

# 自适应交叉融合局部特征的空间目标小样本识别方法

谢文茜,任笑圆,王粲雨,蒋李兵,王 壮\*

(国防科技大学 电子科学学院,湖南 长沙 410073)

**摘要:**针对空间目标在低频次观测的小样本识别场景中,不同姿态下图像表征变化剧烈导致的辨识性特征提取难、图像间特征关联难的问题,提出一种自适应交叉融合局部特征的空间目标小样本识别方法。在现有小样本学习框架上,引入基于自注意力和互注意力的特征交叉融合模块,自适应地学习局部特征之间的相关关系,提高不同姿态下特征提取的判别性和鲁棒性,有效挖掘支持集和查询集之间的相似性,提升存在表征差异条件下的特征关联准确性。同时,在损失函数中引入基于邻域密度的样本标签权重,以解决空间目标数据集中姿态不均衡导致的网络模型学习偏差问题。通过在不同数据集上的验证,证明提出的方法具有更高的识别精度。

**关键词:**空间目标识别;小样本学习;样本不均衡;特征交叉融合;注意力机制

**中图分类号:**TP391 **文献标志码:**A **文章编号:**1001-2486(2025)03-162-11



论  
文  
拓  
展

## Few-shot space target recognition method based on adaptive cross fusion of local features

XIE Wenxi, REN Xiaoyuan, WANG Canyu, JIANG Libing, WANG Zhuang\*

(College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China)

**Abstract:** In the few-shot recognition scenario of space targets observed at low frequency, the drastic changes in the image representation of space targets in different poses challenges to the extraction of discriminative features and the correlation of features between images. To address these issues, the few-shot space target recognition method based on adaptive cross fusion of local features was proposed. Based on the existing few-shot learning framework, the feature cross fusion module based on self-attention and cross-attention was used to adaptively learn the correlation between local features, improve the discriminant and robustness of feature in different poses, effectively explore the similarity between the support set and the query set, and improve the accuracy of feature association with representation differences. Meanwhile, the sample label weight based on neighborhood density was employed into the loss function to solve the learning bias problem of the network model caused by unbalanced space target datasets. Through the verification on different datasets, the proposed method is proved to achieve higher recognition accuracy.

**Keywords:** space target recognition; few-shot learning; unbalanced data; feature cross fusion; attention mechanism

空间目标识别是空间监视的基础任务,同时也是后续特性研判任务的前提。传统的空间目标识别手段是通过目标在轨位置确认其身份,随着航天活动日趋复杂,目标轨道机动频繁发生,临近轨道目标大量增加,仅靠目标轨道信息往往难以准确、唯一地辨别目标身份。近年来,天基监视卫星以其平台灵活、成像距离短的优势受到广泛重视,基于天基光学图像识别空间目标身份是空间

监视的重要手段,具有急迫的现实应用需求。

在实际场景中,轨道交会成像是较为常见的天基光学成像方式,即观测平台和观测目标各自在轨运行中,两者间隔距离符合成像要求时进行成像。由于相对速度快、交会时间短,加之运行轨道的位置约束,一次交会成像数据有限,因此低频次观测目标(如新入轨的空间目标)短期内无法积累大量样本,难以满足传统的识别网络训练需

收稿日期:2023-04-28

基金项目:国家部委基金资助项目(ATR-240302)

第一作者:谢文茜(1999—),女,河南洛阳人,硕士研究生,E-mail:xiewenxi111@163.com

\*通信作者:王壮(1973—),男,江苏镇江人,教授,博士,博士生导师,E-mail:zhuang\_wang@sina.com

引用格式:谢文茜,任笑圆,王粲雨,等. 自适应交叉融合局部特征的空间目标小样本识别方法[J]. 国防科技大学学报, 2025, 47(3): 162-172.

Citation: XIE W X, REN X Y, WANG C Y, et al. Few-shot space target recognition method based on adaptive cross fusion of local features[J]. Journal of National University of Defense Technology, 2025, 47(3): 162-172.

求。基于航天器结构设计具有一定共性准则这一前提,解决以上问题的一个有效思路是将其他目标的历史观测信息迁移到低频次观测目标的识别任务中,即可视为小样本识别问题,从而提高低频次观测目标的识别率。然而,不同姿态下空间目标图像表征变化明显,类内差异显著,如图1所示,给识别任务带来极大挑战。一方面,网络提取目标的辨识性特征难度大幅增加;另一方面,同一特征在不同姿态下差异较大,极易造成关联失配。此外,对于成像时间内在轨姿态调整频繁的目标,观测视角通常无法均匀变化,造成观测样本库中类内姿态空间分布不均衡,从而导致分类器过多关注姿态密集的样本,这也限制了空间目标识别性能的提升。

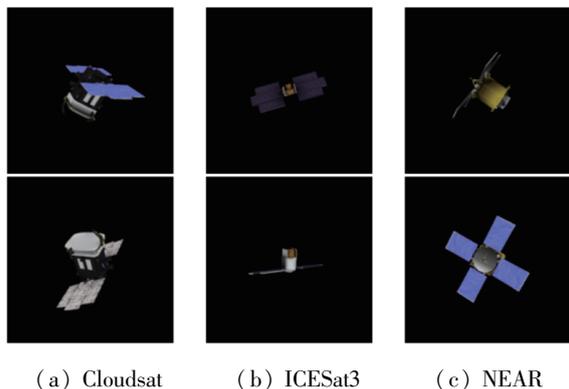


图1 不同姿态下表征差异较大的空间目标

Fig.1 Space targets with large representation difference with different pose

当前,智能方法在空间目标识别领域的应用仍处于初步探索阶段,大多数研究依然沿用地物目标适用的检测识别网络<sup>[1-2]</sup>,对于太空观测特殊场景下低频次观测目标的识别问题缺乏深入研究。Yang等<sup>[3]</sup>提出了融合星本体和星翼部件语义信息的细粒度识别算法来提高识别精度。收集大量数据与部件提取标定的困难,制约了这种方法的性能提升。此外,低频次观测目标样本的缺乏无法满足深度学习模型训练的需要,会导致平均识别精度低,Yang等<sup>[4]</sup>提出用小样本识别方法解决这一问题,并引入带有类内紧凑原则的损失函数缓解类内差异较大的问题,但并未从特征提取层面提出解决方法,识别精度仍然受限于网络的特征表达能力,在面对复杂性高、差异性大的空间目标时识别效果不理想,需要进一步研究如何从特征层面提升空间目标识别的性能。

