同采集, 其中一部分是通过搭载双光吊舱的旋翼无人机, 在300 m和400 m飞行高度下实际采集得到的可见光、红外双波段图像; 另一部分则是基于Google Earth提供的3D模型, 通过模拟飞行过程采集得到的导航标志物飞行器航拍视图, 以解决实际飞行采图成本高、耗时长的问題, 采集过程如图3所示。

首先, 基于导航标志物的经纬度信息, 调用Google Earth中对应的建筑物三维模型; 随后, 设计圆锥螺旋采样路径, 飞行器将沿预设采样路径绕标志性建筑物飞行3周, 高度由500 m逐步递减至200 m; 最后, 进行模拟飞行, 以每秒30帧的速率记录飞行视频, 由此获得多尺度、多视角的

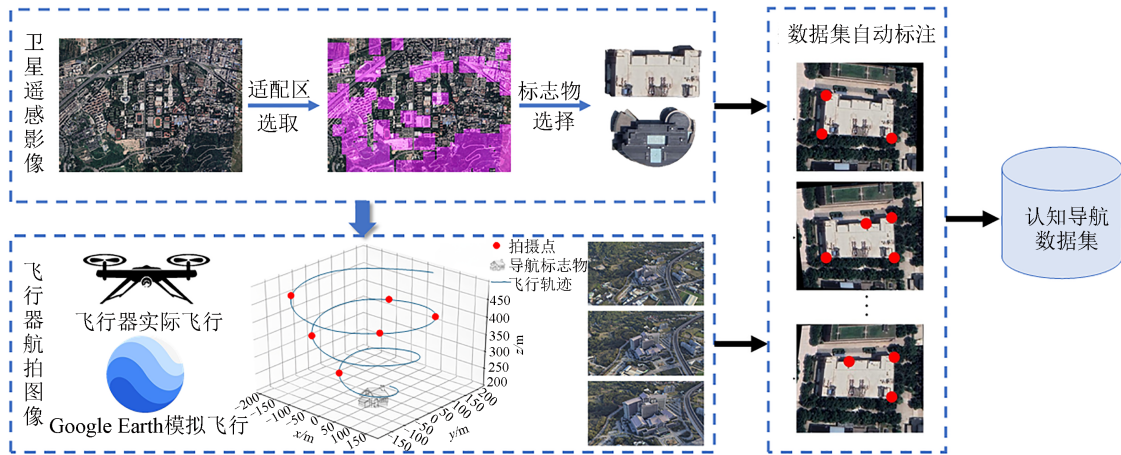


图1 飞行器认知导航数据集制备流程

Fig. 1 Preparation process of aircraft cognitive navigation dataset

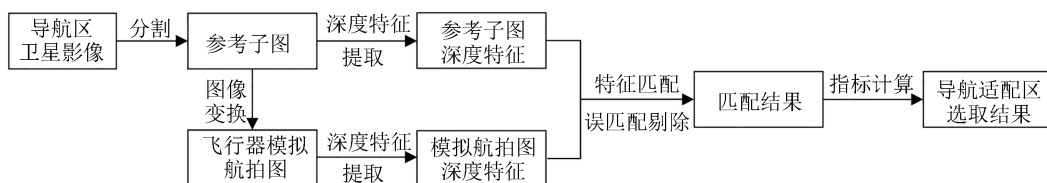


图2 认知导航适配区选取流程

Fig. 2 Selection process of cognitive navigation adaptation zone

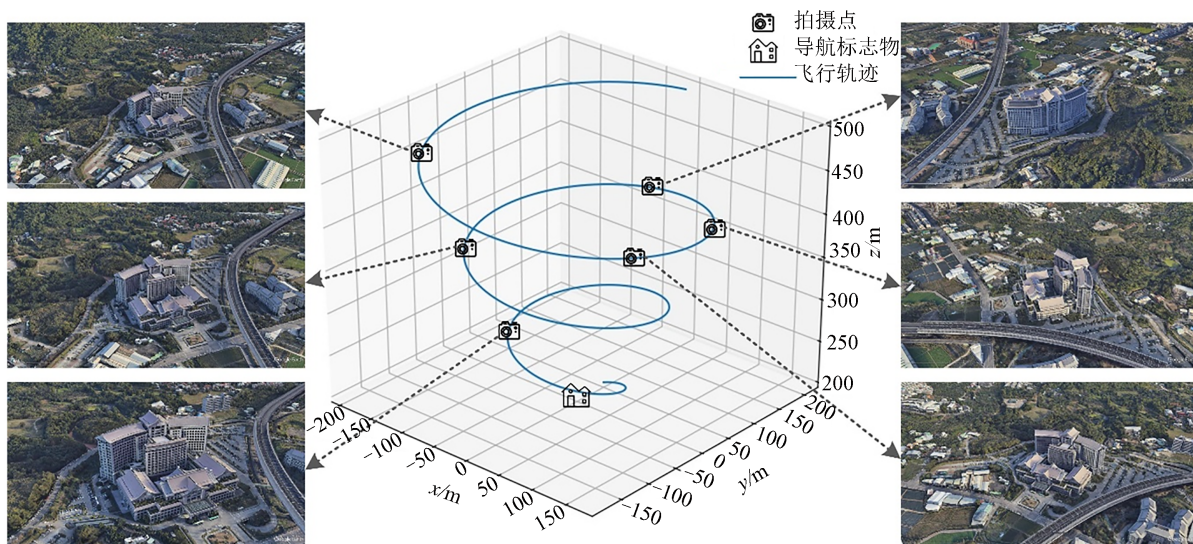


图 3 飞行器视图的模拟仿真生成

Fig. 3 Simulation generation of aircraft views

全维度导航标志物航拍图像序列。

1.4 数据集标注

为满足后续导航算法的学习和训练需求，需要对数据集中导航标志物的特征信息进行标注，具体包括导航标志物卫星-飞行器航拍图像对的匹配特征点对信息以及导航标志物外围特征点集合（即语义信息）。

针对大规模数据集的标注，传统全人工标注方法存在效率低且主观因素影响大的问题^[17]。本文采用级联标注方式，即首先基于 SuperPoint^[18]特征提取算法实现特征点自检测，继而采用特征匹配算法建立特征点对的关联关系，最后基于 La-belme 软件进行特征点的修正和标注。同时，对于导航标志物的语义信息，通过 1.2 节中基于 SAM 的语义分割可获得导航标志物对应的掩码图，掩码外围特征点集合即可作为语义标注信息。语义信息的添加，有助于提高飞行器在复杂场景下的认知能力，为飞行器视觉认知导航提供准确、丰富的地理位置信息和深度特征点数据^[19]。典型导航标志物选取和标注的准确度对于确保飞行器视觉认知导航的性能表现至关重要，下面将重点介绍导航标志物选择和数据集自动标注方法。

2 导航标志物选择及数据集自动标注方法

2.1 基于 SAM 语义分割的典型导航标志物选择

在导航适配区内，基于 Meta AI 团队研发的 SAM 图像语义分割算法，可实现导航标志物精细化提取。该创新性模型通过输入多样化提示信息，

包括点标记、矩形框选、文本描述或既有掩码，无需针对特定任务进行模型微调即能完成未知场景的物体分割，展现出卓越的零样本迁移能力^[20]。基于 SAM 的标志物分割模型结构如图 4 所示。

