

采用动态上采样的弱监督红外小目标检测

赵伟鸣¹ 郑伟¹

¹(中国科学院国家空间科学中心 北京 100190)

摘要 针对红外小目标尺寸小、信噪比低、对比度弱、易受背景杂波干扰而导致的检测难题，本文研究了一种基于动态上采样的弱监督红外小目标检测方法。传统方法依赖复杂手工特征，鲁棒性不足；现有深度学习方法虽性能优越，但需大量像素级标注数据，成本高昂。为此，本文提出一种仅需单点标注的弱监督检测框架。该方法结合标签演化机制，利用卷积神经网络中间预测结果逐步将点标签扩展为像素级目标掩码，显著降低标注负担。同时，引入动态上采样器，通过学习自适应采样偏移与动态范围因子，改善特征重建质量，减少上采样伪影，提升检测精度。在 NUAA-SIRST、NUDT-SIRST 与IRSTD-1K 混合数据集上的实验表明，所提方法在弱监督下可达到全监督模型 70% 以上的交并比 (IoU) 与 90% 以上的检测概率 (Pd)，动态上采样进一步带来 5-10% 的性能提升，并加快模型收敛。本研究为低标注成本下的红外小目标检测提供了一种高效、鲁棒的解决方案。

关键词 红外小目标；目标检测；卷积神经网络；弱监督训练

中图分类号 TP391.4

Weakly Supervised Infrared Small Target Detection Using Dynamic Upsample

ZHAO Yiming¹ ZHENG Wei¹

¹ (National Space Science Centre, Chinese Academy of Sciences, Beijing 100190)

Abstract Infrared small target detection is a critical yet challenging task in various applications such as surveillance, remote sensing, and space situational awareness, due to the targets' small size, low signal-to-noise ratio, weak contrast, and susceptibility to background clutter. Traditional detection methods rely heavily on handcrafted features and meticulous hyperparameter tuning, which often lack adaptability and robustness in complex real-world scenarios. Recent advances in deep learning, particularly Convolutional Neural Networks (CNNs), have shown promising capabilities in feature representation and have been introduced to infrared small target detection. However, most CNN-based approaches require a large amount of pixel-level annotated data for training, which is costly, time-consuming, and often impractical to acquire, especially for small and subtle targets. To address this limitation, this paper proposes a novel weakly supervised framework for infrared small target detection using only single-point annotations. Our method integrates a label evolution mechanism that progressively expands sparse point labels into pixel-level masks by leveraging intermediate CNN predictions. This iterative process mimics the natural "mapping degeneration" phenomenon observed during network training, where the model initially segments target regions coarsely before refining predictions toward precise point locations.

Additionally, we introduce a dynamic upsampling module to replace conventional upsampling operations. This module learns adaptive sampling offsets and incorporates a dynamic range factor to reduce artifacts and improve feature reconstruction fidelity during the decoding stage, thereby enhancing the overall detection accuracy and boundary delineation. The proposed framework is built upon a Dense Nested Attention Network (DNANet) backbone and is evaluated on a composite dataset consisting of NUAA-SIRST, NUDT-SIRST, and IRSTD-1K. Experimental results demonstrate that our approach significantly reduces annotation effort while achieving competitive performance. Under weak supervision, the model attains over 70% of the Intersection-over-Union (IoU) and over 90% of the Detection Probability (Pd) compared to its fully supervised counterpart. The dynamic upsampler further boosts IoU and Pd by 5-10% and accelerates model convergence. Ablation studies confirm the effectiveness of both the label evolution strategy and the dynamic upsampling module across different network architectures. This work contributes a practical and efficient solution for infrared small target detection under low-labeling-cost conditions, offering improved adaptability and robustness suitable for real-world applications where extensive annotation is infeasible.

Key words Infrared small target; target detection; convolutional neural network (CNN); weakly supervised training.