基于度量学习的小样本识别方法<sup>[5-12]</sup>为样本不足条件下的目标识别提供了思路。在小样本

学习网络中引入注意力机制,可以加强网络对辨识性特征的关注,这对于解决空间目标识别中类内差异大的问题具有重要意义。Hou等<sup>[9]</sup>提出的交叉注意网络,解决了图像中包含多个对象时难以准确捕捉到目标的问题。Kang等<sup>[10]</sup>通过基于相似性的自相关和交叉相关,分别学习图像内相关区域和图像之间的关系,以此来突出目标物体区域,从而达到定位目标的效果。Tong等<sup>[11]</sup>提出的DBRN网络设计了两个通道注意力和空间注意力模块,对特征嵌入网络提取的特征进行空间和通道加权,以便模型能够更多地关注两幅图像的相似部分。Ren等<sup>[12]</sup>提出的MLFRNet网络通过随机裁剪提取图像的局部特征,并提出一种多局部特征关系模型来探索局部特征之间的关系。这些网络本质上都是利用点积或余弦相似度计算局部特征之间的相关性,基于支持集和查询集的相似性捕捉、增强特征。这种相关性的计算是预先设定、固定不变的,不能通过网络学习自适应调整。在空间目标识别背景中,如果同一目标在不同姿态下出现形变,仅仅利用图像间的相似性来计算局部特征的注意力权重难以使网络关注到对分类有帮助的特征。Transformer中强大的注意力机制能够长距离地交互上下文和理解内在特征<sup>[13]</sup>,借鉴这种思想,提出参数可学习的自适应特征交叉融合模块,取代原有基于相似性增强特征的做法,加强局部特征之间的交互融合<sup>[14-16]</sup>,帮助网络提取出更具判别性的特征,进而提高决策能力。

综上所述,制约低频次观测空间目标识别的核心问题是样本空间中图像表征类内差异大,以及姿态分布不均衡。针对以上问题,基于特征提取和相似性度量框架的小样本识别基本架构,从自适应交叉融合特征、均衡不同姿态样本关注度的思路出发,提出了一种适用于轨道交会成像场景下的空间目标小样本识别方法。

## 1 问题定义

采用基于任务训练机制的元学习框架实现低频次空间目标识别,本质思想是将其他目标的历史观测信息泛化到低频次空间目标识别任务场景下。空间目标识别场景下存在已积累有大量带标签样本的空间目标数据集 $D_{\text{train}}$ ,其中有 $C_{\text{train}}$ 个类别,用作训练和验证。同时,使用仅有少量标签样本的低频次观测目标数据集 $D_{\text{test}}$ 进行测试,其中有 $C_{\text{test}}$ 个类别,与训练集 $C_{\text{train}}$ 个类别不交叉。

当实验设置为  $N$ -way  $K$ -shot 的形式时,在训练阶段,从数据集  $D_{\text{train}}$  中随机抽取  $N$  个类,每类取  $K$  个做支持集  $S$ ,再抽取 15 个做查询集  $Q$ ,训练模型学习从查询集图像到支持集  $N$  个类标签的映射。在测试阶段,同样从数据集  $D_{\text{test}}$  中选取  $N$  个类别,每个类提供  $K$  个带标签样本做支持集,即可使用模型学习到的映射将剩余的无标签查询集分类为支持集的  $N$  个类之一。实验部分将在 5-way 1-shot 和 5-way 5-shot 的实验设置下给出分类精度。

## 2 整体网络体系

所提方法的网络结构如图 2 所示,主要包含特征提取模块、自适应特征交叉融合模块和相似性度量模块。以 5-way 1-shot 的实验设置为例,训练阶段从  $D_{\text{train}}$  里选取 5 个类别,每个类别随机抽取一个样本组成支持集  $S = \{S_1, S_2, S_3, S_4, S_5\}$ ,查询集  $Q$ 。支持集  $S$  和查询集  $Q$  经过预处理,以

$84 \times 84$  的分辨率送入特征提取模块。特征提取模块采用 ResNet-12 的卷积网络结构<sup>[17]</sup>,得到特征图  $F_S, F_Q \in \mathbb{R}^{h \times w \times c} (h=5, w=5, c=640)$ ,将其展平看作是两组  $m$  个  $d$  维 ( $m=h \times w, d=c$ ) 的局部描述子序列:  $F_S = [s_1, \dots, s_m] \in \mathbb{R}^{d \times m}, F_Q = [q_1, \dots, q_m] \in \mathbb{R}^{d \times m}$ 。特征交叉融合模块自适应地增强和融合局部描述子之间的信息,得到筛选出重要信息的图像局部描述子序列  $F_{A_S} = [A_{S_1}, \dots, A_{S_m}] \in \mathbb{R}^{d \times m}, F_{A_Q} = [A_{Q_1}, \dots, A_{Q_m}] \in \mathbb{R}^{d \times m}$ 。最后,提出局部特征级的相似性度量模块,该模块基于局部特征之间的相似性求得查询集与 5 个类别之间的相似度得分,通过分类器将查询集分类为与之对应的相似度得分最高的类别。图 2 中的虚线仅在训练过程中存在,训练时计算两个分类器损失优化网络,测试时仅取相似性度量分类器的结果作为分类结果。其中,自适应交叉融合模块是网络核心部分,下面将对其进行详细介绍。

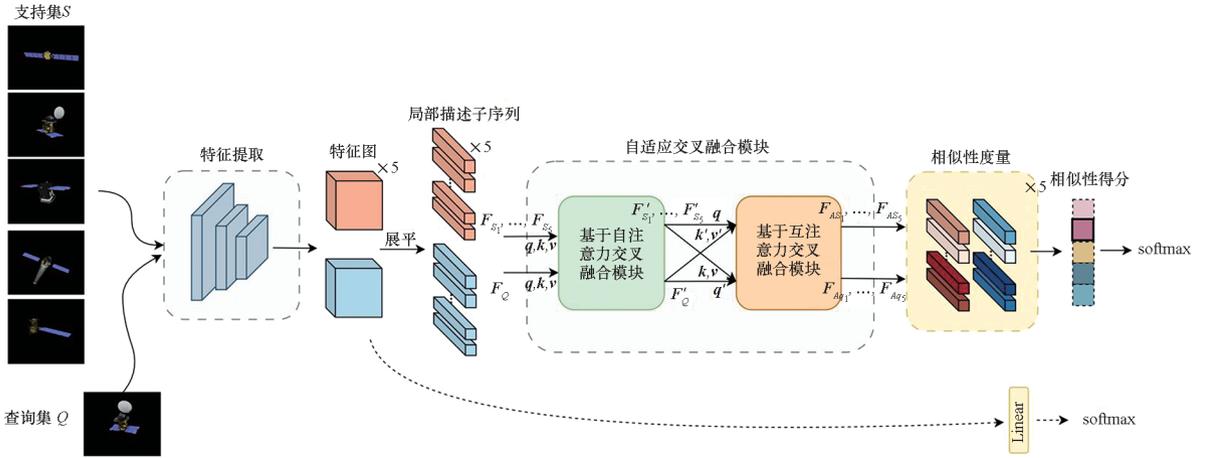


图 2 本文方法网络结构

Fig. 2 Network structure of the proposed method

### 2.1 基于自注意力的特征交叉融合模块

基于自注意力的特征交叉融合模块的作用是通过自适应学习并增强不同空间位置之间的特征相关关系,保留重要特征并消除冗余和噪声信息,准确定位更具辨识性和姿态鲁棒性的深层特征。该模块为后续的物体检测和识别任务提高了性能和可靠性,具体结构如图 3 所示。

