

MAR20: 遥感图像军用飞机目标识别数据集

禹文奇, 程堪, 王美君, 姚艳清, 谢星星, 姚西文, 韩军伟

西北工业大学 自动化学院, 西安 710129

摘要: 遥感图像军用飞机目标识别是对遥感图像中的军用飞机进行定位和细粒度分类, 其在侦察预警、情报分析等领域起着至关重要的作用。但是, 由于数据集匮乏, 遥感图像军用飞机目标识别发展相对缓慢。为推动该领域的研究进展, 本文构建了公开的遥感图像军用飞机目标识别数据集 MAR20 (Military Aircraft Recognition)。该数据集具有以下特点: (1) MAR20 是目前规模最大的遥感图像军用飞机目标识别数据集, 包含 3842 张图像、20 种军用飞机型号以及 22341 个实例, 并且每个目标实例具有水平边界框和有向边界框两种标注方式; (2) 由于所有的细粒度类别均隶属于飞机大类, 因此不同型号的飞机往往具有相似的特征, 导致不同型号目标具有较高的相似性; (3) 由于遥感图像采集过程中受到气候、季节、光照、遮挡、乃至大气散射等因素的影响, 相同型号的目标存在较大的类内差异性。最后, 为建立遥感图像军用飞机目标识别基准, 本文在 MAR20 数据集上评估了 7 种常用的水平框目标识别方法和 8 种有向框目标识别方法。

关键词: 军用飞机, 目标识别, 数据集, 遥感图像, 细粒度识别

中图分类号: TP701/P2

引用格式: 禹文奇, 程堪, 王美君, 姚艳清, 谢星星, 姚西文, 韩军伟. 2023. MAR20: 遥感图像军用飞机目标识别数据集. 遥感学报, 27(12): 2688-2696

Yu W Q, Cheng G, Wang M J, Yao Y Q, Xie X X, Yao X X and Han J W. 2023. MAR20: A benchmark for military aircraft recognition in remote sensing images. *National Remote Sensing Bulletin*, 27(12): 2688-2696 [DOI: 10.11834/jrs.20222139]

1 引言

随着遥感技术的快速发展, 光学遥感图像数据与日俱增, 如何从海量的遥感图像中自动提取有价值的目标信息成为了近年来学者们研究的热点。而遥感图像目标识别是高效提取目标信息的任务之一, 在民用、军事领域中起着至关重要的作用。该任务对遥感图像中感兴趣的目标进行定位和细粒度分类。军用飞机作为最具威胁性的战略目标之一, 飞机目标出现的位置及型号时刻影响着战场的决策。所以军用飞机目标识别任务具有重要的实际应用价值 (Wu 等, 2020b; 姚艳清等, 2021; 周培诚等, 2021)。

高质量、大规模的数据集对于目标识别任务非常重要。当前有很多包含飞机目标的大规模光

学遥感图像目标检测数据集, 如 DIOR (Li 等, 2020)、NWPU VHR-10 (Cheng 等, 2016)、DOTA (Xia 等, 2018)、LEVIR (Zou 和 Shi, 2018)、RSOD (Long 等, 2017) 等, 但是上述数据集没有提供细粒度的飞机型号信息。RarePlanes (Shermeyer 等, 2020) 增加了大量的飞机细节信息, 包括 5 种特性信息 (飞机机翼、发动机、机身、尾翼和功能)、10 种属性以及 33 种子属性, 但是同样没有提供细粒度型号信息。MTARSI (Wu 等, 2020b) 提供了具体的飞机型号, 但是仅仅是对切片图像进行分类, 缺少检测任务中的定位功能。SIMD (Haroon 等, 2020) 包含多种类型的车辆、飞机和船只, 其中飞机的种类是根据功能或外形来分配细粒度的标签, 而不是根据飞机型号进行分配, 且该数据集包含实例数量较少,

收稿日期: 2022-03-30; 预印本: 2022-07-10

基金项目: 国家自然科学基金 (编号: 61772425); 陕西省杰出青年科学基金 (编号: 2021JC-16)