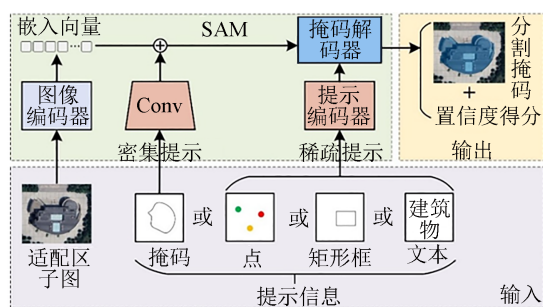


图 4 基于 SAM 的标志物分割模型结构

Fig. 4 Architecture diagram of the SAM-based model

如图 4 所示，SAM 模型主要包括图像编码器、提示编码器和掩码解码器 3 个部分。其中，图像编码器基于掩码自编码器（masked autoencoders, MAE）^[21]预训练的视觉 Transformer（vision transformer, ViT）^[22]，通过层次化特征提取生成图像的特征嵌入表示；提示编码器为双通道体系，其中，稀疏提示通道将点、框等几何要素与文本语义编码作为特征向量，而密集提示通道通过卷积层对输入掩码进行空间压缩（4 倍降采样），最终实现多模态提示信息与图像特征的深度融合；最后，通过轻量化 Transformer 掩码解码器，利用提示自注意力、交叉注意力迭代优化图像和提示的嵌入表示，输出像素级精度的语义分割结果。如图 5 所示，在导航适配区子图内，根据输入的点、矩形框和文本提示信息，基于 SAM 语义分割模型

获得标志性建筑物掩码, 即可选择出具有空间唯一性、视觉显著性和时空稳定性的典型导航标志物。

2.2 基于自监督学习的特征点自动标注

采用基于自监督学习的方式, 无需人工标注训练数据, 可大大降低数据制备的工作量^[23]。基于自监督学习的特征点自动标注原理如图 6 所示,

模型为双通道伪孪生 SuperPoint 网络, 2 个通道具有独立的权重参数, 可分别针对性地提取一种类型图像的特征信息, 因而能够增强对图像之间本质相似特征的捕捉, 有效抑制光照变化、大气干扰等因素导致的表面差异。