将支持集和查询集的特征图展平得到局部描述子向量序列  $F_S, F_Q$ ,送入基于自注意力交叉融合模块,首先加上空间位置编码信息避免展平后丢失位置信息,使用 2D 扩展的标准位置编码<sup>[14]</sup>描述空间位置信息,即

$$f(x, y)^i = \begin{cases} \sin(\omega_n \cdot x), & i = 4n \\ \cos(\omega_n \cdot x), & i = 4n + 1 \\ \sin(\omega_n \cdot y), & i = 4n + 2 \\ \cos(\omega_n \cdot y), & i = 4n + 3 \end{cases} \quad (1)$$

其中,  $f(x, y)^i$  表示特征图中位置为  $(x, y)$  的图像局部描述子中第  $i$  个特征通道的编码信息,  $i$  为特征通道索引。  $\omega_n = \frac{1}{10000^{2n/d}}$ ,  $d$  为图像局部描述子特征通道数,  $n$  取  $[0, \lceil d/4 \rceil]$ ,  $\lceil \cdot \rceil$  表示向上取整。

通过多头注意力衡量各个局部描述子的重要程度,并在特征融合过程中进行加权。设注意力交叉融合模块的输入是图像局部描述子向量序列  $X$ ,经过线性层实现的变换矩阵得到查询向量  $q$ 、

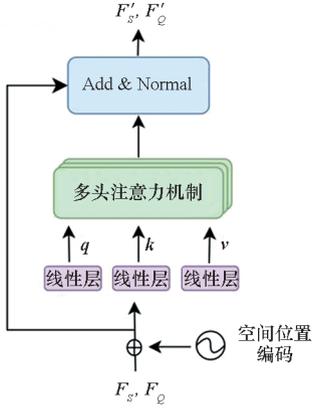


图3 基于自注意力的特征交叉融合模块

Fig. 3 Feature cross fusion module based on self-attention

键向量  $k$  和值向量  $v$ 。类似于信息检索,根据每个值  $v$  对应的查询向量  $q$  与键向量  $k$  的点积计算出注意力权重,并从值向量  $v$  中检索信息。这个过程可表示为:

$$\begin{cases} X \times W^q = q \\ X \times W^k = k \\ X \times W^v = v \end{cases} \quad (2)$$

$$A(q, k, v) = \text{softmax}(q \cdot k^T) v \quad (3)$$

该模块的核心在于通过可学习的变换矩阵,自适应地筛选出对分类有帮助的特征信息,并消除冗余信息。

多头注意力机制是使用多组权重矩阵得到多组  $q, k, v$ , 计算出多个信息矩阵后拼接,并使用线性层进行融合,则多头注意力机制可表示为:

$$M(X, X, X) = \text{concat}(h_1, \dots, h_k) W^O \quad (4)$$

$$h_i = A(XW_i^q, XW_i^k, XW_i^v) \quad (5)$$

其中,  $W_i^q, W_i^k, W_i^v \in \mathbb{R}^{d \times d_k}$ ,  $W^O \in \mathbb{R}^{kd_k \times d}$ ,  $d = 640$ ,  $k = 8$ ,  $d_k = d/k = 80$ , 在 Mini-ImageNet 数据集上设置多头注意力中的融合线性层为两层,其余数据集为一层。

送入自注意力模块的支持集和查询集局部描述子序列分别表示为  $F_s, F_q$ , 经过基于自注意力交叉融合模块后用  $F'_s, F'_q$  表示:

$$F'_s = F_s + M(F_s, F_s, F_s) \quad (6)$$

$$F'_q = F_q + M(F_q, F_q, F_q) \quad (7)$$

## 2.2 基于互注意力的特征交叉融合模块

基于自注意力的交叉融合模块主要考虑了图像内部的信息交互,忽略了不同图像之间存在相关关系的特征,这可能会降低同类图像之间的相似性。因此引入了基于互注意力的特征交叉融合模块,通过对支持集和查询集的图像局部特征序

列进行交叉关注,挖掘出支持集和查询集之间的局部特征相关性,提高模型对图像之间相似性的理解。

如图4所示,查询集和支持集经过可学习的变换矩阵得到各自的查询向量、键向量和值向量,为了增加图像之间的信息交互,用支持集  $F'_s$  的查询向量  $q$  从  $F'_q$  的值向量  $v$  中检索信息,相对应地,用查询集  $F'_q$  的查询向量  $q'$  从  $F'_s$  的值向量  $v$  中检索信息,通过此过程交互融合图像之间的有用信息,增强图像之间相似的部分。经过基于互注意力交叉融合模块的支持集和查询集用  $F_{As}, F_{Aq}$  表示:

$$F_{As} = F'_s + M(F'_s, F'_q, F'_q) \quad (8)$$

$$F_{Aq} = F'_q + M(F'_q, F'_s, F'_s) \quad (9)$$

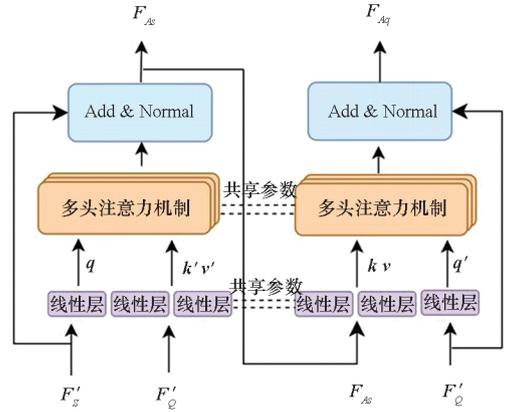


图4 基于互注意力的特征交叉融合模块

Fig. 4 Feature cross fusion module based on cross-attention

## 2.3 基于局部特征相似性的度量模块

相似性度量模块计算支持集和查询集的相似度得分。现有的相似性度量大多先对两幅特征图全局平均求得特征向量,再计算两者之间的余弦相似度,这种方式可能会损失许多有区分度的信息。为了更加准确地描述图像的特征,监督网络中局部特征之间的信息交互,采用基于局部特征相似性的度量模块。每个局部描述子对应图像中的一个局部区域,使用余弦相似度计算两幅图所有对应局部描述子之间的距离,然后取均值作为两幅图的相似度。基于局部特征相似性的度量可以充分地利用每个局部特征向量的信息,同时保留每个局部特征向量之间的相似性和差异性。经过自适应特征交叉融合后得到  $F_{As} = [A_{s1}, \dots, A_{sm}] \in \mathbb{R}^{d \times m}$ ,  $F_{Aq} = [A_{q1}, \dots, A_{qm}] \in \mathbb{R}^{d \times m}$ , 则支持集和查询集的相似性得分计算公式如下:

$$C_{\text{sim}}(\mathbf{F}_{A_s}, \mathbf{F}_{A_q}) = \frac{1}{m} \sum_{i=1}^m \frac{A_{s_i} \cdot A_{q_i}}{\|A_{s_i}\| \cdot \|A_{q_i}\|} \quad (10)$$