第一作者简介: 禹文奇, 研究方向为高分辨率遥感图像理解。E-mail: 1884716321@163.com

通信作者简介: 程堪, 研究方向为高分辨率遥感图像理解。E-mail: gcheng@nwpu.edu.cn

MAR20 数据集共享下载链接: <https://gcheng-nwpu.github.io/>

具体类别是:客机、螺旋桨式飞机、教练机、包机和战斗机。FAIR1M (Sun 等, 2022) 数据集为大规模的遥感图像目标识别数据集, 在飞机大类上, 包含有4个波音型号、4个空客型号、C919、ARJ21和其他飞机等共11个飞机型号, 但是该数据集只能用于民航飞机型号识别, 而不能用于军用飞机型号识别。

表1是上述数据集的统计对比信息, 包含数据集的任务类型、型号数和目标实例数。通过表1可以看出, 目前大规模的军用飞机目标识别数据集仍然极度匮乏, 这严重阻碍了军用飞机目标识别任务的研究进展。因此, 构建一个高质量、大规模的军用飞机目标识别数据集具有重要意义。

表1 包含飞机的目标数据集与MAR20对比

Table 1 Comparison of the current datasets containing aircrafts with MAR20

数据集	任务	型号数	目标实例数
DIOR	检测	1	10104
NWPU VHR-10	检测	1	757
DOTA	检测	1	14085
LEVIR	检测	1	4724
RSOD	检测	1	4993
RarePlanes	检测	1	14707
MTARSI	分类	20	9385
SIMD	识别	5	2513
FAIR1M	识别	11	76195
MAR20	识别	20	22341

为推动军用飞机目标识别领域的研究进展, 本文通过Google Earth收集了3842张高分辨率遥感图像数据, 构建了大规模的军用飞机目标识别数据集MAR20 (Military Aircraft Recognition)。最后, 本文在MAR20数据集上评估了7种常用的水平框识别方法和8种有向框识别方法, 为后面的研究建立了军用飞机目标识别基准。

2 MAR20数据集

2.1 数据集的建立

本文通过Google Earth在美国、俄罗斯等国家的60个军用机场采集了3842张高分辨率遥感图像数据, 图像尺寸大部分为800×800像素。本文挑选了俄罗斯空军基地的6种军用飞机的常见型号, 包

括SU-35战斗机、TU-160轰炸机、TU-22轰炸机、TU-95轰炸机、SU-34战斗轰炸机、SU-24战斗轰炸机, 以及美国空军基地的14种常见军用飞机的型号, 包括C-130运输机、C-17运输机、C-5运输机、F-16战斗机、E-3预警机、B-52轰炸机、P-3C反潜机、B-1B轰炸机、E-8战场联合监视机、F-15战斗机、KC-135空中加油机、F-22战斗机、F/A-18战斗攻击机、KC-10空中加油机。本文将SU-35、C-130、C-17、C-5、F-16、TU-160、E-3、B-52、P-3C、B-1B、E-8、TU-22、F-15、KC-135、F-22、FA-18、TU-95、KC-10、SU-34、SU-24飞机型号简称为A1—A20, 飞机型号样本示例如图1所示。同时表2给出了详细的60个机场包含的主要飞机型号。

2.2 数据集的标注

本文使用开源图像标注工具Labelme (Russell 等, 2008) 共标注了22341个目标实例, 并为每个实例提供了水平边界框和有向边界框两种标注方式。具体来说, 水平边界框的标注格式是 (x_1, y_1, x_2, y_2) , 其中, (x_1, y_1) 是水平边界框的左上角坐标, (x_2, y_2) 是水平边界框的右下角坐标; 有向边界框的标注格式是 $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$, 其中, (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) 、 (x_4, y_4) 分别是有向框的4个顶点坐标。这两种标注分别对应于水平框目标识别和有向框目标识别任务, 研究人员可根据任务需求进行相应的研究与开发。

本文定义目标的水平边界框长边为目标大小, 并将目标大小划分为100像素以下、100—150像素、150—200像素和200像素以上的范围。表3给出了MAR20数据集各型号飞机的目标大小分布情况。通过表3可以看出, 该数据集包含的目标尺度具有非常大的多样性, 增加了目标识别任务的难度。

2.3 数据集的划分

根据不同机场包含的各飞机型号的目标数量, 本文将3842张图像划分为训练集和测试集, 其中, 训练集包括1331张图像和7870个目标实例, 测试集包含2511张图像和14471个目标实例。表4给出了MAR20数据集中各型号飞机在训练集和测试集的实例数量分布情况。