1、引用

红外小目标检测是红外搜索与跟踪（Infrared Search and Track,IRST）系统中的一项关键技术，长期以来因其复杂性和挑战性备受关注。红外小目标通常具有尺寸小、信噪比低、对比度弱等特征，且易被复杂的背景杂波（如云层、地表辐射等）所淹没，这使得其检测任务异常困难。然而，随着红外成像技术的快速发展，红外小目标检测在现代社会中展现出广泛的应用价值，包括交通监控[1]、海上救援[2]、军事监视[3]以及航空航天[4]等领域。特别是在空间科学领域，红外小目标检测技术对空间目标监测、太空碎片追踪以及深空探测任务具有重要意义。

在空间科学中，红外小目标检测的需求日益凸显。例如，在空间目标监测中，红外成像技术可用于识别和跟踪近地轨道

上的卫星、火箭残骸以及其他潜在威胁目标[5]。此外，随着太空活动的增加，太空碎片的数量急剧上升，对航天器的安全构成严重威胁。红外小目标检测技术能够有效识别和追踪这些微小碎片，为航天器的轨道规划和避障提供关键支持。在深空探测任务中，红外探测器常用于捕捉遥远天体的微弱信号，如小行星、彗星等，这些目标的检测对研究太阳系起源和演化具有重要意义[6]。然而，深空环境的复杂性和目标信号的微弱性使得检测任务极具挑战性。

传统的红外小目标检测方法主要依赖于复杂的手工特征设计和精心调整的超参数[7][8][9]。这些方法通常基于图像处理和模式识别技术，如滤波器设计[10]、背景抑制[11]、运动检测和能量检测[12]等。例如滤波器方法通过设计匹配滤波器来增强目标信号，但需要针对不同的应用场景反复

调整参数。背景预测方法试图建立背景模型从而检测前景目标，但在背景复杂多变的情况下效果不佳。虽然这些方法在特定条件下取得了一定的成功，但其适应性和鲁棒性较差。当面对变化多端的环境和复杂的干扰时，检测性能往往显著下降。此外，手工特征的设计过程繁琐，需要大量的领域专家知识，难以适应新兴的检测需求。

近年来，随着深度学习技术的迅猛发展，研究人员开始尝试将其应用于红外小目标检测。深度学习[1][13][14]，特别是卷积神经网络，在计算机视觉领域展现出了强大的特征学习和表示能力[15]。其通过端到端的学习方式，能够自动从大量数据中提取有用的特征，避免了复杂的手工特征设计流程。这为红外小目标检测带来了新的希望。然而，一个深度学习方法通常需要大量的像素级标注数据来进行模型训练[16]，才可以得到一个鲁棒性和适应性比较高的模型。这意味着需要对大量的红外图像进行目标标注，以指导模型学习目标特征。这一过程不仅成本高昂，需要投入大量的人力和物力资源，而且耗时费力，难以及时满足实际应用的需求。

此外，红外小目标本身的特性也增加了数据标注的难度。由于目标尺寸小、特征不明显，对标注人员的注意力和专业性很高，稍有不慎就可能遗漏目标或产生误标。这进一步提高了高质量标注数据的获取门槛。数据不足的问题限制了深度学习模型的泛化能力，导致其在实际应用中的表现不尽如人意。