### 2.4 多分类器损失函数

构建有两个分类器,分别是相似性度量分类器和全连接层分类器。首先介绍相似性度量分类器,输入网络的支持集有  $N$  类,每类有  $K$  个样本,经过特征提取与注意力模块得到嵌入向量  $\mathbf{S}' = [(\mathbf{F}_{A_{s_1}}, \dots, \mathbf{F}_{A_{s_k}}), (\mathbf{F}_{A_{s_2}}, \dots, \mathbf{F}_{A_{s_{2k}}}), \dots, (\mathbf{F}_{A_{s_N}}, \dots, \mathbf{F}_{A_{s_{Nk}}})]$ ,依据文献[7]将同类样本的嵌入向量取平均值作为该类的表征,即类原型中心  $\mathbf{S}'_{\text{pro}} = [\mathbf{F}_{S_1}, \mathbf{F}_{S_2}, \dots, \mathbf{F}_{S_N}]$ ,度量模块输出查询集  $\mathbf{F}_{A_q}$  与  $N$  个类别的相似度即预测值,利用 softmax 函数求得相似性度量分类损失。相似性度量分类器的损失函数计算如下:

$$L_{\text{sim}} = - \sum_{i=1}^N y_i \ln \frac{\exp(C_{\text{sim}}(\mathbf{F}_{S_i}, \mathbf{F}_{A_q})/\tau)}{\sum_{j=1}^N \exp(C_{\text{sim}}(\mathbf{F}_{S_j}, \mathbf{F}_{A_q})/\tau)} \quad (11)$$

其中:  $y_i$  为真实值,只有正确类别  $y_i = 1$ ,其他  $N - 1$  个类别为  $y_i = 0$ ;  $\tau$  为温度系数,在 Mini-ImageNet 的实验中取 0.001,在其他数据集上取 0.01。

此外,针对特征提取模块,在其后构建了单层全连接分类器,以监督特征提取模块提取更为准确的特征。查询集经过特征提取模块得到特征图  $\mathbf{F}_Q$ ,通过全连接层,得到对每个类别的预测结果。全连接分类器的损失函数如下:

$$L_{\text{fc}} = - \sum_{i=1}^{C_{\text{train}}} y_i \ln \frac{\exp(\mathbf{W}_i^T \mathbf{F}_Q + \mathbf{b}_i)}{\sum_{j=1}^{C_{\text{train}}} \exp(\mathbf{W}_j^T \mathbf{F}_Q + \mathbf{b}_j)} \quad (12)$$

其中,  $\mathbf{W}_i$ 、 $\mathbf{b}_i$  分别是第  $i$  个类别的权值矩阵和偏置,一共  $C_{\text{train}}$  个类别。

此外,天基光学成像条件受限,数据集在姿态空间往往数量不均衡,且空间目标由于受姿态属性的影响图像表征差异很大,对于以总体分类精度为目标的网络而言,类内姿态不均衡势必会导致网络在训练时过多关注姿态属性密集的样本,

潜在增加了网络提取特征时的偏差。针对以上问题,引入一种基于邻域密度的样本标签权重来自适应均衡损失函数,利用标签信息提高网络对不平衡数据的识别性能。该方法通过计算每个样本在特征空间中的邻域密度,来确定每个样本在损失函数中的标签权重,降低邻域密度高的样本在损失函数中的占比,增加邻域密度低的样本对损失函数的占比,缓解了数据不均衡的问题。

训练数据集  $D_{\text{train}}$  某一类数据子集为  $D_i (i \in [1, C_{\text{train}}])$ ,其中一个样本的向量为  $\mathbf{x}$ ,欧式距离阈值为  $d$ ,可以用以下公式表示以  $\mathbf{x}$  为中心,同一类内与  $\mathbf{x}$  的欧式距离小于  $d$  的图像样本数量:

$$N(\mathbf{x}, d) = \sum_{\mathbf{y} \in D_i} I(r(\mathbf{x}, \mathbf{y}) \leq d) \quad (13)$$

其中,  $r(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{j=1}^m (x_j - y_j)^2}$  表示图像表征向量之间的欧式距离。  $I(\cdot)$  表示当  $r(\mathbf{x}, \mathbf{y}) \leq d$  时取 1,  $r(\mathbf{x}, \mathbf{y}) > d$  时取 0。以每幅图像为中心距离阈值内的样本数量与该图像在类内姿态属性空间的邻域密度成正比,与在均衡不同姿态样本后的损失函数中所占权重成反比,人为地纠正网络对多数样本的关注偏差,权重计算如下,其中  $\gamma$  是超参数:

$$w = \gamma \frac{1}{N(\mathbf{x}, d)} \quad (14)$$

最终网络的损失函数在基于邻域密度的样本标签权重的加权下,可表示为:

$$L = w \cdot (\lambda L_{\text{sim}} + L_{\text{fc}}) \quad (15)$$

式中,  $\lambda$  是两个分类器损失的权重比,在 Mini-ImageNet 数据集上设为 0.125,在其他数据集上设置为 0.75。

## 3 实验及结果分析

### 3.1 数据集和实验设置

#### 3.1.1 数据集

验证实验中使用了公开数据集 Mini-ImageNet、CUB-200 以及自制的三个空间目标数据集,具体信息如表 1 所示。

表 1 数据集信息

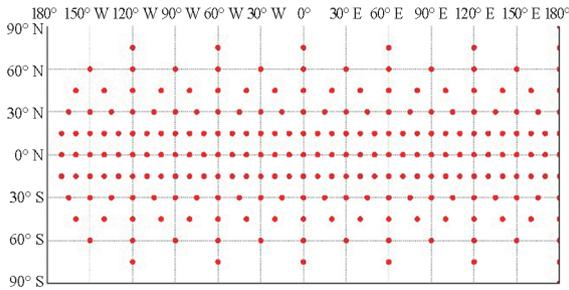
Tab. 1 Dataset information

数据集	数据数量	类别数量	平均类内样本数	训练/验证/测试类别数量
CUB-200	11 788	200	58	100/50/50
Mini-ImageNet	60 000	100	600	64/16/20
BUAA_unbalanced	2 000	20	100	10/5/5
NASA Satellite dataset1	10 800	36	300	20/8/8
NASA Satellite dataset2	7 200	36	200	20/8/8

