

DOI:

# 联合目标分割和关键点检测的飞机型号识别方法

刘思婷<sup>1,2</sup>, 王庆栋<sup>1</sup>, 张力<sup>1</sup>, 韩晓霞<sup>1</sup>, 王保前<sup>1</sup>, 刘玉贤<sup>3</sup>

1. 中国测绘科学研究院, 北京 100830;

2. 兰州交通大学 测绘与地理信息学院, 甘肃 兰州 730070;

3. 深圳市勘察研究院有限公司, 深圳 518026;

**摘要:** 目前, 受限于数据集精细度与网络结构, 深度学习技术仍难以应对飞机目标型号识别这类精细化识别任务。本文针对遥感影像中飞机目标型号识别问题, 提出一种融合目标分割与关键点检测的飞机型号识别方法。该方法有机地结合多任务深度神经网络与条件随机场和模板匹配算法, 利用“预训练+微调+后处理”的方式实现飞机型号的高精度识别。首先, 基于多任务深度神经网络迁移学习技术实现飞机目标物位置、掩膜与关键点信息识别。其次, 为了便于后期高精度模板匹配, 利用本文提出的融合条件随机场的飞机目标掩膜精化算法和基于关键点的姿态调整算法, 实现识别目标的边界精细化与机体姿态调整; 最后, 在本文构建的飞机型号模板库基础上, 将经过精化后处理的飞机掩膜信息与模板库进行匹配, 实现飞机目标的型号识别。为了验证所提方法的有效性, 本文进行了相关实验, 并与传统算法及完全端到端深度学习方法进行了对比, 结果表明, 本文所提方法具有更高准确率, 并且在实用性方面更具优势。

**关键词:** 目标检测; 分割; 关键点检测; 条件随机场; 飞机型号识别;

**中图分类号:**    **文献标识码:**    **文章编号:**

**引用格式:** 刘思婷,王庆栋,张力,韩晓霞,王保前,刘玉贤. 联合目标分割和关键点检测的飞机型号识别方法. 遥感学报, Liu Siting, Wang Qingdong, Zhang Li, Han Xiaoxia, Wang Baoqian and Liu Yuxian. 2022. Aircraft type recognition method by integrating target segmentation and key points detection. National Remote Sensing Bulletin, , , ( ): 00-00

## 1、 引言

飞机目标检测一直是高分辨率遥感影像解译工作中的一个重要任务。受到深度学习技术在目标检测领域的影响, 基于人工设计特征的传统检测方法逐渐被可自动学习目标特征的深度神经网络的方法所替代, 一些经典工作的提出, 如RCNN, Fast-RCNN, YOLO, SSD, 使目标检测在精度方面得到进一步提升(姚群力等, 2019)。近些年, 随着目标检测技术逐渐实用化以及高分辨遥感影像质量的提升, 针对目标物的细粒度检测成为新的研究热点, 而目前针对飞机目标的精细识别, 如飞机目标的姿态检测、型号识别等方面的相关研究较少, 并且由于场景、纹理、光照等复杂因素, 使得这项

任务非常具有挑战性。

目前, 针对遥感影像的飞机目标检测相关研究已有多种类型的算法被提出和使用。研究热点主要集中在候选区域选择和特征提取方法等领域。候选区域选择是指初步提取影像上可能包含飞机目标的区域, 经典的方法有基于滑动窗口的搜索方法(Zhang等, 2014), 基于图像分割的方法(王威等, 2013), 基于图像显著性的方法(朱丹等, 2015)等。特征提取即对候选区域提取相应的特征, 用于飞机目标的确认。特征提取的关键在于提取何种特征, 传统的方法多利用一些通用的低层次特征, 如颜色、形状、纹理等全局特征, 以及SIFT、HOG等局部特征, 这些人工设计的特征需要相关的先验知识, 检测效果不稳定。另外还有一些根据飞机目标的具体

**收稿日期:**        **修回日期:**

**基金项目:** 1、国家重点研发计划资助(编号: 2019YFB1405600); 2、深圳市技术攻关项目(编号: JSGG20191129103003903);

**第一作者简介:** 刘思婷, 1996年生, 女, 硕士研究生, 研究方向为遥感影像智能解译。E-mail:0219760@stu.lzjtu.edu.cn

**通信作者简介:** 王庆栋, 1986年生, 男, 助理研究员, 研究方向为基于点云深度学习的室内外语义三维重建。E-mail:wangqd@casm.ac.cn

特性来设计的相应特征如具有代表性的Hu矩(Wahi等, 2012)、Zernike矩(Mahi等, 2014)等特征, 这类方法在复杂场景中对目标细节描述能力不足, 造成检测精度低。

近些年, 深度学习方法发展迅速, 利用卷积神经网络自动学习图像特征进行目标检测方法取得了不错的效果。其中以检测精度更高的二阶段(Two-stage)检测算法(Ren等, 2017)以及在速度上更快的一阶段(One-stage)检测算法为主。二阶段算法将候选区的选择与卷积神经网络相结合, 而一阶段算法是利用CNN直接预测目标类别与位置的, 二者在检测精度和速度各有优势。以此为基础, 许多算法进行了适应各自任务的改进, 在一些相应的应用场景中提升了检测精度。这些方法广泛地应用在卫星影像、航空影像、SAR影像等多源遥感数据的目标检测任务中。由于卫星影像分辨率各异, 覆盖面广且地物类型复杂, 使用深度神经网络进行目标检测时容易出现小目标的遗漏, “多尺度特征融合”的相关方法(Yang等, 2021; 王冰等, 2021; 姚群力等, 2019)被广泛的提出和采用以实现针对多尺度目标的检测; 航空影像的分辨率更高, 适用于特定场景中的相关目标比如汽车等的检测(Zhang Ruiqian, et al.2021); 由于SAR不受云雨的影响可以获取全天时的影像, 同样是目标检测领域重要的研究对象。但由于其获取的是散射信号, 并且噪声比较大, 图像的可解释性较弱, 使得其目标检测困难。针对此, 相关研究中通过特征融合、散射特性增强等方法(J. Ai, et al.2021; Guo, et al.2019; S. Chen, et al.2016)实现高精度的目标检测。

飞机目标的型号识别是建立在飞机目标检测基础上的任务, 相比于上述的目标检测是一种更加精细化的分类任务, 在机场场景理解、目标解译以及军事情报获取等领域存在大量需求(张凯等, 2020; 苗壮等, 2020), 但目前的研究相对较少。早期李科等人提出了层次化型号识别方法(李科等, 2006), 其原理是依据一定的准则选用一个或几个特征实施逐级分类; 近些年的飞机型号识别多集中于模板匹配或直接通过分类器识别。张名成等人(张名成等, 2006)提出了基于闭合轮廓和部分特征匹配的型号识别方法; Wu Qichang等人依据飞机的边界形状特征与标准模板进行匹配实现了7种型号飞机的识别(Wu等, 2015); 颜荔通过自制训练数据集, 采用基于卷积神经网络弱监督学习提取特征, 再由分类器完成飞机型号识别(颜荔, 2018); 目前深度学习方

法主要依据大量标注数据进行网络训练, 通过调整网络参数得到好的模型, 从而实现端到端的飞机型号识别; 然而现实应用场景中, 精细标注的样本数据相对匮乏、标注繁琐、工作量大, 这在很大程度上制约了端到端网络识别飞机类型的研究。Sun Xian等人提出了一种新的大型基准数据集FAIR1M(Sun等, 2021), 其针对5个大类别和37个子类别进行注释, 并且针对遥感图像中目标物体, 提供了丰富的细粒度类别信息, 其中就包含10种类型的飞机目标。对于飞机型号识别这项任务来说, 该数据集所训练的网络在检测类别的数量上有所限制。为了缓解数据集不足的问题, 深度学习与传统算法相结合的方法逐渐被采用。Fu Kun等人(Fu等, 2019)引入了细粒度视觉分类(FGVC)的思想, 提出利用多类激活映射(MultiCAM)来提取不同类型飞机的判别方法; Zuo Jiawei等人(Zuo等, 2018)提出了一种结合分割网络与关键点检测的新型飞机类型识别方法, 利用神经网络获取分割结果, 结合关键点检测对齐分割结果并采用模板匹配的方法来识别飞机的具体型号。