本研究的主要目标是通过单点监督实现红外小目标检测，利用一种标签演化框架[17]，并结合动态上采样器优化网络结构

[18]。具体来说，利用 CNN 的中间预测逐步扩展点标签，最终获得像素级目标掩码，从而在减少标注成本的同时，保持较高的检测性能。

2、相关工作

有关红外小目标检测的问题经历了几十年的发展。最开始的传统数学方法主要从三个方面设计模型包括基于滤波的方法，基于局部对比度的方法和基于低秩的方法。最近的深度学习的思路更加丰富。与需要精心设计的模型和仔细调整的超参数的传统方法相比，卷积神经网络可以以数据驱动的方式学习输入图像和 GT 标签之间的非线性映射，从而更好地泛化到真实复杂场景。越来越多的注意力放在红外小目标解决方案上。具体来说，Dai 等人[19]专门设计了一个不对称上下文模块，并进一步结合局部对比度测量以提高目标对比度。Li 等人[14]通过重复特征融合保留了目标信息。Zhang 等人[1]聚合边缘信息以实现形状感知的 SIRST 检测。Zhang 等人[2]探索了跨层相关性并采用基于变换器的方法来预测准确的目标掩码。Wu 等人[20]定制了一个 UIU-Net 框架，用于多层次和多尺度特征聚合。现有工作中心集中网络模型架构设计上，以追求在完全监督方式下的性能。然而，现有的深度学习确实减少了对数学模型的依赖，但它们却需要大量精确标注的数据来支撑模型的训练。如果缺乏足够的高质量数据，模型可能会出现过拟合或无法有效泛化的问题，无法充分利用深度模型的整合能力。

为了降低标注成本并提高检测效率，寻找可用的弱监督方法显得尤为重要[21][22]。弱监督学习指的是利用不完全、不精确或部分标注的数据来训练模型，而不是

完全监督要求数据集精确标注。使用弱监督方法不仅可以降低标注成本，还能够提升检测效率。由于减少了对大量精确标注数据的需求，数据准备的周期大大缩短，模型可以更快速地迭代和部署。此外，弱监督方法还有助于提高模型的鲁棒性和泛化能力。通过接触更多的未标注或弱标注数据，模型能够学习到更为广泛的特征和更加复杂的背景变化。相较于完全监督，弱监督得出的模型往往准确率比较差。在面对复杂多变的实际环境时，要使用弱监督模型的同时要保持良好的检测性能。

本文采用的弱监督分割是点级弱监督分割[21][23]。图像分割的弱监督分割有着丰富的实例。具体来说，Bearman 等人[24]首次尝试在点级监督训练损失函数中引入客观性潜力，以提升分割性能。Qian 等人[25]通过距离度量损失利用几个标记点的语义信息来实现场景解析。Zhang 等人[26]提

出了一种内外指导方法，通过五个精心设计的点击实现实例分割。Cheng 等人[13]设计了在框注释内提供十个随机采样的二进制点注释用于实例分割。Ying 等人[17]在端到端方式下充分利用红外小目标的局部对比度先验，通过单个粗略点逐步演化伪掩码。

3、方法与理论

3.1 整体架构

对于给定输入图像和相应的 GT 点标签，我们使用 DNANet[27] 作为基线 SIRST 检测网络进行训练。基于 U-Net 架构改进的密集嵌套网络 DNANet 主要包含三个核心组件，输入信息依次通过特征提取模块、特征金字塔融合模块、以及八连通域聚类算法最后生成检测结果，其创新性网络架构设计显著提升了多尺度目标检测性能。网络架构示意图如图 1 所示