CUB-200 数据集包含鸟类的 200 个子类,图像分辨率较高,是小样本识别常用的公开数据集,同时也是细粒度视觉分类任务广泛使用的数据集。

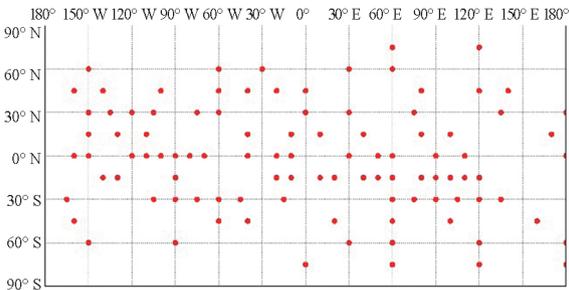
Mini-ImageNet 数据集是从 ImageNet 数据集抽取的 100 类组成的,常用于小样本图像分类,每幅图像的尺寸为 84 像素 × 84 像素。

BUAA-SID1.0 (以下简称 BUAA) 是由张浩鹏等<sup>[18]</sup>制作的空间目标全视点仿真图像序列,包含 20 颗卫星、230 个采样视点对应的灰度图像,图像分辨率为 320 像素 × 240 像素。为了验证所提方法对姿态不均衡数据集的效果,将 BUAA 中部分视点的图像删去,以得到姿态不均衡的数据集 BUAA\_unbalanced。依据文献[18]中用经纬度表示视点,如图 5 所示。



(a) BUAA 抽取前的 230 个视点

(a) 230 viewpoints before extraction on BUAA



(b) BUAA 抽取后的 100 个视点

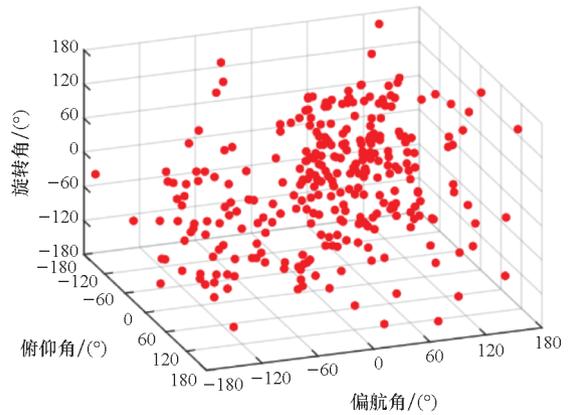
(b) 100 viewpoints after extraction on BUAA

图 5 数据集采样视点

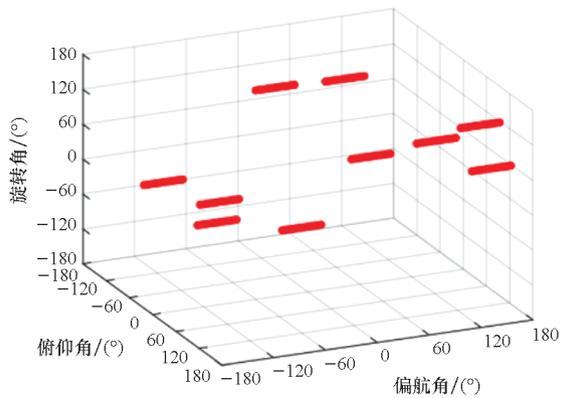
Fig. 5 Sampling viewpoints of the dataset

NASA Satellite 由于目前开源的空间目标数据集有限,利用 Blender 软件对从 NASA 官网获取的 36 个 NASA 卫星的 3D 模型进行渲染制作了两组数据集,第一组包含了随机生成的 300 个不均衡的姿态,第二组数据集遵循航天动力学轨道运动规律<sup>[19]</sup>模拟真实场景中空间目标与天基空间监视平台 10 次交会时成像图像中空间目标的姿态,每次交会成像 20 幅,共 200 个姿态。以下分别简称为 NASA Satellite dataset1、NASA Satellite dataset2。两组数据集的姿态属性用欧拉角来描

述,如图 6 所示。



(a) NASA Satellite dataset1



(b) NASA Satellite dataset2

图 6 空间目标数据集样本姿态

Fig. 6 Sample pose in space target dataset

### 3.1.2 实验设置

实验均在 Ubuntu20.04 平台、编程语言为 Python3.7、深度学习框架为 Pytorch1.11.0 版本、显卡为 NVIDIA A100 的环境下进行。采取 5-way 1-shot、5-way 5-shot 两种参数设置下的实验评估所提方法,最终精度以随机抽取的测试集 95% 置信区间的平均分类精度为准。采用随机梯度下降 (stochastic gradient descent, SGD) 作为优化器,在公开数据集上动量值为 0.9,权值衰减值设为 0.0005,初始学习率设置为 0.1;在空间目标数据集上权值衰减值设为 0.001,初始学习率设置为 0.05。在 5-way 1-shot 时迭代次数为 80,在第 60、70 次迭代时学习率分别降为 0.01、0.002;在 5-way 5-shot 时迭代次数为 60,在第 40、50 次迭代时学习率分别降为 0.01、0.002。Mini-ImageNet、CUB-200、空间目标数据集的 BatchSize 分别设为 64、32、128。在空间目标数据集上实验时, BUAA\_unbalanced、NASA Satellite dataset1、NASA Satellite dataset2 分别设置  $d = 600$ 、 $d = 3\ 000$ 、 $d = 4\ 500$ ,  $\gamma = 2$ 、 $\gamma = 4$ 、 $\gamma = 5$ ,公开数据集不引入基于邻域密度的样本标签权重。

### 3.2 实验及结果分析

#### 3.2.1 对比实验

表 2 ~ 4 展示了所提方法及目前主流的小样本识别方法在两个公开数据集和空间目标数据集 NASA Satellite dataset2 上的实验结果。其中, RN<sup>[8]</sup>和 DN4<sup>[20]</sup>是广泛应用于小样本学习任务中的经典方法,具有不错的性能表现; CAN<sup>[9]</sup>、RENet<sup>[10]</sup>和 DBRN<sup>[11]</sup>是基于相似性增强特征的代表方法; LMPNet<sup>[21]</sup>和 DeepEMD<sup>[22]</sup>、CGRN<sup>[23]</sup>是表现优异的前沿小样本识别方法。以上方法在公开数据集 Mini-ImageNet、CUB-200 上的实验精度均引自对应文献。