总体来看, 现有飞机型号识别研究主要存在以下问题: 1) 基于遥感影像的飞机型号识别率低。遥感影像视角单一, 飞机外观相似度较高, 导致高分辨率遥感影像飞机型号识别时区分困难。2) 现有飞机目标检测结果并不能很好服务于更精细的飞机型号检测任务, 如何充分利用并优化目标检测结果实现准确的飞机型号识别仍有待研究; 3) 相关特征提取中, 依赖多种神经网络, 模型训练与调优难度高。针对以上问题, 本文提出一种融合随机条件场CRF与机体关键点检测的飞机型号识别方法。该方法在利用深度神经网络Mask R-CNN(He等, 2020)迁移学习获取的目标掩膜的基础上, 对其边界信息进行精化, 并利用机体关键点信息进行模板匹配, 实现精确的飞机型号识别。为了构建不同型号飞机模板库, 本文选择公开可用的MTARSI数据集(Wu等, 2020)制作各类型号飞机的标准模板。本文主要贡献有以下几点: (1)提出融合关键点检测与目标检测的飞机目标高精度检测框架, 实现飞机目标检测与型号识别; (2)基于迁移学习提出利用面向飞机目标物的边界检测与关键点检测方法; (3)在此基础上, 提出融合条件随机场与目标特征的模板匹配技术, 优化飞机目标边界与姿态, 实现飞机型号匹配精度的有效提升。为了验证该方法的有效性与实用性, 本文利用RSOD数据集和自建的模板库进行

了多组实验，并与现有相关算法进行比较和分析，实验表明，与现有方法相比本文方法在11种飞机类型的识别在性能与精度方面都有较显著提升。并且与目前完全端到端的型号识别深度神经网络方法相比，本文所提方法在稳健性，实用性等方面更具优势。

## 2 基于目标分割的型号识别算法

飞机型号识别作为高精细度的目标检测任务，不仅涉及到飞机目标检测，还涉及到难度较大的飞

机目标分割与关键点的检测，因此，本文围绕这些精细检测任务设计了相应的飞机型号识别框架。该框架主要由两大模块组成：1) 多任务飞机目标快速检测模块，涉及到飞机目标检测，飞机轮廓特征提取与目标物关键点特征提取等；2) 融合条件随机场(CRF)与目标特征的模板匹配模块，其基于分割掩膜与关键点特征，利用CRF算法对目标边界与姿态进行优化，实现检测目标与样本库的高精度模板匹配，最终实现飞机目标型号的准确识别。飞机型号识别框架具体如图1所示。

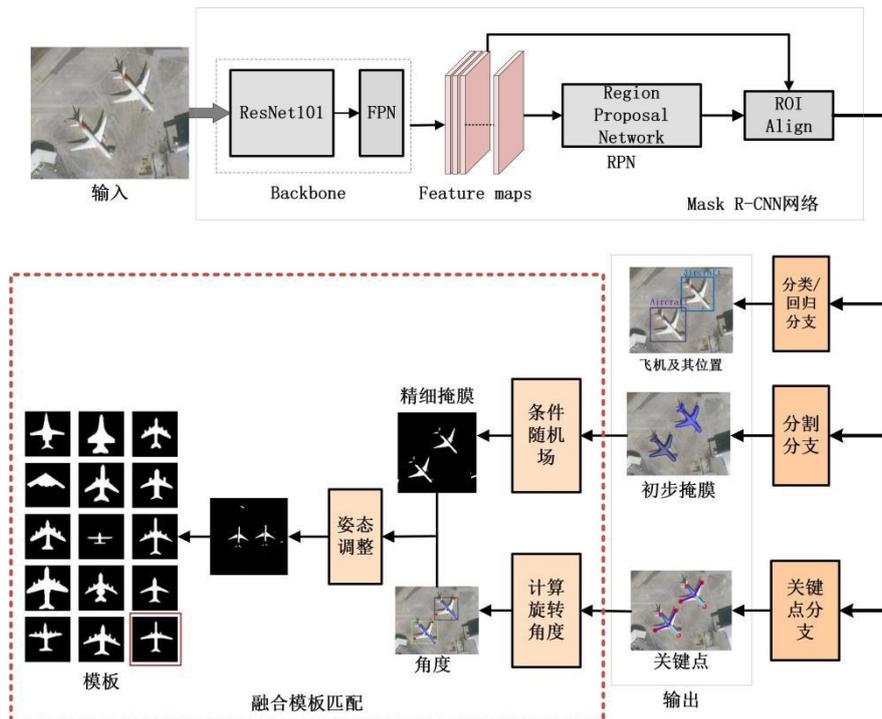


图 1.本文飞机型号识别框架

Figure 1. The aircraft type recognition framework

### 2.1 基于多任务学习的飞机目标快速检测

目前，针对飞机目标检测任务的相关研究较多，但大多数研究不涉及进一步高层次特征识别，因此，相关研究的检测任务相对单一，获得的特征信息不够丰富。为了配合本文后续的飞机型号识别高精细度检测任务，本文采用多任务的Mask R-CNN网络，在进行飞机目标检测的同时，实现飞机目标的分割与关键点检测，为飞机型号识别提供丰富的边界轮廓信息与关键点信息。

#### 2.1.1 基于 Mask R-CNN 迁移学习的飞机目标检测与分割

Mask R-CNN是在Faster R-CNN(Ren等, 2017)的基础上提出的目标检测网络，其结构如图2所示。它主要有三部分：残差网络组成的主干特征提取网络(Backbone)，用于候选框生成与筛选的区域建议

网络RPN(Region proposal network)，以及最后用于检测、分割等任务的ROI分类器(Classifier)。

主干特征提取网络(Backbone)采用的是残差网络模块，这主要是因为随着网络深度增加，网络性能会受到梯度消失效应的影响，导致前面网络层的权重无法得到有效更新，出现网络退化问题。深度残差网络(He等, 2016)可以一定程度上解决网络退化问题，它由一系列残差块组成，通过残差部分的跨层连接把低层特征传到高层构造残差映射，可在一定程度上保护信息的完整性。本文中选择具有强特征提取能力的深度残差网络ResNet101进行特征提取，结合特征金字塔网络FPN(Feature Pyramid Networks)共同作为主干特征提取网络。RPN网络以Backbone输出的特征图为输入，针对特征图的每个位置产生目标的候选框，并且针对不同尺度的特征

图使用不同尺寸的窗口(anchors)。如特征层{P2, P3, P4, P5, P6}分别对应的anchor尺度为 $\{32^2, 64^2, 128^2, 256^2, 512^2\}$ , 使用三种长宽比例 $\{1:2, 1:1, 2:1\}$ , 所以生成具有不同尺度和宽高比的anchor共有15种。