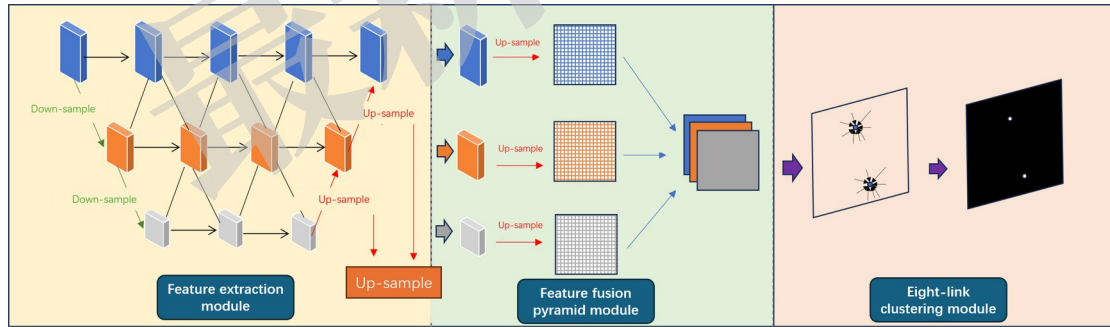


图 1 整体网络架构

Fig. 1 Overall network architecture

首先，网络采用独特的密集嵌套交互模块(DNIM)作为特征提取主干，该模块通过堆叠多级 U 形子网构建层次化特征空间。每个子网通过编码器-解码器结构分别处理不同感受野的特征信息，其中浅层子网专注于捕捉微小目标的精细纹理，而深层子

网则负责提取大尺寸目标的语义特征。这种分层处理机制有效解决了传统单路径网络在跨尺度目标检测中的表征局限性。

在特征融合机制上，DNIM 模块突破了传统跳跃连接的线性传输模式，创新性地构建了密集交叉连接网络。具体而言，在编码器与解码器的特征传输路径中设置了多个交互节点，通过横向连接实现跨层特征的交织融合。每个节点不仅接收本层特征，还整合了来自相邻层级的特征信息，形成类似神经网络突触连接的密集交互结构。这种多层次特征融合机制具有两个显著优势：其一，通过特征复用机制增强了对微小目标的表征能力，有效抑制了深层网络中的特征稀释现象；其二，构建了丰富的特征梯度，使网络能够自适应地调整不同尺度目标的关注权重。

完成特征提取模块之后，网络使用了一个特征金字塔融合模块来聚合生成的多层特征。首先将各层特征图经过上采样到相同尺寸。然后，将具有丰富空间和剖面信息的浅层的高分辨率特征和具有丰富抽象语义的深层特征采用通道拼接的方式整合在一起，最终生成包含丰富空间细节和深度语义的全局特征图。

最后，在目标定位阶段，引入八连通邻域聚类算法对所有像素点进行目标分割，计算每个目标的质心。首先对特征图进行阈值分割得到二值图像，随后扫描所有有效像素点：对于每个激活像素，系统会递归检查其八个邻域方向，将满足灰度相似性条件的相邻像素归并为同一目标区域。完成全图扫描后，系统通过计算各连通区域的一阶矩确定目标质心坐标。

3.2 动态上采样方法

3.2.1 初始采样位置

在 pytorch 内置上采样函数中，对于 X 中的一个点，其 s^2 个采样位置都固定分配

在了相同的初始位置（即特征图 X 中的标准网格点）。这种上采样方法忽略了原始特征图 X 邻近点之间的位置关系，当图片放大时，补充的像素知识简单的由器附近

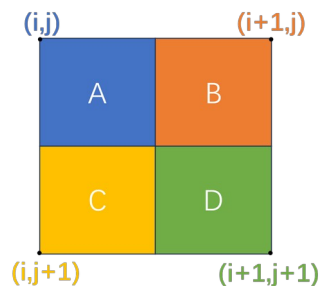


图 2 最临近插值法

Fig. 2 The nearest interpolation

的颜色填充，填充效果如图 2 所示。导致初始采样位置分布不均匀，出现明显的锯齿，画面过度不平滑。如果生成的偏移量都是零，则上采样特征相当于最近邻插值得到的特征。因此，这种初步初始化可以称为“最邻近插值法”。

针对这个问题，我们将初始位置更改为双线性初始化，双线性插值就是做两次线性变换，先在 X 轴上做一次线性变换，求出每一行的 R 点，再通过一次线性变换求出在该区域中的 P 点，

$$f(x, y) \approx \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2) \quad (1)$$

通过这种插值方法可以保留原始特征图临

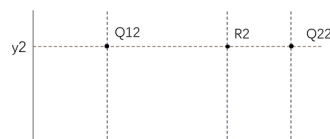
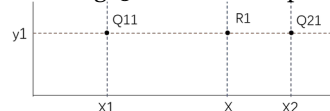