图1 MAR20数据集中的20种型号样例

Fig. 1 Samples of 20 aircraft types from the MAR20 dataset

表2 各机场包含的MAR20飞机型号

Table 2 Aircraft types of MAR20 included in each airport

机场名称	飞机型号	机场名称	飞机型号	机场名称	飞机型号	机场名称	飞机型号
布图利诺夫卡区	SU-35,SU-34	恩格斯	TU-160,TU-95	多弗	F-22	戴斯	B-1B
乌克兰卡	TU-95	麦克德菲尔德	C-17	兰利	C-5	瓦丁顿	E-3,E-8
共青城	SU-35,SU-34	费尔柴尔德	KC-135	麦克迪尔	KC-135	嘉手纳	KC-135,F-22
赫梅明	SU-35,SU-34,SU-24	芒廷霍姆	F-15	埃格林	F-15,F-16,F-22,C-130	托雷洪	FA-18
哈巴罗夫斯克	SU-34,SU-24	麦奎尔	KC-10	小岩城	C-130	谢泼德	F-15,F-16,C-130
阿里克谢耶夫卡	TU-95,TU-160	爱德华兹	F-16	泰德尔	F-22	巴克利	F-16
利佩茨克	SU-34,SU-35,SU-24	比尔	KC-135,C-5	埃尔斯沃思	B-1B	沃斯堡	C-130,FA-18
梁赞	TU-95,TU-22	希尔	F-22	罗宾斯	E-8,C-5,C-130	内利斯	F-15,F-16,F-22
基洛夫区	TU-22	卢克	F-16	廷克	E-3,B-52,B-1B	厚木	P-3C
喀山	TU-22	戴维斯蒙森	美国大多数型号	埃尔蒙多夫	E-8,E-3,F-15,F-16,F-22	查尔斯顿	C-17
阿赫图宾斯克	SU-35,SU-34	拉克兰	C-5	奥特菲尔德	F/A-18	阿尔特斯	KC-135,C-17
德米特里耶夫卡	SU-24	欧福特	E-8	勒莫尔	F/A-18	特拉维斯	C-5,KC-10
卡赞采沃	SU-24	斯科特	KC-135,C-130,C-17	木古角	P-3C	凯利	C-5
尼科利斯科耶	SU-35,SU-34	巴克斯戴尔	B-52	安德森	KC-135,B-52	斯图尔特	C-17,C-5
伊尔门湖	TU-22	安德鲁斯	KC-135	西卡姆	KC-135,C-17	韦斯特沃	C-5

表3 MAR20各型号飞机的目标大小分布情况
Table 3 Distribution of object size for each type of aircrafts on the MAR20 dataset

型号	100像素以下	100—150像素	150—200像素	200像素以上
A1	759	759	110	18
A2	648	774	255	52
A3	608	448	86	34
A4	101	305	135	101
A5	736	441	78	7
A6	352	80	8	1
A7	204	260	193	23
A8	347	362	182	53
A9	407	463	168	48
A10	292	532	80	20
A11	152	199	117	39
A12	281	347	66	8
A13	843	551	206	52
A14	550	825	328	75
A15	396	189	32	1
A16	1220	1270	138	4
A17	584	542	230	41
A18	184	82	32	10
A19	721	417	86	12
A20	587	357	37	0

尽管本文使用 Google Earth 获得了不同气候、不同季节和不同光照的遥感影像,但是同一机场同一型号的军用飞机仍然存在相似的空间信息和背景信息。若同一机场的飞机被分配到训练集和测试集中,将会严重影响目标识别算法的泛化性

表4 MAR20数据集各型号飞机在训练集和测试集的实例数量分布

Table 4 Distribution of the aircraft numbers of each type in the training set and test set of the MAR20 dataset

型号	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20
训练集实例数	554	736	316	179	435	107	191	365	475	368	151	240	561	659	200	785	439	160	519	430
测试集实例数	1092	993	860	463	827	334	489	579	611	556	356	462	1091	1119	418	1847	958	148	717	551
总和	1646	1729	1176	642	1262	441	680	944	1086	924	507	702	1652	1778	618	2632	1397	308	1236	981

3 MAR20数据集上的算法评估

3.1 评价指标

本文采用与 PASCAL VOC 2007 (Everingham 等, 2010) 相同的评价指标, 即平均精度均值 mAP (mean Average Precision), 它是指不同型号目标的平均精度 AP (Average Precision) 的均值, 而平均精度定义为 precision-recall 曲线下的面积。