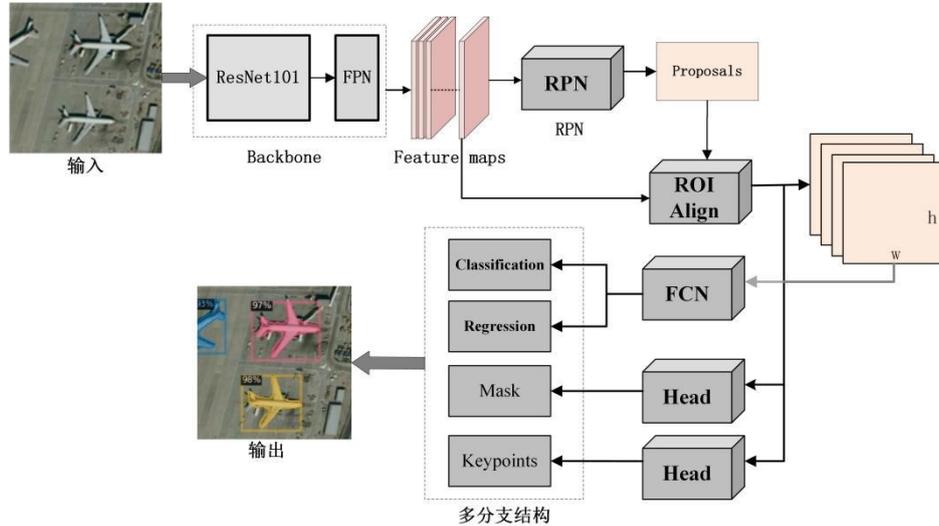


图 2.Mask R-CNN 网络结构

Figure 2. Network structure of Mask R-CNN

ROI分类器可以根据需求设置多任务的输出, 除飞机目标检测外, 飞机的掩膜信息(mask)和关键点信息(key points)分别由分割分支和关键点分支计算。分割时, 对于每一个建议框预测 $k \times m^2$ 个二值mask(其中: $k$ 表示目标类别数、 $m$ 表示特征图的大小)。对于 $m^2$ 的mask使用RoIAlign进行像素对齐, 因为RoIAlign取代之前的RoIpooling, 应用双线性插值对特征图特定区域池化, 并结合全连接网络进行像素级分割预测从而保证RoI特征能够更准确映射到原图。mask损失采用平均交叉熵损失(the average binary cross-entropy loss), 对于每一个ROI, 如果检测出ROI属于某个类别, 就只使用该类别的交叉熵损失作为loss进行训练。在多任务输出中, Classification实现目标分类, 本文中只关注飞机单类目标; Regression的作用是回归飞机目标的位置, 即表示目标位置的矩形框, 二者分别作为分类和回归分支; 两个Head分别是指分割分支和关键点分支预测损失的部分:

“预训练+微调”指利用预训练模型初始化权重, 并在此基础上进行微调的一种学习方法, 其属于一种迁移学习的方法。在源域和目标域具有一定相似性的基础上, 使其共享一部分网络参数, 从而降低模型训练的复杂度。在实际应用中常针对具体任务的特点对一些经典网络进行修改, 针对任务的特点加入领域适配层, 然后进行联合训练, 减少训练耗

将anchor与Ground Truth计算交并比(IOUS), 根据其阈值判定是否为前景框, 实现候选框筛选, 然后再对候选框进行分类与回归。

时。因此, 本文在构建数据集的基础上, 对Mask R-CNN网络的预训练模型进行调整: 使Backbone和RPN部分共享源域的此部分参数, 对Classifier部分的权重进行重新初始化, 将最后输出的网络层数调整为适配本文任务的目标类别数和关键点数量。采用自建数据集进行飞机目标检测与分割的迁移训练, 并通过模型调优, 使性能达到最佳。

### 2.1.2 基于 Mask R-CNN 的飞机目标关键点检测

关键点检测是计算机视觉领域的常见任务, 其主要用于人体骨骼关键点检测、人脸关键点检测以及手部等特定物体关键点检测。对于人体关键点检测, 关键点用于人体姿态估计, 关键点的位置和数量可自定义。在二维的多人关键点检测中, 检测方法总体上可以分成两个类型, 其一是自上而下的方法(top-down), 即先进行目标检测(人), 再对每个检测到的目标进行单人的关键点检测(单人姿态估计), 点的位置预测用坐标回归的方式来解决。代表性的方法有G-RMI(Papandreou等, 2017)、RMPE(Fang等, 2017)、CPN(Chen等, 2018); 其二是自下而上的方法(bottom-up)即先检测所有关键点, 将关键点建模成热力图, 通过像素分类任务, 回归热力图分布得到关键点位置, 然后再对关键点进行分组关联, 得到每个点的标签, 比如PAF(Cao等, 2017)、Part Segmentation(Xia等, 2017)等。这两类算法得到的关键点能够很好地刻画人体的姿态信息。受关键点

检测在人体姿态估计方面应用的启发, 本文提出面向飞机的关键点检测方法, 利用该方法提取的关键点可用于飞机位置姿态调整, 对提高飞机型号模板匹配有显著提升作用。

相比于人体姿态和形状多变的特点, 飞机在结构形态上具有较稳定的分布, 更适用于关键点检测技术。目前用于关键点检测的方法较多, 考虑到框架的简洁性与易用性, 降低网络训练难度, 本文在Mask R-CNN网络检测飞机目标的基础上, 增加关键点检测分支进行关键点检测。该分支将关键点位置建模为单个像素的掩膜, 预测每个关键点的类型, 这种关键点检测方法属于一种自上而下的检测方法, 它是在目标检测的基础上, 针对掩膜结果区分前景像素来计算关键点位置。相较于其他关键点检测方法, 基于Mask R-CNN网络的飞机目标关键点检测具有一显著优势: 在目标检测网络的基础上添加关键点分支预测目标的关键点信息, 使目标检测、分割和关键点检测任务共享大部分网络参数, 避免耗时的重复训练, 在很大程度上降低了模型调优复杂度。

关键点分支是对每个掩膜mask检测出 $k$ 个关键点, 其中的每个关键点对应一个one-hot编码的二值掩膜( $m \times m$ ), 通过训练使其中只有一个像素被标记为前景。训练时, 一个批次中预测目标数为 $n$ , 表示预测关键点信息的张量尺寸为 $(n, k, m, m)$ , 与真实关键点计算交叉熵损失函数, 其公式为式(1), 得到损失后根据关键点数量进行归一化。关键点分支部分由8个 $3 \times 3 \times 512$ 卷积层的堆栈、1个反卷积层和2倍双线性上采样组成, 产生 $56 \times 56$ 的输出。即每一个关键点其使用一个 $56 \times 56$ 的向量表示, 向量仅关键点位置为1, 其余位置为0。如图3所示。

$$H(p, q) = - \sum_{i=1}^n \sum_{j=1}^k P(x_j) \log(Q(x_j)) \quad (1)$$

其中,  $P(x)$ 表示真实关键点(gt-key points)分布,  $Q(x)$ 表示预测的关键点分布,  $k$ 表示关键点数量,  $n$ 为批次中实例的数量。

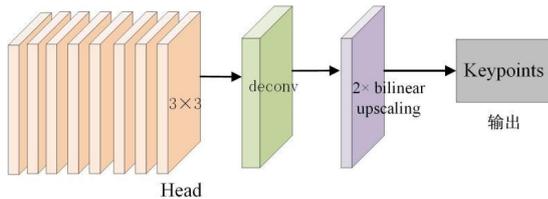


图3. 关键点检测分支的结构

Figure 3. The structure of the key points detection branch

为了准确的获得飞机姿态信息, 本文定义飞机的5个关键点, 分别是位于头部、中部和尾部三个点, 以及两侧机翼的两个点, 按照头部、机身中部、尾部、左翼和右翼的顺序分别标记为1-5号点, 方便计算和表示, 关键点的位置如图4所示。



图4. 飞机关键点位置

Figure 4. Location of key points of aircraft target

## 2.2 融合条件随机场 CRF 与目标特征的模板匹配方法

### 2.2.1 基于条件随机场的目标掩膜精化

通过前文的飞机目标快速检测, 获得了飞机的掩膜信息, 但受到Mask R-CNN网络卷积过程中诸多因素的影响——影像池化、感受野限制等, 使得获得的目标掩膜边界存在误差。针对这一问题, 本文利用条件随机场(Conditional Random Fields, CRF)算法对Mask R-CNN分割结果的目标掩膜进行精化。

条件随机场CRF(Lafferty等, 2001)是由Lafferty等人提出的一种基于判别式的概率无向图模型, 在无独立性假设的前提下能够有效处理特征间的关系(孙莹莹, 2018)。CRF模型在表达图像多特征与空间上下文信息方面具有优势, 多用于解决语义标注问题。在图像分割方面, 条件随机场在提取像素特征信息的同时, 能够利用其邻域关系来获取待观察区域的空间邻域信息, 使提取到的特征更为完整、信息利用更充分(李航, 2019), 目前已有一些学者将CRF应用与CNN分割结果后处理工作中。鉴于此, 本文将其用于飞机分割掩膜的边界精化, 将目标检测网络输出的掩膜信息进行细化。本文利用分割网络生成的掩膜来构建CRF的能量函数 $E(x)$ 如式(2)。

$$E(x) = \sum_i \theta_i(x_i) + \sum_{ij} \theta_{ij}(x_i, y_j) \quad (2)$$

其中 $\theta_i(x_i)$ 为一元势函数, 此处使用网络分割的掩膜结果作为一元势函数,  $\theta_{ij}(x_i, y_j)$ 为配对势

函数, 其公式为式(3)。

$$\theta_{ij} = \mu(x_i, y_j) \left[ \omega_a e^{-\frac{\|P_i - P_j\|^2}{2\sigma_a^2} - \frac{\|I_i - I_j\|^2}{2\sigma_b^2}} + \omega_b e^{-\frac{\|P_i - P_j\|^2}{2\sigma_\gamma^2}} \right] \quad (3)$$