自动标注前, 利用包含丰富角点特征的人工合成数据集对 SuperPoint 模型进行预训练, 生成

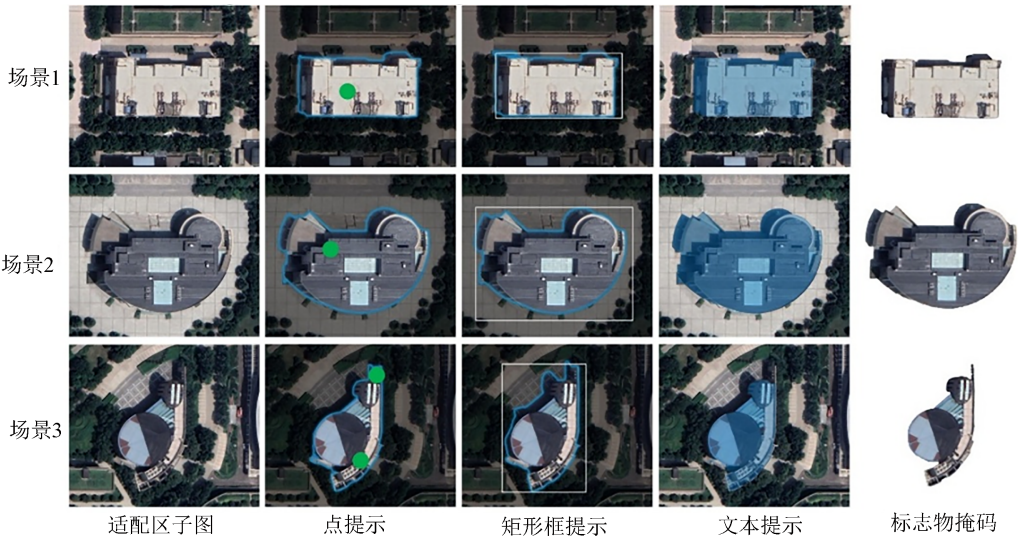


图 5 典型导航标志物分割结果
Fig. 5 Segmentation results of typical navigation landmarks

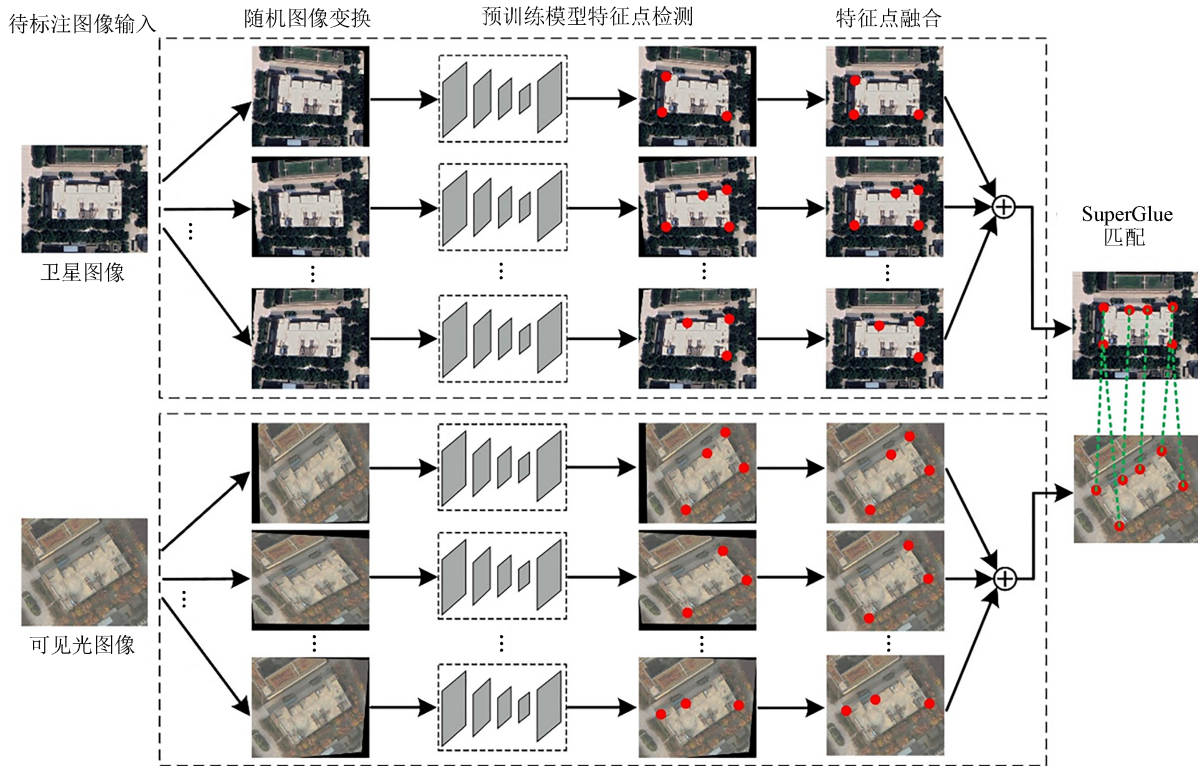


图 6 基于自监督学习的特征点自动标注原理
Fig. 6 Principle of automatic feature point annotation based on self-supervised learning

待标注图像的伪真值特征点标签。标注过程中，首先将待标注的图像及其伪真值特征点标签进行随机单应性变换，并输入至预训练的 SuperPoint 网络；接着，双通道网络分别输出提取到的特征点，通过计算与伪标签的损失值，监督 2 个通道的模型参数进行更新和调整；然后，对所有变形图像的特征点进行融合，得到待标注图像的特征点；最后，通过 SuperGlue^[24] 算法得到待标注异源图像对间特征点的匹配关系，同时剔除不共有的特征点，完成匹配特征点对的自动标注。