图 3 双线性插值

Fig. 3 Bilinear interpolation



近点的位置关系，从而使上采样得到的图片过渡更加平滑。

3.2.2 动态范围因子

经过最后归一化层的处理，输出特征的值通常位于以 0 为中心的 $[-1,1]$ 范围内。因此，在上述初始采样点生成的过程中，局部 s^2 个采样位置的偏移范围可能会显著重叠，如图 4 所示。

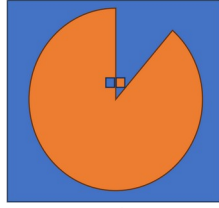


图 4 偏移范围重叠

Fig. 4 Offset range overlap

这种重叠通常是不均匀的，会影响到对边界预测，并且此类错误会逐阶段传播并导致预测伪影，如图 5 所示。

为了减少伪影的出现，可以通过将偏移量乘以设置好的“静态范围因子”，通

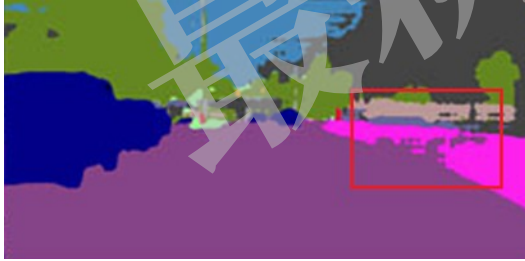


图 5 预测伪影

Fig. 5 Prediction artifacts

过处理来尽量满足重叠与非重叠之间的理

论边界条件。为了增加偏移量的灵活性，可以进一步通过线性投影输入特征来生成每个点的“动态范围因子”。通过使用 sigmoid 函数和预设 0.5 的静态因子，动态范围的取值在以 0.25 为中心的 $[0,0.5]$ 范围内，与静态范围相比，更好的满足重叠与非重叠之间的理论边界条件。

3.2.3 偏移生成过程

给定一个大小为 $C \times H \times W$ 的特征图 X ，首先使用线性投影来产生大小为 $2gs^2 \times H \times W$ 的偏移量 O 。

$$O = \lim_{ear(X)}, \quad (2)$$

然后对集合进行像素重排以满足空间大小 $2 \times sH \times sW$ 的尺寸。采样集 S 是偏移量 O 与原始采样网格 G 的和。

$$S = G + O. \quad (3)$$

最后使用 `grid_sample` 函数使用采样集 S 中的位置对之前的输入特征 X 进行重新采样，生成大小为 $C \times sH \times sW$ 的上采样特征图 X' ，