式(3)中,  $x_i$  表示深度网络分割结果中像素的标签分配, 如果  $x_i = y_j$ ,  $\mu(x_i, y_j) = 1$ , 否则  $\mu(x_i, y_j) = 0$ 。其中有两个高斯核, 第一个内核取决于像素的位置  $P$  和 RGB 颜色值  $I$ , 而第二个内核仅取决于像素的位置  $P$ 。 $\sigma_a$ 、 $\sigma_b$  和  $\sigma_\gamma$  是超参数, 用于控制高斯核的尺度,  $\omega_a$  和  $\omega_b$  是用于控制两个内核的比例。CRF 通过迭代结构对输入采用多次循环迭代的方式进行求解, 使输入图像中的所有像素类都根据能量函数重新分配, 并充分利用颜色信息对每个像素进行分类。采用 CRF 后, 掩膜结果呈现很多目标轮廓的细节信息, 使分割得更精细, 如图 5 所示。

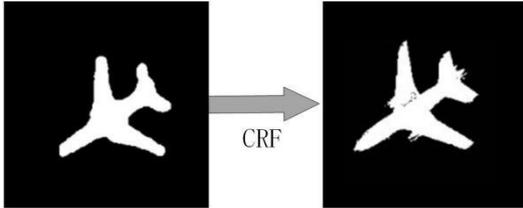


图 5. 目标掩膜精化

Figure 5. Refinement of the target mask

### 2.2.2 目标姿态调整与模板匹配

由于模板匹配对目标姿态较为敏感, 为了提高模板匹配精度, 需要对目标掩膜进行姿态调整。利用关键点, 将飞机头部关键点与尾部关键点连线或者机翼两点连线作为基准, 本文优先选择准确率高的关键点连线作为基准, 以飞机头部与尾部连线为例, 将此两点所在直线确定为目标的方向, 检测时采用两点连线与垂直方向夹角值作为方向角, 利用此角度将目标姿态调整到统一的垂直方向, 如图 6 所示。

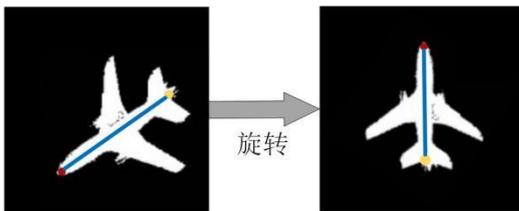


图 6. 目标姿态调整

Figure 6. Posture adjustment of the target

由于每种类型的飞机都有独特的形状和尺度特征, 模板匹配是一种解决飞机类型识别问题的实

用方法(Wu等, 2015)。对于模板匹配, 其模板制作相较于现有方法, 更为方便, 也便于后续模板库扩展。将经过边界精化与姿态调整后的飞机掩膜以二值图像的形式与飞机模板进行匹配。由于得到的掩膜结果中飞机的尺度与影像的分辨率是相关联的, 在采集模板过程中, 采用较高分辨率影像进行采集 (0.5-2m), 便于根据待匹配目标的分辨率进行调整。

模板匹配本质上是一种相似性计算, 通过计算待测试数据与模板数据的差异, 与设置的匹配阈值进行比较, 判定是否与标准模板一致, 匹配方法的差异在于差值的计算。本文的匹配方法为归一化差值平方和匹配方法, 其计算由公式 (4)。

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2}{\sqrt{\sum_{x', y'} T(x', y')^2 * \sum_{x', y'} I(x + x', y + y')^2}} \quad (4)$$

其中,  $T(x', y')$  为模板,  $I(x, y)$  表示图像,  $R(x, y)$  表示匹配结果。对于归一化差值平方和匹配方法来说, 差异  $R$  的值越小, 表明匹配效果越好。将待识别目标与模板库中所有模板计算差异, 与目标匹配效果最好的模板型号确定为待测飞机目标的型号。

## 3 实验与结果分析

### 3.1 飞机目标检测

#### 3.2.1 模型训练

为了获得网络良好的飞机检测性能, 本文采用 RSOD-Data 数据集(Y. Long等, 2017)和 MTARSI 数据集构建符合本文需要的实验数据集。其一是 RSOD-Data 数据集, 由武汉大学发布的用于遥感图像物体检测的数据集, 包含飞机、操场等四类目标, 其中飞机图像共 446 张, 共包含 4993 架飞机, 并且其中的飞机多为民用客机, 此数据集主要用来训练网络检测飞机目标的能力。其二是 MTARSI 数据集, 其包含多机型的遥感影像片段, 包括波音客机和 B-1 轰炸机、C-5 运输机、KC-135 加油机、E-3 预警机等各类军用机型, 由于此数据集具有飞机型号信息, 主要用于后续飞机型号识别任务中。为了丰富训练数据, 本文联合使用这两种数据集进行深度神经网络训练获得对飞机目标的检测能力。由于数据集中的数据不包含掩膜和关键点标签, 故针对飞机目标自行标注, 样本标注采用 labelme 工具, 增加标注目标的掩膜信息和关键点信息, 标注完成后统一转换成 COCO 数据格式。实验中仅选取部分数据: 共标

记目标数量738，作为训练集。在迁移训练过程中，将上述训练集作为目标域，MS COCO数据集的检测任务作为源域，对Mask R-CNN网络进行迁移训练：采用源域数据集目标检测任务的权重作为预训练模型进行初始化，在模型微调时使用初始学习率为0.002进行训练，训练过程中动态调整学习率。

### 3.2.1 掩膜和关键点检测结果

掩膜和关键点测试结果如图7所示，其中(a)列为测试原图，(b)列为掩膜结果，(c)列中红色点为关键点检测结果，关键点之间按照点号顺序由直线连接，以更清楚地表示。

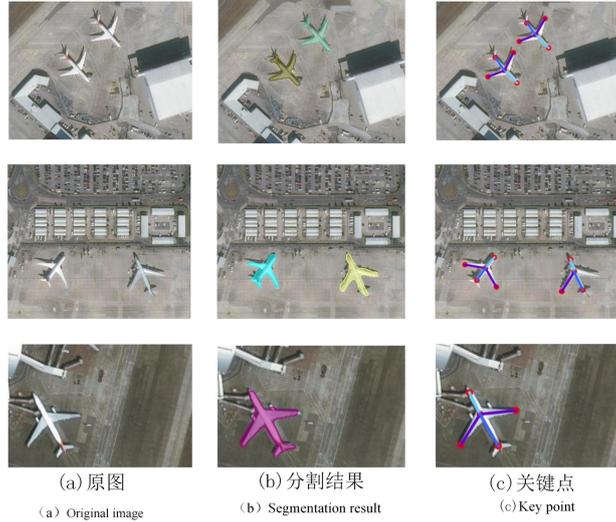


图 7.测试数据的掩膜和关键点检测结果

Figure 7. Mask and key points detection results on test data

表 1. 飞机目标检测精度

Table 1. Aircraft target detection accuracy

Score_thresh	精度 (P)	召回率(R)	虚警率
0.50	93.2%	100%	6.8%
0.70	97.4%	98.4%	2.6%
0.80	98.9%	97.9%	1.0%
0.90	100%	92.2%	0

Mask R-CNN网络进行飞机检测的精度如表1所示。在测试时，若取不同置信度阈值(Score\_thresh)，结果会自动筛选置信度高于阈值的目标。由表可知，阈值低时召回率更高，但准确率会有所降低，后续实验中为了使检测阶段得到高召回率，取阈值为0.5。从图7中的(b)列测试结果可知，经迁移训练后的得到的飞机掩膜已经具有基本飞机形状。