3 认知导航数据集结构及分析评估

3.1 飞行器认知导航数据集结构

本文认知导航数据集结构设计借鉴了 Word-Net 语义知识体系^[25]。作为自然语言处理领域的经典词库，WordNet 以同义词集合为基本语义单元，通过层级化的语义网络组织词汇资源，其蕴含的丰富语义关联特性可显著提升词汇检索的准确性与效率。基于这一优势，本文构建了三级递进式的认知导航数据架构，如图 7 所示。其中，顶层架构依据数据应用场景划分为训练集与测试集 2 大模块，中间层根据导航标志物类别进行分类，底层结构根据图像获取方式将数据细分为卫星视图、飞行器航拍视图等，形成涵盖多尺度、多视角、多场景的多维数据体系。这种层次化的数据组织架构深度融合了 WordNet 的语义关联特性，不仅强化了数据间的语义关联性，更有助于实现数据资源的系统性整合与高效管理。具体的认知导航数据由图像和标注文件 2 个部分组成，其中，图像

为多种场景下的导航标志物图像文件，标注文件则清晰、准确地描述了对应图像包含的视觉信息，包括经纬度信息、特征点信息、导航标志物语义信息等。

本文所构建的三级递进式认知导航数据集结构能够有效强化图像数据的结构化表征与语义检索效率。在具体应用场景中，飞行器根据任务需求，通过语义信息匹配实现导航标志物的精准快速检索。当面对新的导航场景时，可在保持原有数据结构稳定的前提下，动态扩展更新数据节点，从而保证数据集的兼容性与可拓展性。

3.2 飞行器认知导航数据集分析评估

本文制备的飞行器认知导航数据集包含陕西西安和安徽凤阳城郊区域的 8 个认知导航场景，每个场景下选择 5 个标志性建筑作为导航标志物，数据集组成如表 1 所示，认知导航数据集样本示例如图 8 所示。

表 1 认知导航数据集组成

Table 1 Composition of the cognitive navigation dataset

视图类型	训练集			测试集		
	图像数量	导航标志物	场景数量	图像数量	导航标志物	场景数量
飞行器	4 095	25	5	2 484	15	3
卫星	25	25	5	15	15	3

飞行器认知导航数据集由 40 张地理位置信息已知的卫星视图图像和 6 579 张飞行器视图图像组成。其中，训练集的 5 个认知导航场景中共有卫星视图图像 25 张、飞行器视图图像 4 095 张，其中包含的飞行器视图通过无人机实际飞行和 Google

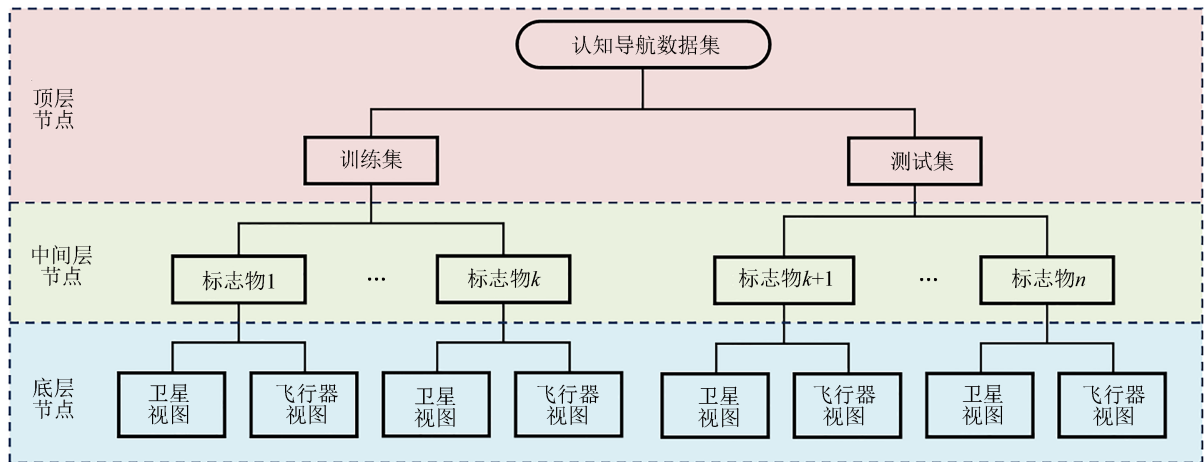


图 7 认知导航数据集结构

Fig. 7 Structure of the cognitive navigation dataset

Earth 模拟飞行 2 种方式采集得到, 含 1 390 张红外和 2 705 张可见光图像。测试集的 3 个场景共有卫星视图图像 15 张、飞行器视图图像 2 484 张, 其中飞行器视图均由无人机实际飞行采集而来, 含 828 张红外和 1 656 张可见光图像。