可以看出所提方法在 3 个数据集上的整体性能优于其他方法,在 CUB-200 数据集上,所提方法在 5-way 1-shot 和 5-way 5-shot 实验中分别高于最先进方法 1.65%、0.24%;在 Mini-ImageNet 数据集和 NASA Satellite dataset1 上,在 5-way 1-shot 实验均取得当前最好的结果,这说明所提方法在 5-way 1-shot 的实验设置下效果更好,能够在有限的样本中获得尽可能有效的信息,表现出良好的泛化能力和样本利用率。同时,与基于相似性融合特征的 CAN、RENet 和 DBRN 网络相比,提出的方法均取得了更好的结果,这表明自适应特征交叉融合机制可以在少数样本中更有效地提取到辨识性特征并增加特征关联的正确性,从而提高了识别精度。

#### 3.2.2 消融实验

为了验证基于自注意力交叉融合模块和基于互注意力交叉融合模块的有效性,在 CUB-200 数据集的 5-way 1-shot 实验中进行模块消融实验,结果如表 5 所示,仅使用基于自注意力交叉融合模块可以帮助网络提取更有辨识性的特征,因而分类精度提升了 0.79%。其中,“√”表示实验中用到了此模块。仅使用基于互注意力交叉融合模块可以帮助网络增加支持集和查询集的相似性,有小幅度的提升。当两个模块共同协作时,有效定位显著特征和挖掘图像之间的相似性相互辅助,网络的分类效果达到最好,提升了 1.4%。

表 2 不同方法在 CUB-200 数据集上的分类精度

Tab. 2 Recognition accuracy of different methods on the CUB-200 dataset

网络模型	%	
	5-way 1-shot	5-way 5-shot
DN4	46.84 ± 0.81	74.92 ± 0.64
DeepEMD	75.65 ± 0.83	88.69 ± 0.50
LMPNet	65.59 ± 0.13	68.19 ± 0.23
RENet	79.49 ± 0.44	91.11 ± 0.24
本文方法	<b>81.14 ± 1.02</b>	<b>91.35 ± 0.54</b>

表 3 不同方法在 Mini-ImageNet 数据集上的分类精度

Tab. 3 Recognition accuracy of different methods on the Mini-ImageNet dataset

网络模型	%	
	5-way 1-shot	5-way 5-shot
RN	50.44 ± 0.82	65.32 ± 0.70
DN4	51.24 ± 0.74	71.02 ± 0.64
CAN	63.85 ± 0.48	79.44 ± 0.48
DBRN	52.74 ± 0.87	68.43 ± 0.64
LMPNet	62.74 ± 0.11	<b>80.23 ± 0.52</b>
CGRN	50.85 ± 0.86	64.13 ± 0.70
本文方法	<b>64.62 ± 1.01</b>	79.57 ± 0.71

表 5 在 CUB-200 数据集上的模块消融实验

Tab. 5 Ablation experiment on the CUB-200 dataset

自注意力模块	互注意力模块	分类精度/%
		79.697
√		80.487
	√	79.830
√	√	<b>81.140</b>

表 4 不同方法在 NASA Satellite dataset2 数据集上的分类精度

Tab. 4 Recognition accuracy of different methods on the NASA Satellite dataset2

网络模型	%	
	5-way 1-shot	5-way 5-shot
DN4	73.94 ± 0.74	<b>92.64 ± 0.43</b>
RN	73.70 ± 0.81	86.44 ± 0.60
CAN	80.49 ± 1.56	84.44 ± 1.30
DeepEMD	80.34 ± 0.25	92.24 ± 0.49
本文方法	<b>82.59 ± 0.88</b>	91.90 ± 0.66

此外,为了验证基于邻域密度的样本标签权重对不平衡数据集有效,分别在三组姿态不平衡的空间目标数据集上进行权重消融实验。数据集中部分空间目标的图像表征受姿态影响较大,导致该类别的分类精度较低,为了观察引入基于邻域密度的损失权重对这种姿态不平衡是否有效,除了分类精度,还引入宏平均  $F_1$  来衡量网络对各个类别空间目标的综合分类性能。多分类任务中常用  $F_1$  来综合评价准确率  $P$  和召回率  $R$  两个指标,宏平均  $F_1$  是对所有类别的  $F_1$  平均求得,反映了各个类别的综合性能,计算公式如下:

$$F_1 = \frac{2PR}{P+R} \quad (16)$$

在5-way 1-shot 和5-way 5-shot 数据集上进行消融实验,结果如表6~7所示。从表6和表7可以看出,提出的自适应损失均衡化方法在不同数据的空间目标识别任务中均取得了较好的性能提升。在BUAA\_unbalanced、NASA Satellite dataset1、NASA Satellite dataset2三个数据集5-way 1-shot 的实验中,精度分别提升了3.43%、1.82%、0.99%,宏平均  $F_1$  也相应提升了2.65%、1.89%、2.34%。此外,在5-way 5-shot 实验中,精度分别提升了1.283%、1.767%、1.483%,宏平均  $F_1$  提升2.0%、3.08%、4.12%。可以看出,通过引入邻域密度信息可以自适应损失均衡化,使网络更加关注于邻域稀疏区域的样本,进而提高了网络的识别性能。

表6 在空间目标数据集上5-way 1-shot 的权重消融实验  
Tab.6 Weight ablation experiment in 5-way 1-shot task on space target datasets

数据集	%			
	未加权重		加权重	
	精度	宏平均 $F_1$	精度	宏平均 $F_1$
BUAA_unbalanced	70.770	59.87	74.200	62.52
NASA Satellite dataset1	81.763	64.99	83.583	66.88
NASA Satellite dataset2	81.603	65.31	82.593	67.65

表7 在空间目标数据集上5-way 5-shot 的权重消融实验  
Tab.7 Weight ablation experiment in 5-way 5-shot task on space target datasets

数据集	%			
	未加权重		加权重	
	精度	宏平均 $F_1$	精度	宏平均 $F_1$
BUAA_unbalanced	87.110	78.73	88.393	80.73
NASA Satellite dataset1	88.690	75.44	90.457	78.52
NASA Satellite dataset2	90.420	78.58	91.903	82.70

四组实验训练时的损失函数曲线图如图7所示,随着迭代次数的增加损失值呈下降趋势,在一定迭代次数之后趋于稳定,可以证明本文模型的有效性。还提供了所提方法在NASA Satellite dataset2数据集5-way 5-shot 的实验中未加权重与加权重情况下的分类结果混淆矩阵图,如图8所示,可以