在本文中，我们关注关键点的检测精度，根据

测试结果评定其精度，并且认为预测关键点与真实关键点之间的距离小于阈值的点为正确检测的点。表2统计了171个目标的关键点的检测效果，由统计结果可知，头部、中部、尾部的3个关键点的精度较高，而机翼处的两个关键点精度较低，这是由于机翼处关键点的局部信息区分性较弱，造成飞机不同关键点的检测难易程度不一样，其检测难度比头部和尾部的关键点检测难度高。并且，自上而下的关键点定位依赖于检测算法的提出的候选框，可能造成检测不准和重复检测等现象。本文实验中1号点与3号点(头部和尾部)同时预测正确的目标数量为156，准确率为91.2%。根据飞机各关键点的置信度，本文选择头部和尾部两个置信度高的关键点来调整飞机的姿态。

表 2. 飞机各个关键点检测精度

Table 2. each of the aircraft key points detection accuracy

点号(位置)	1(头部)	2(中部)	3(尾部)	4(左翼)	5(右翼)
关键点总数	171	171	171	171	171
正确检出点数	167	156	163	115	97
准确率	97.7%	91.2%	95.3%	67.2%	56.7%

### 3.2 CRF 飞机掩膜精化

深度网络在遥感影像中检测出的飞机目标掩膜经过CRF精化后的结果如图8所示。图中显示了RSOD数据集中3幅影像及其飞机检测和分割的结果

果，由结果可知，CRF算法能够获得目标更多细节信息，详细的边界信息有助于后续的模板匹配操作，使其可靠性更高。

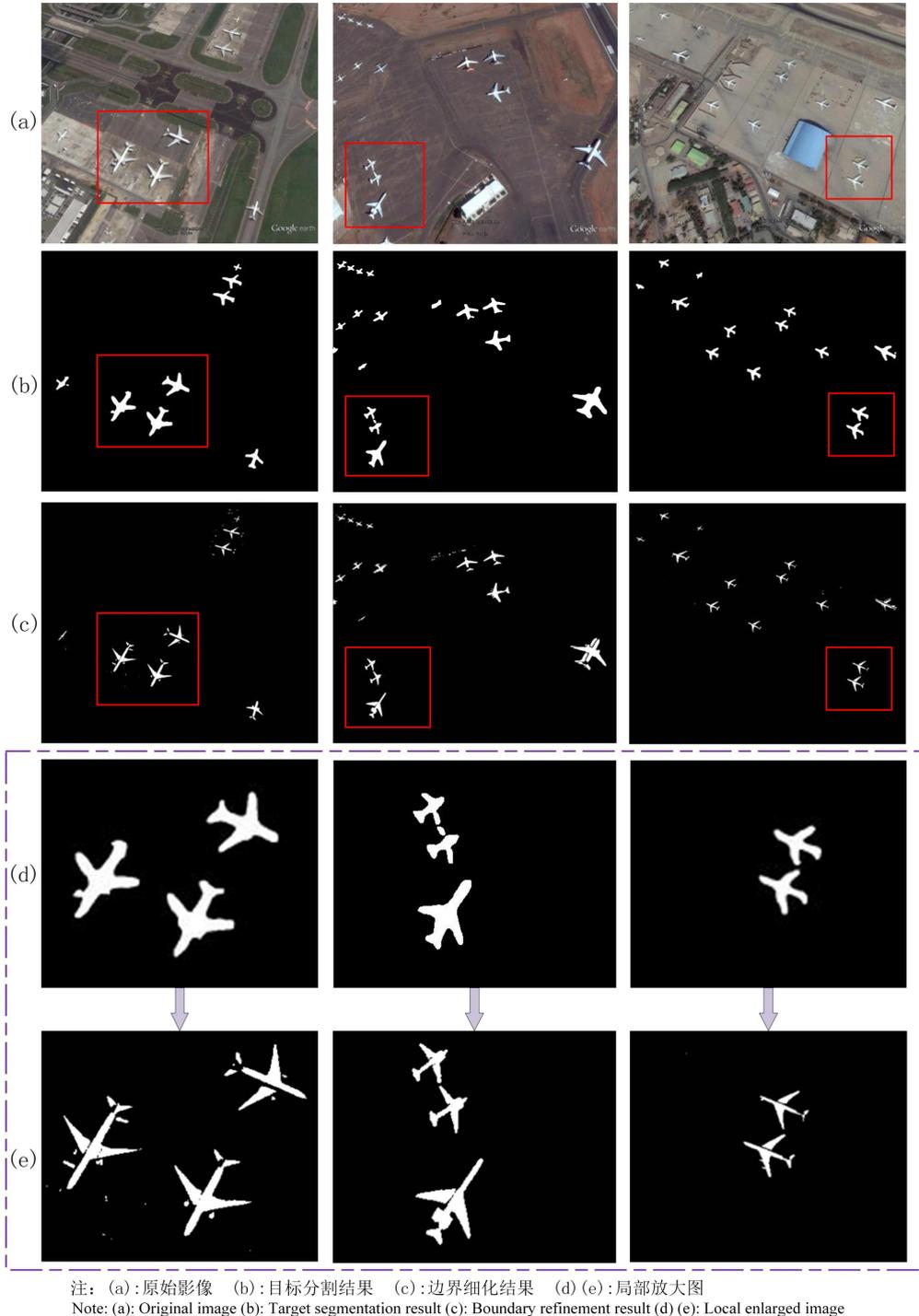


图 8.CRF 精化掩膜边界

Figure 8. CRF refines the mask boundary

### 3.3 模板匹配结果

#### 3.3.1 模板库构建

由于MTARSI数据集(Wu等, 2020)中包含多种类型的飞机型号标签, 本文基于此数据集构建飞机

型号模板库。该数据集包含数千张多种飞机类型的图像, 其在场景、姿态、空间位置等方面具有较丰富的多样性, 因此从中选取一部分数据构建模板库, 部分作为测试数据。本文选择其中的11种类型的飞

机图像进行模板的构建。对于同一种类型的飞机而言，由于姿态，角度以及光照强度等条件的影响，在遥感影像上会呈现差异，因此在制作模板时，同

一种类型可以包含多个模板，部分模板图像示例如图9所示。



图 9. 11 种飞机的部分模板

Figure 9. The 11 types of aircraft templates

### 3.3.2 模板匹配结果

首先对MTARSI数据集中的部分数据进行测试，通过模板匹配进行飞机型号识别的结果如图10所示。11种类型飞机的测试结果统计在图11的混淆矩阵中。结果显示，有10种类型的飞机型号识别准确率均能达到85%以上，识别准确率最低的是F-22侦察机，接近80%。实验发现，结构简单、形状独特的飞机型号识别准确率更高，比如B-2、B-1等，而

结构复杂、与其它类型形状相似度高的飞机型号识别准确率较低，比如E-3侦察机。并且，飞机目标的尺寸也会影响识别准确率，比如F-22侦察机的尺寸是所列类型当中最小的，从影像中获取的边界信息较难满足精细掩膜的要求，导致型号识别准确率较低。此外，飞机的阴影和机身纹理等情况也会影响精细掩膜的提取，在进行模板匹配时可针对不同类型设定不同阈值，使识别更准确。