本文制备的认知导航数据集与其他飞行器视觉导航数据集的对比如表 2 所示, 表中详细展示了各数据集在图像数量、图像来源、视图类型、目标类型及经纬度信息方面的具体参数。

通过分析图 8 和表 2 可以看出, 对比众多经典的飞行器视觉导航数据集, 本文所构建的认知导航数据集具有以下优势: 1) 图像数量充足, 可有效支撑飞行器认知导航算法的训练和测试; 2) 图像来源丰富, 除常见的卫星、可见光图像外, 还采集了相应的红外图像, 提供了更加丰富的环境

感知与导航信息; 3) 导航场景多样, 包含城市、郊区的多种建筑物和地形, 支撑认知导航算法在未知环境中的泛化能力; 4) 时空信息完善, 数据集自动标注了导航标志物的丰富深度特征信息及精确地理位置, 可有效增强无人机对复杂场景的语义理解能力, 支持图像间的精细匹配, 可实现高精度地理定位。

与传统视觉导航数据集相比, 本文数据集还可用于探索多模态数据融合导航、飞行器视角目标定位等, 有助于飞行器开展复杂任务执行、多模态感知与自主决策导航。但目前该数据集还存在一些局限性, 如部分数据基于模拟环境仿真生成, 与真实复杂场景仍存在差距; 标注环节还需要结合人工选择和修正, 尚未实现完全自动标注, 后续仍需在数据采集及标注方面进一步完善。