能评估。所以在划分数据集时, 本文首先随机筛选机场, 再将图像数据分配到训练集和测试集中。

另外值得说明的是, 受功能用途、造价等因素的影响, 个别型号飞机的生产和保有数量较少, 再加上一些典型目标自身的敏感性, 导致获取到的目标样本数据存在一定程度的样本不平衡问题, 如 A6、A11、A12、A15、A18 型号的飞机实例数较其他型号较少。

2.4 数据集特点

MAR20 数据集具有以下 3 个特点: (1) 大规模: MAR20 是目前公开的规模最大的遥感图像军用飞机目标识别数据集, 包含 3842 张图像、20 种军用飞机型号以及 22341 个实例。此外, 该数据集提供了水平边界框和有向边界框两种标注方式, 分别对应水平框目标识别任务与有向框目标识别任务, 方便研究人员进行不同算法的性能。(2) 不同型号目标类间差异小: 由于这 20 个细粒度类别均属于同一个飞机大类, 因此不同型号的飞机往往具有相似的特征, 这导致细粒度类别之间具有较高的相似性。不同飞机型号的区别主要体现在局部部件的不同, 例如, SU-35 和 SU-34 飞机型号的差别是头部的前鸭翼。(3) 相同型号目标类内差异大: 由于遥感图像采集过程中受到气候、季节、光照、遮挡、乃至大气散射等因素的影响, 相同型号的目标往往视觉差异大, 存在较高的类内差异性。

precision 和 recall 的计算如式 (1) 和式 (2) 所示:

$$\text{precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (2)$$

式中, TP 代表的是真正例, FN 代表的是假负例, FP 代表的是假正例。

3.2 实验设置

本文使用一张显存为 12 G 的 NVIDIA Titan X

GPU 进行训练和测试。以外, 本文使用的 CPU 为 Intel Xeon Silver 4114 2.20-GHz。在水平框目标识别和有向框目标识别任务中, 本文分别使用开源目标检测库 MMDetection (Chen 等, 2019) 和 OBBDetection (Xie 等, 2021) 作为算法平台进行实验。

3.2.1 水平框方法实验设置

在水平框目标识别任务中, 本文选取了 7 种常用方法, 包括 Faster R-CNN (Ren 等, 2017)、RetinaNet (Lin 等, 2017b)、ATSS (Zhang 等, 2020)、FCOS (Tian 等, 2019)、Cascade R-CNN (Cai 和 Vasconcelos, 2018)、TSD (Song 等, 2020) 和 Double-Head (Wu 等, 2020a)。

上述 7 种基准方法均使用 ResNet-50 (He 等, 2016) 作为骨干网络, 并采用特征金字塔网络 (Lin 等, 2017a) 结构。同时, 骨干网络加载 ImageNet (Russakovsky 等, 2015) 预训练模型, 实现参数初始化。参数设置上, 本文选择随机梯度下降 SGD (Stochastic Gradient Descent) 作为网络优化器, 并设置动量参数为 0.9, 衰减系数为 0.0001, batch-size 为 2。因为本文采用了多种类型的基准方法, 所以对不同的方法设置了不同的迭代数和学习率。对于 Faster R-CNN、Cascade R-CNN、TSD 和 Double-Head, 本文设置初始学习率为 0.0025, 迭代数为 12, 分别在第 8 次迭代和第 11 次迭代降低学习率; 对于 ATSS, 设置初始学习率为 0.00125, 迭代数、降低学习率的迭代次数与 Faster R-CNN 相同; 对于 RetinaNet 和 FCOS, 本文设置初始学习率为 0.00125, 迭代数为 36, 分别在第 24 次迭代和第 33 次迭代降低学习率。

3.2.2 有向框方法实验设置

有向框目标识别任务选取 Faster R-CNN-O (Ren 等, 2017)、RetinaNet-O (Lin 等, 2017b)、RoI Transformer (Ding 等, 2019)、Gliding Vertex (Xu 等, 2021)、Double-Head-O (Wu 等, 2020a)、Oriented R-CNN (Xie 等, 2021)、FCOS-O (Tian 等, 2019) 和 S²A-Net (Han 等, 2022) 作为本文的基准方法。其中, Faster R-CNN-O、RetinaNet-O、Double-Head-O 和 FCOS-O 分别通过在原有检测头上增加了角度预测分支实现有向目标识别。在特征提取骨干网络、预训练模型参数初始化、