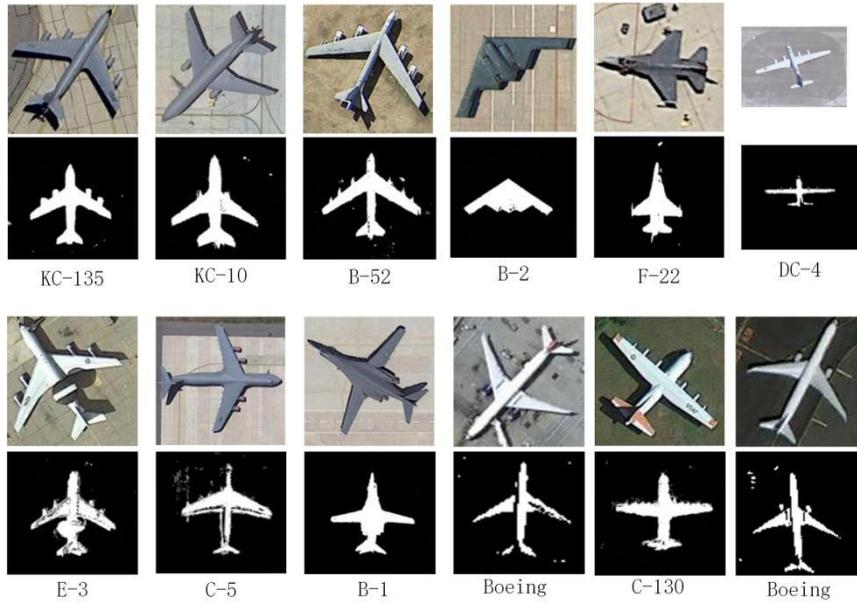


图 10.数据集中飞机型号识别结果示例

Figure 10. Example of aircraft type recognition results on our data set

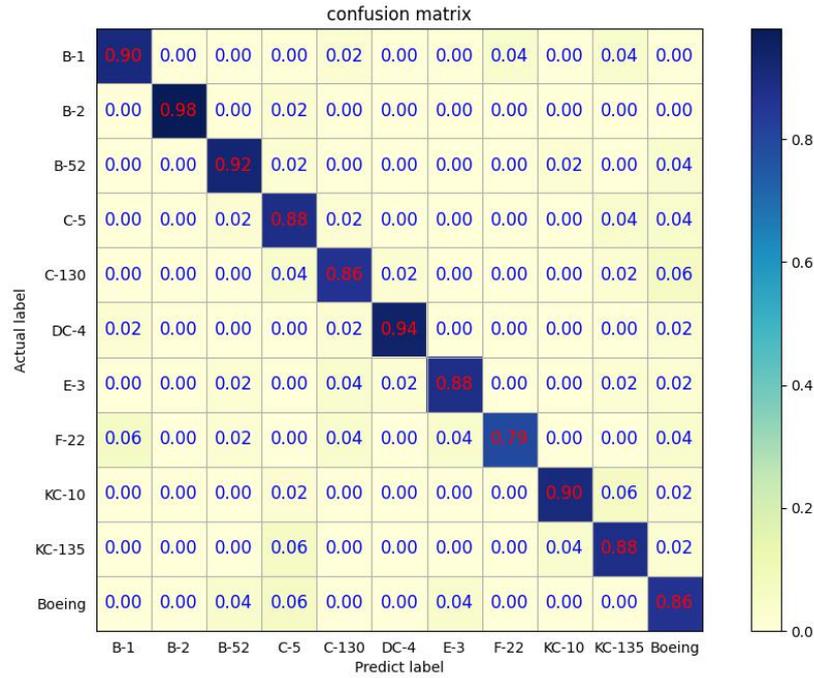


图 11.本文方法的混淆矩阵

Figure 11. Confusion matrix of our method

除数据集外, 本文对高分遥感影像进行飞机目标检测与识别的实验, 测试结果如图12所示, 其中(c)列中红色框包含的对象为型号识别正确的飞机, 蓝色框则表示未正确识别, (d)列为飞机型号识别结

果。由结果可知, 本文的方法在高分辨率遥感影像上进行飞机目标检测与型号识别方面具有较好的效果。

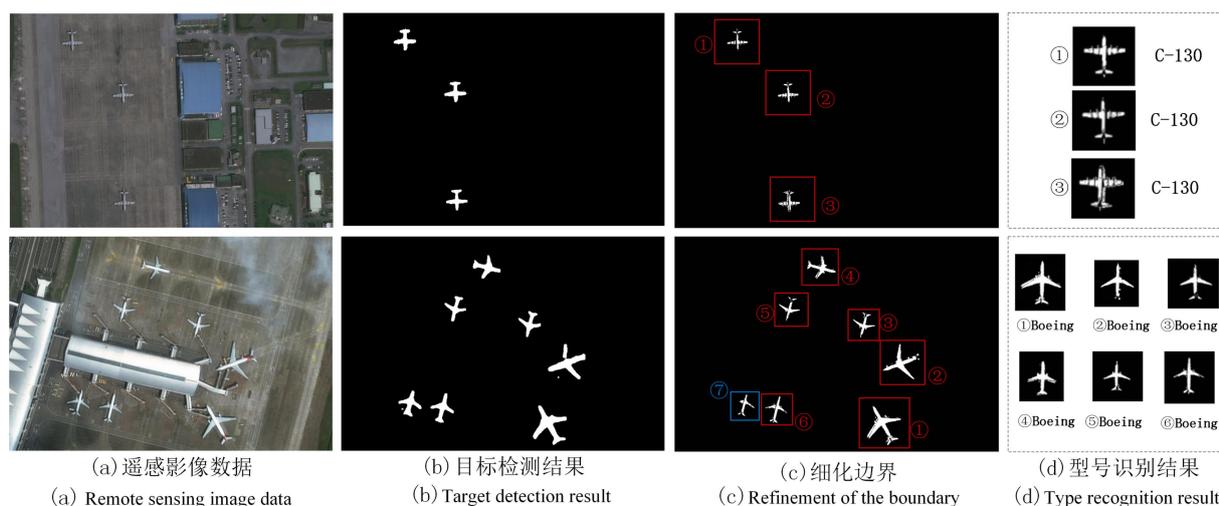


图 12.型号识别结果

Figure 12. Type recognition results

### 3.4 飞机型号识别性能评估

#### 3.4.1 方法对比

为了验证本文方法与其他方法识别性能的优劣,将所提出的方法与其他不同类型的方法进行比较。使用本文采用的MTARSI数据集作为训练和测试数据进行型号识别对比实验。我们采用计算zernike特征向量进行SVM分类的方法、应用Mask R-CNN网络进行端到端(end-to-end)的型号识别方法,以及应用YOLOv5网络进行端到端识别的方法与本文方法进行对比。

实验时选择MTARSI数据集的一部分作为训练和测试数据,另选一部分作为测试数据,对11种类型飞机共110张图像进行了型号预测,测试时输入测试集图像,输出为预测的飞机型号,如图13所示,实验参数和细节展示在表3中,其中飞机类别数量均为11类,训练集的数量由于标注难度不同所以样本数量不同,比如机器学习分类的方法不需要手动标注,所以样本数量较多,而端到端网络由于标注比较耗时,所以样本数量较少。

①机器学习分类方法SVM:通过计算zernike特征向量进行SVM分类,此方法不能直接用于从影像上判断飞机类型,此处仅在数据集进行实验,作为本文方法的一个对比。

②端到端的神经网络识别方法:Mask R-CNN(end-to-end):将11种类型的飞机图像和类别标签用于训练,输出包括飞机型号、矩形位置框以

及分割结果;

③端到端的深度学习识别方法:YOLOv5(end-to-end):同样采用11种类型的飞机图像和类别标签用于训练,输出包括飞机型号、矩形位置框,效果见图13。

⑥本文方法在目标检测模型训练阶段只需要检测出飞机类,不需要检出具体的型号,故模型的训练和调优更容易,具体型号由后续的方法检出,且输入数据集中的飞机图像,或者是高分辨率遥感影像,均具有较好的效果,如论文中图12的结果所示。

另外两种识别方法分别为:④GAN+ROI+SVM方法是一种基于条件生成对抗网络(GAN)的飞机类型识别框架,首先检测飞机关键点用于生成飞机掩膜并定位飞机的位置,其次,在未标记的飞机图像及其相应的掩码上训练具有感兴趣区域(ROI)加权损失函数的条件GAN,再对ROI进行特征提取多尺度特征,最后采用线性支持向量机(SVM)分类器根据每个样本的特征对其进行分类。⑤DBN with MSSRBM And sparse BP方法利用逐像素无监督特征学习算法来训练混合结构稀疏受限玻尔兹曼机(RBM)。通过堆叠几层RBM,构建DBNs的深度生成模型。使用具有稀疏惩罚的反向传播(BP)算法训练整个网络,最后,深度模型生成图像及其标签的良好联合分布。

表 3.对比实验参数

Table 3. Comparative experimental parameters

方法	方法类型	类别数	输入	输出	训练集数量	测试集数量
zernike 矩+SVM	机器学习分类	11	目标图像	型号	2200	110
Mask R-CNN (end-to-end)	目标检测+分割	11	图像	位置+掩膜+型号	738	110
YOLOv5 (end-to-end)	目标检测	11	图像	位置+型号	320	110
本文方法	目标检测+分割 +关键点检测	11	图像	位置+掩膜+ 关键点+型号	738	110