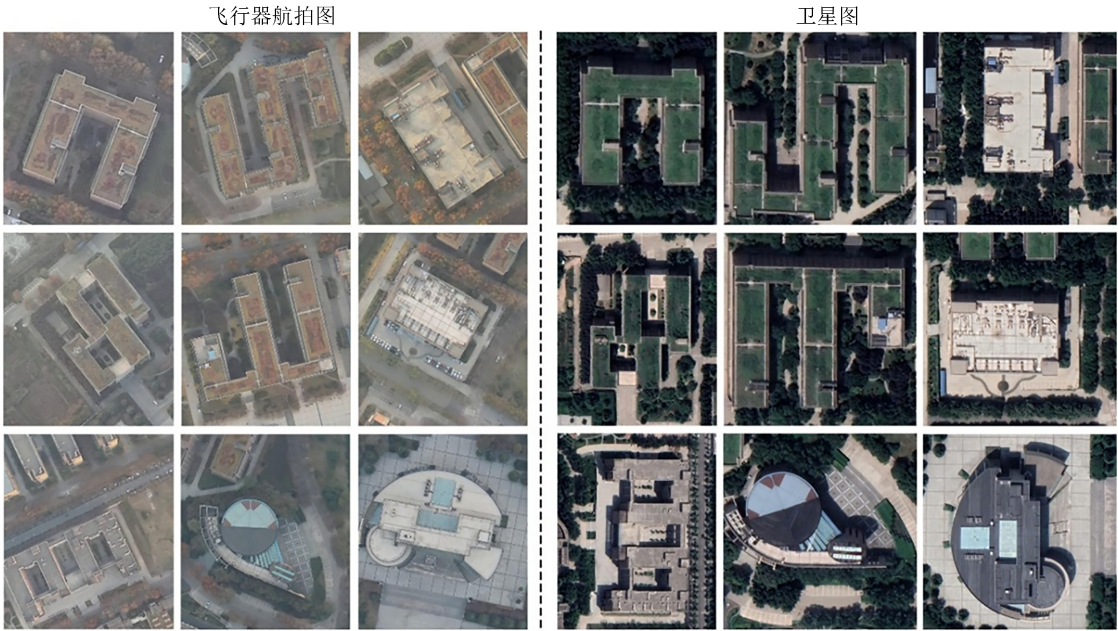


图 8 认知导航数据集样本示例

Fig. 8 Examples of the cognitive navigation dataset

表 2 本文数据集与其他飞行器视觉导航数据集的对比

Table 2 Comparison between the proposed dataset and other aircraft visual navigation datasets					
参数	本文数据集	University-1652 ^[8]	SUES-200 ^[9]	CVUSA ^[5]	CVACT ^[6]
训练集数量	5×164.8	701×71.64	120×51	35 500×2	35 500×2
图像来源	卫星、可见光、红外	卫星、可见光	卫星、可见光	卫星、可见光	卫星、可见光
视图类型	卫星、飞行器	卫星、飞行器、地面	卫星、飞行器	卫星、地面	卫星、地面
目标类型	标志建筑物	建筑物	—	—	—
经纬度信息	有	无	无	无	无

注: 训练集数量=目标场景数×单个场景的平均图片数。

4 结论

为满足新型飞行器视觉认知导航技术的发展需求, 本文通过导航适配区选取、导航标志物分割、飞行器视图采集、数据自动标注及数据集层次化结构与格式规范设计, 构建了一个多源、多视角、多场景的飞行器视觉认知导航数据集。数据集包含卫星及飞行器 2 种视图, 总量 6 619 张, 有卫星、可见光、红外 3 种图像, 覆盖了大量城镇及郊区复杂场景, 标注了精确的地理位置信息与丰富的深度语义特征, 与其他常见飞行器视觉导航数据集相比, 在图像模态、覆盖场景、时空信息等方面具有明显优势, 为飞行器实现快速、精确的视觉认知导航奠定了坚实的数据基础。后续, 可考虑增加并继续扩充真实采集数据, 同时引入自适应的图像标注算法, 以进一步提高认知导航数据集的制备质量。

参考文献

- [1] 刘学俊, 贺伟, 贺小龙. 基于卫星影像的无人机全局位姿估计方法[J]. 测绘与空间地理信息, 2025, 48(3): 83-86.
LIU X J, HE W, HE X L. UAV global pose estimation method based on satellite image[J]. Geomatics & Spatial Information Technology, 2025, 48(3): 83-86.
- [2] LI Q G, YANG X G, LU R T, et al. Vision-ICE: air-ground integrated intelligent cognition visual enhancement system based on a UAV[J]. Drones, 2023, 7(4): 1-17.
- [3] JIN Y F, YU L, CHEN Z, et al. A mono SLAM method based on depth estimation by DenseNet-CNN[J]. IEEE Sensors Journal, 2022, 22(3): 2447-2455.
- [4] 周博文, 李阳, 马鑫骥, 等. 深度学习的跨视角地理定位方法综述[J]. 中国图象图形学报, 2024, 29(12): 3543-3563.
ZHOU B W, LI Y, MA X J, et al. A survey of cross-view geo-localization methods based on deep learning[J]. Journal of Image and Graphics, 2024, 29(12): 3543-3563.
- [5] WORKMAN S, SOUVENIR R, JACOBS N. Wide-area image geolocalization with aerial reference imagery[C]//2015 IEEE International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2015: 3961-3969.