特征金字塔网络、网络优化器和 batch-size 上, 本文采用与水平框目标识别相同的参数。对于 Faster R-CNN-O、RoI Transformer、Gliding Vertex、Double-Head-O、Oriented R-CNN 和 S²A-Net, 本文设置初始学习率为 0.005, 迭代数为 36, 分别在第 24 次迭代和第 33 次迭代降低学习率; 对于 RetinaNet-O 和 FCOS-O, 本文设置初始学习率为 0.0025, 迭代数与其他方法相同。

3.3 实验结果

3.3.1 基准方法实验结果及分析

表 5 和表 6 分别给出了 7 种水平框目标识别方法和 8 种有向框目标识别方法在 MAR20 数据集上的识别结果。在水平框目标识别任务中, 效果较好的是 Faster R-CNN、Cascade R-CNN 和 TSD; 在有向框识别任务中, 效果较好的是 Faster R-CNN-O、RoI Transformer 和 Oriented R-CNN。此外, 图 2 和图 3 分别给出了 Faster R-CNN 和 Faster R-CNN-O 在 MAR20 数据集上的部分识别结果。

通过水平框与有向框目标识别任务的实验对比可知, 双阶段方法在目标识别任务中更加有效。另外, 针对分类和回归任务特征需求的不同, TSD 实现了目标空间位置的解耦, 通过为两任务筛选适合的特征可有效提升目标识别性能。除此以外, 通过级联多个检测头并逐级提升样本质量也可显著提升目标识别精度。通过分析上述方法可以看出, 如何获得更具判别性的分类特征是实现精细识别的核心。

此外, 本文以水平框目标识别任务中 Faster R-CNN 的实验数据为例, 进一步分析实验结果。通过表 5 的实验数据可以看出, A8、A9、A10、A17 型号的 AP 值接近甚至超过了 90%。虽然上述型号的实例数量不是最多的, 但是它们的目标特征更具判别性, 使得网络更容易区分。而 A13、A15 和 A18 型号识别效果较差, 3 种型号的 AP 值都低于 70%。导致这一结果的主要原因如下: (1) A13 和 A15 型号属于战斗机, 与其他的战斗机型 A5 和 A16 有着较高的类间相似度, 使得网络提取的特征判别性较弱, 造成了上述两种型号的目标识别精度较低; (2) A18 的实例数量远低于其他型号目标的实例数量, 且该型号容易与 A4 型号混淆, 所以该型号的识别精度较低。

表 5 水平框飞机目标识别结果

Table 5 Aircraft recognition results with horizontal bounding boxes

型号	单阶段			多阶段			
	RetinaNet	FCOS	ATSS	Faster R-CNN	TSD	Cascade R-CNN	Double-Head
A1	70.7	71.1	69.9	79.7	87.2	79.6	82.3
A2	82.1	79.4	83.7	81.6	86.7	81.4	81.2
A3	45.7	67.1	70.3	84.7	93.9	82.2	79.0
A4	46.1	46.4	50.2	72.0	71.8	66.7	63.4
A5	70.4	65.1	59.1	75.0	75.7	76.5	76.8
A6	76.7	84.8	85.1	88.0	93.6	87.9	87.9
A7	50.7	72.1	82.6	87.1	96.5	88.7	85.5
A8	81.2	83.3	80.0	89.7	91.3	86.9	83.6
A9	50.8	80.5	84.8	89.6	89.2	89.2	85.0
A10	88.9	84.5	90.0	90.6	98.1	90.7	90.5
A11	36.2	57.8	61.4	82.5	87.3	81.8	79.9
A12	77.0	71.3	82.7	86.4	87.8	85.1	85.4
A13	62.7	58.5	68.1	66.6	67.1	66.5	66.0
A14	71.8	75.3	78.1	85.8	89.6	78.6	84.6
A15	25.6	33.8	40.5	45.8	34.5	42.7	42.2
A16	85.0	83.7	87.0	87.8	91.0	88.2	88.4
A17	85.5	87.1	89.5	90.2	95.5	90.1	89.8
A18	27.9	38.6	36.1	54.2	58.5	42.6	44.1
A19	68.1	69.2	67.7	76.6	79.6	78.3	74.4
A20	77.7	66.4	77.5	78.6	80.1	76.7	77.2
mAP	64.0	68.8	72.2	79.6	82.7	78.0	77.4