$$X' = \text{grid_sample}(X, S). \quad (4)$$

动态上采样流程如图 6 所示。

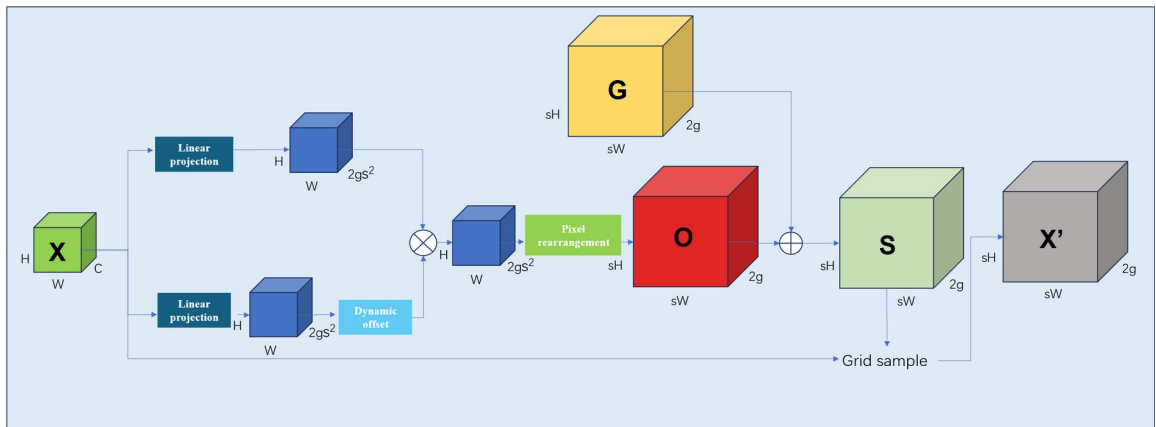


图 6 动态上采样过程

最新录用

3.3 标签演化网络

3.3.1 映射退化现象

在单点标签监督下，卷积神经网络通常会在初期以较低的置信度对目标周围的像素簇进行分割（区域-区域映射），随后逐步收敛到高置信度的真值点标签预测（区域-点映射）[17]。这种“映射退化”现象的原因可以归结为以下几点：

1) 红外系统的成像特点，目标仅包含强度信息，缺乏结构和纹理细节，导致目标区域内像素高度相似；

2) 红外小目标的局部对比度较高，目标区域的像素与背景杂波之间形成明显对比；

3) 卷积神经网络倾向于先学习简单的映射，然后逐步过渡到更复杂的映射。

相比区域-点映射，区域-区域映射较为简单，因此往往作为区域-点映射的过渡阶段。前两个原因解释了点标签外的区域映射产生的原因，而最后一个因素则导致映射的退化的现象。

3.3.2 标签演化流程

利用“映射退化”现象，可以建立一

个单点监督下的标签演化框架网络。首先学习将目标附近的像素簇分割出来，并以点标签为监督信号，通过标签更新过程，当前网络的预测结果用来调整点标签。接着，更新后的标签将作为监督信号继续训练网络，直到进行下一轮标签更新。通过反复迭代标签更新，网络逐步优化预测结果，最终生成与真实目标位置高度一致的像素级掩模。在这个过程中，网络会经历从区域到点的映射退化，通过候选像素提取、误报消除和加权求和等手段，不断提升标签质量，最终实现精确的端到端目标检测。

为了便于理解，采用 1 维曲线说明标签演化框架的具体流程。如下图 7 所示，虚线左侧的子图表示网络预测结果，其中黑色曲线表示标签更新框架内的网络预测结果，而灰色曲线表示网络在没有标签演化下的网络预测结果。在虚线的右侧，第一列和第二列分别表示当前标签和更新的标签，黑色箭头表示每一轮标签更新。标签更新总体框架可概括如下：在点标签作为监督的情况下，在第 1 轮标签更新中，当前的网络预测 1st Pred 用于更新当前点标签 Point Label 以生成第 1 个更新标签 1st