- [6] LIU L, LI H D. Lending orientation to neural networks for cross-view geo-localization[C]//2019 IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2019: 5624-5633.
- [7] ZHU S J, YANG T, CHEN C. VIGOR: cross-view image geo-localization beyond one-to-one retrieval[C]//2021 IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2021: 3640-3649.
- [8] ZHENG Z D, WEI Y C, YANG Y. University-1652: a multi-view multi-source benchmark for drone-based geo-localization[C]//Proceedings of the 28th ACM International Conference on Multimedia. New York, USA: ACM, 2020: 1395-1403.
- [9] ZHU R Z, YIN L, YANG M Z, et al. SUES-200: a multi-height multi-scene cross-view image benchmark across drone and satellite[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(9): 4825-4839.
- [10] ZENG Z L, WANG Z, YANG F, et al. Geo-localization via ground-to-satellite cross-view image retrieval[J]. IEEE Transactions on Multimedia, 2022, 25: 2176-2188.
- [11] LI J S, YANG C F, QI B J, et al. 4SCIG: a four-branch framework to reduce the interference of sky area in cross-view image geo-localization[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 4703818.
- [12] ZHANG X X, WANG C Y, JIN J, et al. Object detection of VisDrone by stronger feature extraction FasterRCNN[J]. Journal of Electronic Imaging, 2023, 32: 013018.
- [13] HUANG D H, CHEN Y, LIU Y, et al. Adaptive assignment for geometry aware local feature matching[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2023: 5425-5434.
- [14] 李邦杰, 武健, 张大巧, 等. 基于卷积神经网络的景象匹配区适配性筛选方法[J]. 火箭军工程大学学报, 2024, 38(3): 45-50, 59.
LI B J, WU J, ZHANG D Q, et al. A CNN-based scene matching adaptability screening method[J]. Journal of Rocket Force University of Engineering, 2024, 38(3): 45-50, 59.
- [15] CHEN Y, HUANG D H, XU S, et al. Guide local feature matching by overlap estimation[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022, 36(1): 365-373.
- [16] KIRILLOV A, MINTUN E, RAVI N, et al.