注: 表中加粗数值表示最优结果。

表 6 有向框飞机目标识别结果

Table 6 Aircraft recognition results with oriented bounding boxes

型号	单阶段			多阶段				
	FCOS-O	RetinaNet-O	S ² A-Net	Faster R-CNN-O	Double-Head-O	Oriented R-CNN	RoI Transformer	Gliding Vertex
A1	68.50	79.04	82.62	85.01	86.35	86.05	85.40	85.85
A2	79.70	84.31	81.59	81.63	80.85	81.73	81.53	81.53
A3	61.00	71.05	86.21	87.47	88.90	88.08	87.61	86.80
A4	52.34	54.72	80.75	70.68	82.54	69.57	78.33	76.35
A5	64.00	73.18	76.86	79.63	76.04	75.61	80.45	72.22
A6	83.30	86.59	90.00	90.58	90.06	89.92	90.49	89.90
A7	72.90	75.57	84.73	89.71	89.76	90.49	90.24	89.84
A8	82.28	85.51	85.70	89.82	87.28	89.54	87.58	89.38
A9	81.11	88.65	88.70	90.40	89.20	89.78	87.93	89.14
A10	84.55	85.84	90.84	90.89	90.78	90.91	90.89	90.77
A11	67.70	68.20	81.67	85.54	84.35	87.62	85.88	86.20
A12	78.77	73.22	86.09	88.08	86.18	88.39	89.29	87.45
A13	60.86	63.51	69.59	68.39	65.76	67.52	67.24	64.94
A14	81.26	79.72	85.25	88.27	87.42	88.50	88.20	88.28
A15	32.82	24.10	47.69	42.44	44.13	46.33	47.85	47.01
A16	81.84	84.85	88.10	88.86	87.49	88.27	89.11	87.84
A17	90.60	90.32	90.20	90.45	90.25	90.59	90.46	90.40
A18	51.06	49.19	61.98	62.23	56.40	70.50	74.59	64.94
A19	68.12	74.96	83.59	78.25	82.82	78.72	81.30	83.90
A20	71.10	76.07	79.84	77.71	76.69	80.25	80.00	76.83
mAP	70.69	73.43	81.10	81.35	81.16	81.92	82.72	81.48

注: 表中加粗数值表示最优结果。



图2 水平框飞机目标识别结果可视化(绿框代表正确识别结果,红框代表漏检,蓝框代表虚警)
Fig. 2 The visualization of aircraft recognition results with horizontal boxes (The green boxes represent true positives , red boxes represent false negatives and blue boxes represent false positives)



图3 有向框飞机目标识别结果可视化(绿框代表正确识别结果,红框代表漏检,蓝框代表虚警)
Fig. 3 The visualization of aircraft recognition results with oriented boxes (The green boxes represent true positives , red boxes represent false negatives and blue boxes represent false positives)

4 结 论

遥感图像军用飞机目标识别在侦察预警等领域具有重要的应用作用。但受限于缺乏大规模的军用飞机目标识别数据集,该领域研究进展相对缓慢。为此,本文通过 Google Earth 在全球 60 个军用机场收集了 3842 张高分辨率遥感图像数据,建

立了大规模公开的遥感图像军用飞机目标识别数据集 MAR20, 包含了 20 种飞机型号和 22341 个目标实例。在数据标注方面, MAR20 提供了水平边界框和有向边界框两种标注方式,分别对应于水平框与有向框目标识别任务,方便研究人员根据任务需求灵活使用。

在 MAR20 数据集上,本文评估了 7 种水平框

目标识别方法和8种有向框目标识别方法。最后,我们希望本文建立的公开数据集能够促进光学遥感图像军用飞机目标识别领域的研究进展。

参考文献(References)