注：(a) 正确识别、(b) 错误识别/未识别  
 Note: (a) Correct identification、(b) Misidentified/Unidentified

图 13.Mask R-CNN 与 YOLOv5 端到端检测结果示例图

Figure 13. End-to-end detection results of Mask R-CNN and YOLOv5

表 4.不同方法的识别精度

Table4. Recognition accuracy of different methods

序号		识别飞机类别数	型号识别平均精度
①	zernike 矩+SVM	11	73.60%
②	Mask R-CNN(end-to-end)	11	68.30%
③	YOLOv5(end-to-end)	11	79.04%
④	GAN+ROI+SVM(ZhangY, et al.2018)	8	92.73%
⑤	DBN with MSSRBM And sparse BP (Wenhui Diao, et al.2015)	5	88.93%
⑥	本文方法	11	89.0%

图13展示了深度神经网络端到端方法的检测结果示例，表4显示了几种方法的对比结果。由结果可知，对于第一种方法，zernike矩在复杂场景下细节描述能力不够，造成检测精度不高；两种端到端的检测网络在飞机目标检测上具有较稳定的性能，但在型号识别的任务上，有较多的错误识别。

(1) 飞机外部形状相似度高，较大的差异在于发动机数量、机身和翼展的尺寸、以及机身纹理几个方面。采用深度网络做端到端型号识别时，由于遥感影像分辨率的限制，使得不同型号飞机的类间差异太小，导致网络能够检测出飞机目标但不能准确

地判断其型号。(2) 本文的方法利用目标检测网络良好的检测性能，首先检测出飞机，将更加困难的型号识别任务用更加稳妥的模板匹配方法来实现，使方法的稳定性和实用性方面有所增强，正是本文设计的方法的优势所在。

方法④与方法⑤分别在8种和5种型号的飞机上有92.73%和88.93%的识别准确率，但其类别广泛性比本文的11类略少。从结果看，本文方法综合性最好，受益于对掩膜的后处理与基于关键点的飞机姿态优化，最后充分利用飞机的形状和尺寸信息进行型号判断，因此优于其他方法，在测试数据集

中11种类型的飞机达到了89%的平均准确率。并且，本文的方法在飞机类别上更容易扩展，更适用于从遥感影像上检测和识别飞机类型。

### 3.4.2 消融实验

在本文数据集上，对所提的型号识别框架中的CRF模块和姿态调整模块进行了对识别性能影响的实验，将是否进行边界细化处理的掩膜用于本文识别框架，并比较结果，结果见表5。由结果可知，若不采用CRF模块对深度神经网络分割所得的飞机掩膜进行精细化，型号识别的准确率只有56.42%。由在高分辨率遥感影像中，飞机目标仍然是一种较小的目标，通过深度网络分割所得的飞机掩膜，与对应型号的飞机真实掩膜相比边界比较粗略，难以适用对掩膜精度较高的模板匹配策略。在这种情况下，类似于B-2轰炸机这类外形独特的飞机型号识

别正确率较高，而类似于KC-10和KC-135这样的外形相似度较高的飞机，就无法准确识别其型号。由此可见，CRF模块在本文的识别框架中的细化掩膜边界的重要作用，一方面将多任务的目标检测网络的结果综合利用起来，另一方面使粗略的掩膜信息细化后得到了更重要的利用。

此外，对于飞机掩膜的姿态调整模块也进行了对识别性能影响的实验。若不进行姿态调整，而使用多姿态的模板进行模板匹配。即在标准模板库中，对每一类机型扩充多方向的标准掩膜用于模板匹配，从而实现待测目标的型号识别，由此得到的飞机型号识别准确率为71.8%。由对比可知，利用关键点调整待测飞机的掩膜姿态比扩充模板库中标准模板的方向准确率更高，且匹配时计算量更小、整个识别流程更高效。

表 5.型号识别框架消融实验对比结果

Table 5. Ablation study of our method

Mask r-cnn 网络	CRF 掩膜精化	姿态调整	模板匹配	型号识别率
√		√	√	56.42%
√	√		√	71.80%
√	√	√	√	89.0%

## 4 结论

针对现有遥感影像飞机目标型号识别方法难度较大的问题，本文提出一种融合目标分割与关键点检测的飞机型号识别方法。该方法有机地结合多任务深度神经网络与条件随机场和模板匹配算法，利用“预训练+微调+后处理”的方式实现飞机型号的高精度识别。结果表明，（1）联合深度学习目标分割与关键点检测进行型号识别，能够改善和弥补传统方法和端到端方法在飞机精细检测任务中精度低的问题。（2）本文的方法在可识别的飞机类型的数量上有所提高，并且可通过增加相应的标准模板来扩展新类型飞机的检测。（3）通过有效结合现有深度学习技术与传统方法，在基于遥感影像的飞机目标精细检测方面具有实用性与有效性，同时对于其他类似的重点目标如海面上舰船目标、重点军事设施等的精细检测具有一定的借鉴意义。

尽管本文方法在飞机型号识别数量和识别精度上均有提高，但仍有一些不足，后续研究将从以下方面进行改进：（1）目前，可识别的飞机类别的广泛性仍有很大的局限性，数据集和算法设计应向着更广泛、更全面的类型方向发展，扩充标准数据集等。（2）作为军事目标，飞机存在遮挡、涂

装等反侦察手段，光学影像在这些情况下难以发挥作用。随着SAR卫星的广泛应用，它获取数据时不会受限于天气和光照条件，数据获取稳定，所以针对飞机这类目标的细粒度检测可以更多地利用SAR数据，或者光学与SAR的结合，达到更准确、更稳健的检测效果。

## 参考文献 (References)

- Chen Yilun, et al. Cascaded Pyramid Network for Multi-Person Pose Estimation [J]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7103–7112.
- Cao Zhe, et al. Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields [J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 1302–1310.
- Chen H.Wang, F.Xu and Y.Jin, Target Classification Using the Deep Convolutional Networks for SAR Images [J]. IEEE Transactions on Geoscience and Remote Sensing, vol.54, no.8, pp.4806–4817, Aug.2016.
- Fu Kun, et al. MultiCAM: Multiple Class Activation Mapping for Aircraft Recognition in Remote Sensing Images [J].