- Segment anything[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ, USA: IEEE, 2023: 4015-4026.
- [17] 黄攀, 范继伟, 杨小冈, 等. 智能图像导引头目标数据集制备技术研究[J]. 战术导弹技术, 2022(2): 105-112.
- HUANG P, FAN J W, YANG X G, et al. Research on target dataset preparation technology of intelligent image seekers[J]. Tactical Missile Technology, 2022(2): 105-112.
- [18] DETONE D, MALISIEWICZ T, RABINOVICH A. SuperPoint: self-supervised interest point detection and description[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway, NJ, USA: IEEE, 2018: 224-236.
- [19] FAN B, ZHOU J J, FENG W S, et al. Learning semantic-aware local features for long term visual localization[J]. IEEE Transactions on Image Processing, 2022, 31: 4842-4855.
- [20] 蔡伟, 高蔚洁, 蒋昕昊, 等. 可见光条件下伪装目标智能化分割算法综述[J]. 火箭军工程大学学报, 2024, 38(2): 72-87.
- CAI W, GAO W J, JIANG X H, et al. Review of intelligent camouflaged object segmentation algorithms under visible light[J]. Journal of Rocket Force University of Engineering, 2024, 38(2): 72-87.
- [21] HE K M, CHEN X L, XIE S N, et al. Masked autoencoders are scalable vision learners[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 15979-15988.
- [22] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: transformers for image recognition at scale[C/OL]. (2021-06-03)[2025-04-30]. <https://doi.org/10.48550/arXiv.2010.11929>.
- [23] REVAUD J, LEROY V, WEINZAEPFEL P, et al. PUMP: pyramidal and uniqueness matching priors for unsupervised learning of local descriptors[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2022: 3916-3926.
- [24] SARLIN P E, DETONE D, MALISIEWICZ T, et al. SuperGlue: learning feature matching with graph neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2020: 4938-4947.
- [25] 张琪, 范永胜, 金独亮. 基于MMR和WordNet的新闻文本摘要生成研究[J]. 西南师范大学学报(自然科学版), 2023, 48(5): 77-86.
- ZHANG Q, FAN Y S, JIN D L. Research on news text summarizations generation based on MMR and WordNet[J]. Journal of Southwest China Normal University (Natural Science Edition), 2023, 48(5): 77-86.

(编辑: 刘建民)