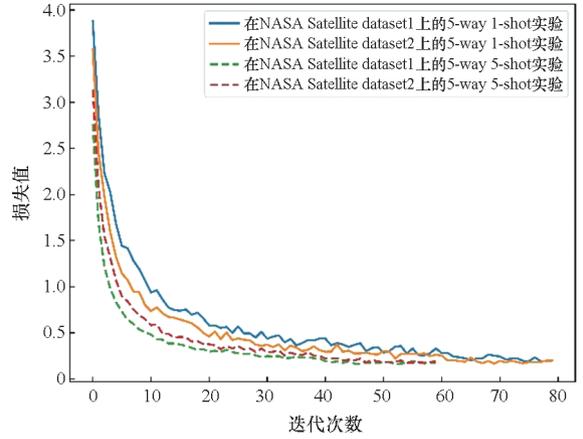
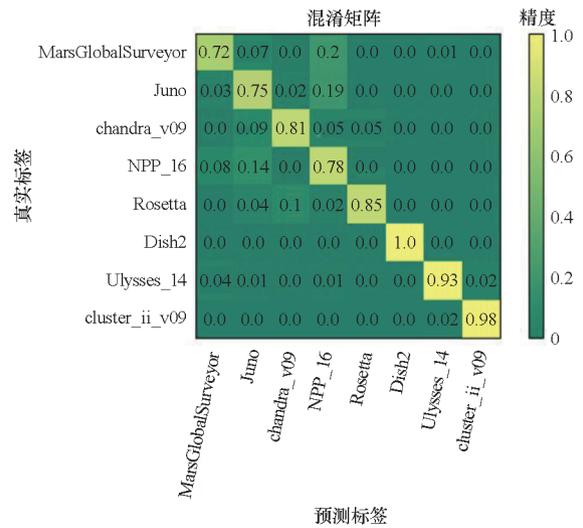


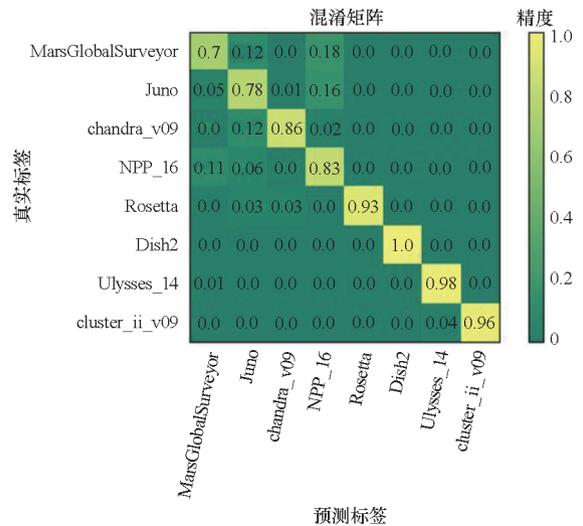
图7 模型训练损失函数曲线图

Fig.7 Curve of model training loss function



(a) 未加基于邻域密度的损失权重

(a) Without weights based on neighborhood density



(b) 加基于邻域密度的损失权重

(b) With weights based on neighborhood density

图8 本文方法在NASA Satellite dataset2数据集上的分类结果混淆矩阵图

Fig.8 Confusion matrix diagram of the proposed method on NASA Satellite dataset2

直观地看到加权重之后网络对不同类别之间的混淆有一定的减少,虽然对 MarsGlobalSurveyor 和 cluster\_ii\_v09 这两类卫星的分类精度有微弱下降,但其余类别的分类精度均有所提高,验证了所提出方法的有效性。

### 3.2.3 样本数量实验

为了探索提出小样本方法的目标识别性能与样本数量的关系,在 NASA Satellite dataset2 数据集 5-way 的实验中设置不同的  $K$ -shot 对模型进行训练并测试,识别精度在表 8 中给出。

表 8 在 NASA Satellite dataset2 数据集 5-way  $K$ -shot 的性能比较

Tab. 8 Performances comparison with 5-way  $K$ -shot on NASA Satellite dataset2

$K$ -shot	精度	%
1-shot	81.603	
5-shot	90.420	
10-shot	92.723	
15-shot	95.843	
20-shot	96.457	
30-shot	95.630	

从表 8 可以看出,模型的分类精度随着每类的样本数量增加而提高,在 15-shot 时开始呈现一

个相对稳定的趋势。可以看出每类 15 个样本可以提供足够的类别信息,能够比较准确地代表整个类别,得到一个理想的精度。

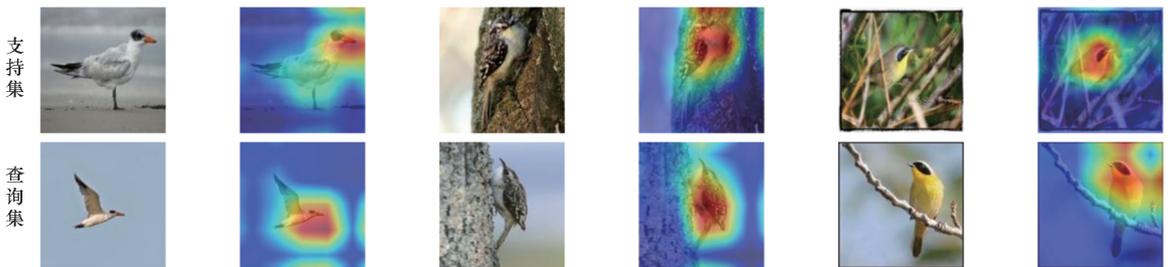
### 3.2.4 可视化实验

为了更加清晰地展示自适应局部特征交叉融合模块对模型性能的提升,在多组数据集 5-way 1-shot 的实验结果中选取了三对支持集和查询集经过交叉融合后的特征图,并通过双线性插值得到的输入图像大小的特征热力图,如图 9 所示。在这些特征热力图中,不同的颜色代表不同强度的特征响应,红色表示某个区域在样本分类中更加重要,蓝色则表示该区域分类贡献较小,可视化地展示了自适应局部特征交叉融合模块能够更好地捕捉支持集和查询集之间相似的辨识性特征。在 Mini-ImageNet 数据集中,蚂蚁、汽车的形态甚至数量变化后,通过局部特征之间信息的交互,网络依旧可以找到相似的目标部分;在 CUB-200 数据集上,网络更加关注的是鸟类的喙、斑纹等更具辨识性的特征,即使在鸟类的姿态或者背景发生变化时依然能正确识别。在空间目标数据集中,由于图像表征中同一卫星姿态的不同,帆板或者主体部分的形态、纹理也会发生变化,但是网络能够自适应地关注一些变化较小的部件,从而实现目标的有效识别。综上所述,所提出的方法在各数据集中都取得了出色的性能表现,充分验证了提取辨识性特征以及提高特征关联准确性的优势。