- Cai Z W and Vasconcelos N. 2018. Cascade R-CNN: delving into high quality object detection//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 6154-6162 [DOI: 10.1109/cvpr.2018.00644]
- Chen K, Wang J Q, Pang J M, Cao Y H, Xiong Y, Li X X, Sun S Y, Feng W S, Liu Z W, Xu J R, Zhang Z, Cheng D Z, Zhu C C, Cheng T H, Zhao Q J, Li B Y, Lu X, Zhu R, Wu Y, Dai J F, Wang J D, Shi J P, Ouyang W L, Loy C C and Lin D H. 2019. MMDetection: open MMLab detection toolbox and benchmark. arXiv preprint arXiv: 1906.07155
- Cheng G, Zhou P C and Han J W. 2016. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(12): 7405-7415 [DOI: 10.1109/TGRS.2016.2601622]
- Ding J, Xue N, Long Y, Xia G S and Lu Q K. 2019. Learning RoI transformer for oriented object detection in aerial images//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE: 2844-2853 [DOI: 10.1109/CVPR.2019.00296]
- Everingham M, Van Gool L, Williams C K I, Winn J and Zisserman A. 2010. The Pascal visual object classes (VOC) challenge. *International Journal of Computer Vision*, 88(2): 303-338 [DOI: 10.1007/s11263-009-0275-4]
- Han J M, Ding J, Li J and Xia G S. 2022. Align deep features for oriented object detection. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5602511 [DOI: 10.1109/TGRS.2021.3062048]
- Haroon M, Shahzad M and Fraz M M. 2020. Multisized object detection using spaceborne optical imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 3032-3046 [DOI: 10.1109/JSTARS.2020.3000317]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE: 770-778 [DOI: 10.1109/cvpr.2016.90]
- Li K, Wan G, Cheng G, Meng L Q and Han J W. 2020. Object detection in optical remote sensing images: a survey and a new benchmark. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159: 296-307 [DOI: 10.1016/j.isprsjprs.2019.11.023]
- Lin T Y, Dollár P, Girshick R, He K M, Hariharan B and Belongie S. 2017a. Feature pyramid networks for object detection//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE: 936-944 [DOI: 10.1109/cvpr.2017.106]
- Lin T Y, Goyal P, Girshick R, He K M and Dollár P. 2017b. Focal loss for dense object detection//Proceedings of the 2017 IEEE International Conference on Computer Vision. Venice: IEEE: 2999-3007 [DOI: 10.1109/iccv.2017.324]
- Long Y, Gong Y P, Xiao Z F and Liu Q. 2017. Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 55(5): 2486-2498 [DOI: 10.1109/tgrs.2016.2645610]
- Ren S Q, He K M, Girshick R and Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1137-1149 [DOI: 10.1109/tpami.2016.2577031]
- Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S A, Huang Z H, Karpathy A, Khosla A, Bernstein M, Berg A C and Fei-Fei L. 2015. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3): 211-252 [DOI: 10.1007/s11263-015-0816-y]
- Russell B C, Torralba A, Murphy K P and Freeman W T. 2008. LabelMe: a database and web-based tool for image annotation. *International Journal of Computer Vision*, 77(1/3): 157-173 [DOI: 10.1007/s11263-007-0090-8]
- Shermeyer J, Hossler T, Van Etten A, Hogan D, Lewis R and Kim D. 2020. RarePlanes: synthetic data takes flight. arXiv preprint arXiv: 2006.02963
- Song G L, Liu Y and Wang X G. 2020. Revisiting the sibling head in object detector//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 11560-11569 [DOI: 10.1109/cvpr42600.2020.01158]
- Sun X, Wang P J, Yan Z Y, Xu F, Wang R P, Diao W H, Chen J, Li J H, Feng Y C, Xu T, Weinmann M, Hinz S, Wang C and Fu K. 2022. FAIRIM: a benchmark dataset for fine-grained object recognition in high-resolution remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 184: 116-130 [DOI: 10.1016/j.isprsjprs.2021.12.004]
- Tian Z, Shen C H, Chen H and He T. 2019. FCOS: fully convolutional one-stage object detection//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE: 9626-9635 [DOI: 10.1109/iccv.2019.00972]
- Wu Y, Chen Y P, Yuan L, Liu Z C, Wang L J, Li H Z and Fu Y. 2020a. Rethinking classification and localization for object detection//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 10183-10192 [DOI: 10.1109/cvpr42600.2020.01020]
- Wu Z Z, Wan S H, Wang X F, Tan M, Zou L, Li X L and Chen Y. 2020b. A benchmark data set for aircraft type recognition from remote sensing images. *Applied Soft Computing*, 89: 106132 [DOI: 10.1016/j.asoc.2020.106132]
- Xia G S, Bai X, Ding J, Zhu Z, Belongie S, Luo J B, Dateu M, Pelillo M and Zhang L P. 2018. DOTA: a large-scale dataset for object detection in aerial images//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 3974-3983 [DOI: 10.1109/CVPR.2018.00418]
- Xie X X, Cheng G, Wang J B, Yao X W and Han J W. 2021. Oriented