- Remote Sensing, vol. 11, no. 5, 2019, p.544.
- Fang Hao-Shu, et al. RMPE: Regional Multi-Person Pose Estimation [J]. 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2353–2362.
- Guo H.Wang and F.Xu, Scattering Enhanced Attention Pyramid Network for Aircraft Detection in SAR Images [J]. IEEE Transactions on Geoscience and Remote Sensing, vol.59, no.9, pp.7570-7587, Sept.2021, doi: 10.1109/TGRS.2020.3027762.
- He Kaiming, Georgia Gkioxari, Piotr Dollar, and Ross Girshick. 2020. Mask R-CNN [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 42 (2): 386-97.
- He Kaiming, et al. Deep Residual Learning for Image Recognition [J]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.
- Hang Li, Statistical Learning Methods (2nd Edition), Tsinghua University Press [M], 2019.05.(李航, 统计学习方法(第2版), 清华大学出版社 [M], 2019.05.)
- J. Ai, R. Tian, Q. Luo, J. Jin and B. Tang, Multi-Scale Rotation-Invariant Haar-Like Feature Integrated CNN-Based Ship Detection Algorithm of Multiple-Target Environment in SAR Imagery [J]. IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 12, pp. 10070-10087, Dec. 2019.
- Li Ke, Wang Runsheng, Wang Cheng. A tree classifier method for aircraft model recognition [J]. Computer Engineering and Science, 2006(11):136-139. (李科, 王润生, 王程. 一种用于飞机型号识别的树分类器方法 [J]. 计算机工程与科学, 2006(11):136-139.)
- Lafferty, John D, et al. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data [J]. ICML'01 Proceedings of the Eighteenth International Conference on Machine Learning, 2001, pp. 282-289.
- Mahi, Habib, et al. Zernike Moments and SVM for Shape Classification in Very High Resolution Satellite Images [J]. The International Arab Journal of Information Technology, vol. 11, 2014, pp. 43-51.
- Papandreou, George, et al. Towards Accurate Multi-Person Pose Estimation in the Wild [J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3711–3719.
- Ren Shaoqing, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, 2017, pp. 1137–1149.
- Sun X, Wang P, Yan Z, et al. FAIR1M: A Benchmark Dataset for Fine-grained Object Recognition in High-Resolution Remote Sensing Imagery [J]. arXiv preprint arXiv:2103.05569, 2021.
- Sun Yingying. Semantic segmentation of high-resolution SAR images based on multi-scale FCN-CRF and reinforcement learning [D]. Xidian University, 2018. (孙莹莹. 基于多尺度 FCN-CRF 和强化学习的高分辨率 SAR 图像语义分割 [D]. 西安电子科技大学, 2018.)
- Wang Wei, Li Yuanlei, Liang Fang, et al. A ground aircraft target recognition method based on region segmentation [J]. Computer Engineering, 2013, 39(06):214-218. (王威, 李元垒, 梁方, 张连峰. 一种基于区域分割的地面飞机目标识别方法 [J]. 计算机工程, 2013, 39(06):214-218.)
- Wahi, Amitabh, et al. Rotated Object Recognition-Based on Hu Moment Invariants Using Artificial Neural System [J]. 2012 World Congress on Information and Communication Technologies, 2012, pp. 45-49.
- Wang B, Zhou Y, Zhang H N, etc. Remote sensing image aircraft target detection method based on improved SSD framework [J]. Firepower and Command Control, 2021, 46(01): 14-19. (王冰, 周焰, 张怀念等. 基于改进 SSD 框架的遥感影像飞机目标检测方法 [J]. 火力与指挥控制, 2021, 46(01):14-19.)
- Wu, Qichang, et al. Aircraft Recognition in High-Resolution Optical Satellite Remote Sensing Images [J]. IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 1, 2015, pp. 112-116.
- Wu, Zhi-Ze, et al. A Benchmark Data Set for Aircraft Type Recognition from Remote Sensing Images [J]. Applied Soft Computing, vol. 89, 2020, p. 106132.
- Wenhui Diao, Xian Sun, et al. (2015) Object recognition in remote sensing images using sparse deep belief networks, Remote Sensing Letters [J], 6:10, 745-754, DOI: 10.1080/2150704X.2015.1072288.
- Xia Fangting, et al. Joint Multi-Person Pose Estimation and Semantic Part Segmentation [J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 6080–6089.
- YAO Q L, HU X, LEI H. Aircraft detection in remote sensing imagery with multi-scale feature fusion convolutional neural networks [J]. Acta Ceodaetica et Sinica, 2019, 48(10):1266-1274. (姚群力, 胡显, 雷宏. 基于多尺度融

- 合特征卷积神经网络的遥感图像飞机目标检测[J]. 测绘学报, 2019, 48(10):1266-1274.)
- Yang Yumin, et al. Remote Sensing Image Aircraft Target Detection Based on GIoU-YOLO V3 [J]. 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP), 2021, pp. 474-478.
- Yan Li. Research on Remote Sensing Image Aircraft Target Recognition Based on Convolutional Neural Network[D]. University of Science and Technology of China, 2018. (颜荔. 基于卷积神经网络的遥感图像飞机目标识别研究[D]. 中国科学技术大学, 2018.)
- Y. Long, Y. Gong, Z. Xiao and Q. Liu. Accurate Object Localization in Remote Sensing Images Based on Convolutional Neural Networks[J]. IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 5, May 2017, pp. 2486-2498.
- Zhang Wanceng, et al. Object Detection in High-Resolution Remote Sensing Images Using Rotation Invariant Parts Based Model [J]. IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 1, 2014, pp. 74-78.
- Zhu Dan, Wang Bin, Zhang Liming. Remote sensing image airport target detection based on linear proximity parallelism and GBVS saliency [J]. Journal of Infrared and Millimeter Waves, 2015, 34(3):375-384. (朱丹, 王斌, 张立明. 基于直线邻近平行性和 GBVS 显著性的遥感图像机场目标检测 [J]. 红外与毫米波学报, 2015, 34(3):375-384.)
- Zhang, Ruiqian, et al. Multi-Scale Adversarial Network for Vehicle Detection in UAV Imagery [J]. Isprs Journal of Photogrammetry and Remote Sensing, vol. 180, 2021, pp. 283-295.
- Zhang Kai, Liu Hao, Yang Xi, Li Shaoyi, etc. Aerial infrared target key parts recognition algorithm based on key point detection network [J]. Journal of Northwestern Polytechnical University, 2020, 38(06):1154-1162. (张凯, 刘昊, 杨曦, 李少毅等. 基于关键点检测网络的空中红外目标要害部位识别算法[J]. 西北工业大学学报, 2020, 38(06):1154-1162.)
- Zhang Mingcheng, Wu Xiuqing, Wang Pengwei. Aircraft recognition based on closed contour extraction and partial feature matching [J]. Computer Simulation, 2006(11):193-197. (张名成, 吴秀清, 王鹏伟. 基于闭合轮廓提取和部分特征匹配的飞机识别 [J]. 计算机仿真, 2006(11):193-197.)
- Zhang Y.; Sun, H. et al. Aircraft Type Recognition in Remote Sensing Images Based on Feature Learning with Conditional Generative Adversarial Networks[J]. Remote Sens. 2018, 10, 1123.
- Zuo Jiawei, et al. Aircraft Type Recognition Based on Segmentation With Deep Convolutional Neural Networks [J]. IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 2, 2018, pp. 282-286.

## Aircraft type recognition method by integrating target segmentation and key points detection

LIU Siting<sup>1,2</sup>, WANG Qingdong<sup>1</sup>, ZHANG Li<sup>1</sup>, HAN Xiaoxia<sup>1</sup>, WANG Baoqian<sup>1</sup>, LIU Yuxian<sup>3</sup>

1. Chinese Academy of Surveying and Mapping, Beijing 100830, China;

2. Faculty of Geomatics, Lanzhou Jiaotong University, Lanzhou 730070, China;

3. Shenzhen Investigation & Research Institute Co., Ltd, Shenzhen 518026, China

**Abstract:** Objective: Aircraft detection by deep learning is a hot field in remote sensing image analysis. However, due to the limited perspective of satellite imagery and high similarity in appearance, aircraft type recognize is still a challenging task. The existing deep learning methods cannot be satisfied with the fine-grained aircraft type recognition tasks well, which require a refined lables for datasets. Thus, for recognizing aircraft type recognition in remote sensing images, in this paper, we propose an target segmentation and key points detection integrated aircraft type recognition method.

Method: The method organically combines the multi-task deep neural network with the conditional random field and template matching algorithm, and achieves the high-precision recognition of the aircraft type by means of "pre-training, fine-tuning, post-processing". First, based on the multi-task learning and transfer learning technology, the aircraft target position, mask and key points recognition are realized. Secondly, in order to facilitate the high-precision template matching in the later stage, the aircraft target mask refinement algorithm and the key-points based mask attitude adjustment algorithm are proposed to achieve the boundary refinement of the recognition target and the aircraft target mask attitude adjustment.

18 Finally, based on the aircraft type template library constructed in this paper, the refined aircraft mask information is matched  
19 with the template library to recognize the aircraft type.

20 Results: The algorithm was applied to the MTARSI data set and remote sensing images for verification. The results  
21 showed that the recognition accuracy of 11 types is 89%. Aircraft with simple structure and unique shape have higher  
22 recognition accuracy, such as B-2, B-1, etc., while aircraft with complex structure and high similarity with other types of  
23 shapes have lower recognition accuracy, such as E-3 reconnaissance aircraft. The algorithm is compared with traditional  
24 algorithms and end-to-end deep learning methods, and the results show that our method on the 11 types of aircraft get 15.4%  
25 and 20.7% more accurate than the two contrasting methods, respectively.

26 Conclusion: The use of target segmentation and key points information has achieved good results in model recognition  
27 on high-resolution remote sensing images. However, limitations remain in the breadth of identifiable aircraft types.  
28 Therefore, further research is needed.

29 **Key words:** Object detection; segmentation; key points detection; conditional random field; aircraft type recognition;

30 Supported by: 1、 National Key R&D Program of China (No.2019YFB1405600); 2、 Shenzhen Technical Research Project  
31 (No.JSGG20191129103003903);