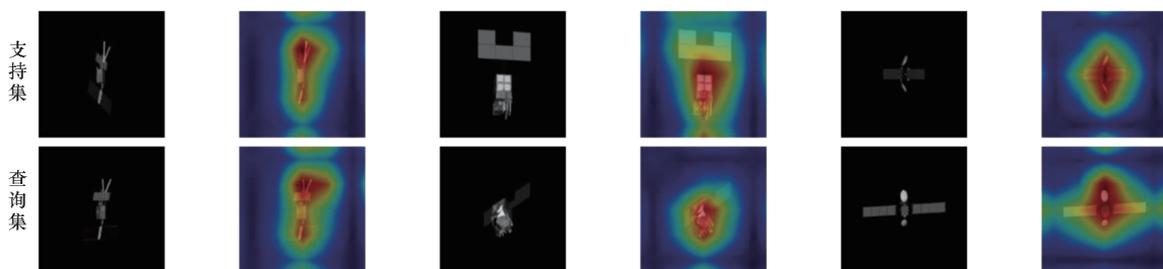
(a) Mini-ImageNet 数据集上的可视化结果

(a) Visualization of the proposed method on Mini-ImageNet dataset



(b) CUB-200 数据集上的可视化结果

(b) Visualization of the proposed method on CUB-200 dataset



(c) BUAA\_unbalanced 数据集上的可视化结果

(c) Visualization of the proposed method on BUAA\_unbalanced dataset

图9 本文方法在多个数据集上的可视化结果

Fig.9 Visualization of the proposed method on different datasets

## 4 结论

在低频次观测的空间目标识别的背景下,针对现有的小样本学习方法无法适应空间目标数据集类内表征差异大且姿态属性不均衡的问题,提出了一种自适应特征交叉融合的空间目标小样本识别方法,并进行了所提方法与主流小样本识别方法的详细对比实验,以及模块消融实验。结果表明,提出的方法能够显著提升网络提取特征的辨识性和姿态鲁棒性,并且能够缓解空间目标数据集姿态属性不均衡造成的网络学习偏差问题。在空间目标识别领域,本文方法的提出为解决小样本识别的问题提供了新的思路和方法。

## 参考文献 (References)

- [1] ZENG H Y, XIA Y. Space target recognition based on deep learning [C]//Proceedings of the 2017 20th International Conference on Information Fusion (Fusion), 2017: 1-5.
- [2] 李林泽,张涛. 基于深度学习的空间非合作目标特征检测与识别[J]. 智能系统学报, 2020, 15(6): 1154-1162.  
LI L Z, ZHANG T. Feature detection and recognition of spatial noncooperative objects based on deep learning [J]. CAAI Transactions on Intelligent Systems, 2020, 15(6): 1154-1162. (in Chinese)
- [3] YANG X, WU T, WANG N N, et al. HCNN-PSI: a hybrid CNN with partial semantic information for space target recognition [J]. Pattern Recognition, 2020, 108: 107531.
- [4] YANG X, NAN X T, SONG B. D2N4: a discriminative deep nearest neighbor neural network for few-shot space target recognition [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(5): 3667-3676.
- [5] KOCH G, ZEMEL R, SALAKHUTDINOV R. Siamese neural networks for one-shot image recognition [C]//Proceedings of the 32nd International Conference on Machine Learning, 2015.
- [6] VINIYALS O, BLUNDELL C, LILLICRAP T, et al. Matching

networks for one shot learning [J]. Advances in Neural Information Processing Systems, 2016: 3637-3645.

- [7] SNELL J, SWERSKY K, ZEMEL R. Prototypical networks for few-shot learning [EB/OL]. (2017-06-19) [2023-04-21]. <https://arxiv.org/pdf/1703.05175>.
- [8] SUNG F, YANG Y X, ZHANG L, et al. Learning to compare: relation network for few-shot learning [C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 1199-1208.
- [9] HOU R B, CHANG H, MA B P, et al. Cross attention network for few-shot classification [C]//Proceedings of the 33rd International Conference on Neural Information Processing Systems, 2019: 4003-4014.
- [10] KANG D, KWON H, MIN J H, et al. Relational embedding for few-shot classification [C]//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 8802-8813.
- [11] TONG Y F, TIAN H, JIANG X R, et al. Dual branch relation network with feature weighting for few-shot learning [C]//Proceedings of the 2021 7th International Conference on Computer and Communications (ICCC), 2021: 1743-1751.
- [12] REN L, DUAN G D, HUANG T X, et al. Multi-local feature relation network for few-shot learning [J]. Neural Computing and Applications, 2022, 34(10): 7393-7403.
- [13] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 6000-6010.
- [14] SUN J M, SHEN Z H, WANG Y A, et al. LoFTR: detector-free local feature matching with transformers [C]//Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 8918-8927.
- [15] PETIT O, THOME N, RAMBOUR C, et al. U-Net transformer: self and cross attention for medical image segmentation [C]//Proceedings of the Machine Learning in Medical Imaging: 12th International Workshop, 2021: 267-276.
- [16] CHEN X, YAN B, ZHU J W, et al. Transformer tracking [C]//Proceedings of the 2021 IEEE/CVF

- Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 8122 – 8131.
- [17] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [ C ] // Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770 – 778.
- [18] 张浩鹏, 刘正一, 姜志国. BUAA-SID1.0 空间目标图像数据库 [ J ]. 航天返回与遥感, 2010, 31(4): 65 – 71.
- ZHANG H P, LIU Z Y, JIANG Z G. BUAA-SID1.0 space object image dataset [ J ]. Spacecraft Recovery & Remote Sensing, 2010, 31(4): 65 – 71. (in Chinese)
- [19] DE RUITER A H, DAMAREN C, FORBES J R. Spacecraft dynamics and control: an introduction [ M ]. Hoboken, USA: John Wiley & Sons, 2012.
- [20] LI W B, WANG L, XU J L, et al. Revisiting local descriptor based image-to-class measure for few-shot learning [ C ] // Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 7253 – 7260.
- [21] HUANG H W, WU Z K, LI W B, et al. Local descriptor-based multi-prototype network for few-shot learning [ J ]. Pattern Recognition, 2021, 116: 107935.
- [22] ZHANG C, CAI Y J, LIN G S, et al. DeepEMD: few-shot image classification with differentiable earth mover's distance and structured classifiers [ C ] // Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 12200 – 12210.
- [23] JIA X, SU Y L, ZHAO H. Few-shot learning via relation network based on coarse-grained granulation [ J ]. Applied Intelligence, 2023, 53(1): 996 – 1008.