- R-CNN for object detection//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE: 3500-3509 [DOI: 10.1109/ICCV48922.2021.00350]
- Xu Y C, Fu M T, Wang Q M, Wang Y K, Chen K, Xia G S and Bai X. 2021. Gliding vertex on the horizontal bounding box for multi-oriented object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(4): 1452-1459 [DOI: 10.1109/TPAMI.2020.2974745]
- Yao Y Q, Cheng G, Xie X X and Han J W. 2021. Optical remote sensing image object detection based on multi-resolution feature fusion. *Journal of Remote Sensing*, 25(5): 1124-1137 (姚艳清, 程焱, 谢星星, 韩军伟. 2021. 多分辨率特征融合的光学遥感图像目标检测. *遥感学报*, 25(5): 1124-1137) [DOI: 10.11834/jrs.20210505]
- Zhang S F, Chi C, Yao Y Q, Lei Z and Li S Z. 2020. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE: 9756-9765 [DOI: 10.1109/cvpr42600.2020.00978]
- Zhou P C, Cheng G, Yao X W and Han J W. 2021. Machine learning paradigms in high-resolution remote sensing image interpretation. *National Remote Sensing Bulletin*, 25(1): 182-197 (周培诚, 程焱, 姚西文, 韩军伟. 2021. 高分辨率遥感影像解译中的机器学习范式. *遥感学报*, 25(1): 182-197) [DOI: 10.11834/jrs.20210164]
- Zou Z X and Shi Z W. 2018. Random access memories: a new paradigm for target detection in high resolution aerial remote sensing images. *IEEE Transactions on Image Processing*, 27(3): 1100-1111 [DOI: 10.1109/TIP.2017.2773199]

MAR20: A benchmark for military aircraft recognition in remote sensing images

YU Wenqi, CHENG Gong, WANG Meijun, YAO Yanqing, XIE Xingxing, YAO Xiwen, HAN Junwei

School of Automation, Northwestern Polytechnical University, Xi'an 710129, China

Abstract: Military aircraft recognition in remote sensing images locates military aircraft in remote sensing images and classify them at a fine-grained level. It plays a vital role in reconnaissance and early warning, intelligence analysis, and other fields. However, the development of military aircraft recognition in remote sensing images is relatively slow due to the lack of publicly available datasets. Therefore, constructing a high-quality and large-scale military aircraft recognition dataset is important.

This study constructs a public remote sensing image military aircraft recognition dataset called MAR20 to promote the research progress in this field. The dataset has the following characteristics: (1) MAR20 is currently the largest remote sensing image military aircraft recognition dataset, which includes 3842 images, 20 types, and 22341 instances. Each instance has a horizontal bounding box and also an oriented bounding box. (2) Given that all fine-grained types belong to the aircraft category, different types of aircraft often have similar characteristics, which result in high similarity of different types of targets. (3) Large intra-class differences exist between targets of the same type due to the influence of climate, season, illumination, occlusion, and even the atmospheric scattering in the process of remote sensing imaging.

To establish a benchmark for military aircraft recognition in remote sensing images, this paper study evaluates seven commonly used horizontal object recognition methods, namely, Faster R-CNN, RetinaNet, ATSS, FCOS, Cascade R-CNN, TSD, and Double-Head, as well as eight oriented object recognition methods, namely, Faster R-CNN-O, RetinaNet-O, RoI Transformer, Gliding Vertex, Double-Head-O, Oriented R-CNN, FCOS-O, and S2A-Net, on the MAR20 dataset. Through experimental comparisons in the tasks of horizontal object recognition and oriented object recognition, two-stage methods are proven to be more effective in target recognition than one-stage methods.

In this study, 3842 high-resolution remote sensing images were collected from 60 military airports around the world through Google Earth, and a large-scale publicly available remote sensing image military aircraft recognition dataset, named MAR20, was established. In terms of data annotation, MAR20 provides two annotation methods, namely, horizontal bounding boxes and oriented bounding boxes, which correspond to the tasks of horizontal target recognition and oriented target recognition. We hope that the MAR20 dataset established in this study could promote the research progress in this field. MAR20 can be downloaded at <https://gcheng-nwpu.github.io/>.

Key words: military aircraft, object recognition, dataset, remote sensing images, fine-grained recognition

Supported by National Natural Science Foundation of China (No. 61772425); Shaanxi Science Foundation for Distinguished Young Scholars (No. 2021JC-16)