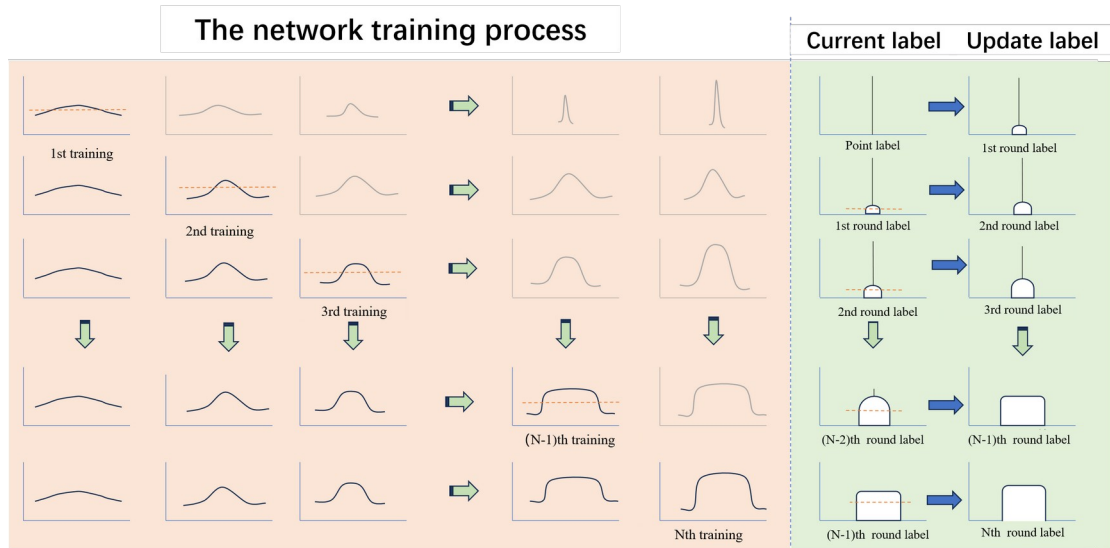


图 7 标签演化框架

Fig.7 Label evolution framework

Label。随后该标签用于监督网络训练，直到第 2 轮标签更新。通过迭代标签更新和网络训练，网络预测最终可以近似更新后

4、实验

4.1 数据集与评价指标

我们在训练过程中使用的数据集和之后的测试集，都是整合了三个公开数据集 NUAA-SIRST[2]、NUDT-SIRST [14] 和IRSTD-1K[1]的部分内容，具体整合的图像种类和数量在下表 1 中展示。

我们使用了两个像素级指标（即交并比 (IoU) 和像素精度 (PA)）和两个目标级指标（即检测概率 (Pd) 和误报率 (Fa)）进行性能评估。其中 IoU 是预测区域与真实区域的重叠面积与它们的并集面积的比值；PA 是正确分类的像素数占总像素数的比例；Pd 是模型正确检测到的目标数占真实目标总数的比例；Fa 是模型错误检测的目标数占所有检测结果的比例。其中 TP：真阳性数，在 label 中为阳性，在预测值中也为

的伪掩码标签，并且网络可以同时学习以端到端的方式实现像素级红外小目标检测。

阳性的个数；TN：真阴性数，在 label 中为阴性，在预测值中也为阴性的个数；FP：假阳性数，在 label 中为阴性，在预测值中为阳性的个数；FN：假阴性数，在 label 中为阳性，在预测值中为阴性的个数。

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{TP}{TP + FP + FN}, \quad (5)$$

$$PA = \frac{TP + TN}{TP + TN + FP + FN}, \quad (6)$$

$$Pd = \frac{TP}{TP + FN}, \quad (7)$$

$$Fa = \frac{FP}{TN + FP}. \quad (8)$$

表 1 数据集详细信息

Table1 datasets

数据集名称	图片类型	背景	标签类型	图片数目
NUAA-SIRST	实景	云/城市/海	手工粗标注	427
IRSTD-1K	合成	云/城市/河流	手工粗标注	1000
NUDT-SIRST	合成	云/城市/海/田野	实景准确标注	1327

4.2 实验细节

在训练期间，所有输入图像都被标准化，并随机裁剪成 256×256 大小的图像块作为网络输入。为扩充训练集，通过随机翻转和旋转来增强训练数据。为了可以动

态调整每个参数的学习率，所有网络都通过 Adam 方法进行学习率的动态优化。学习率最初设置为 5×10^{-4} ，并在第 100 个周期时减少十倍。批量大小被设置为 16。我们在 200 个周期后停止训练。所有模型都是在带有 Nvidia GeForce 4090D GPU 的

PC 上使用 PyTorch 实现的。

5.分析

5.1 标签演化框架分析

我们比较了在弱监督下，分别使用和不使用标签演化流程的 DNA-Net [27]在数据集上的平均结果,实验结果如表 2 所示。通过全监督评价指标作为对照，不使用标签演化流程的 DNA-Net 网络的评价指标很低，几乎无法正常实现对目标的检测，这表明在弱监督下，对红外小目标检测是一项具有挑战性的任务。相比之下，使用标签演化流程的 DNA-Net 网络在弱监督下的

所有评级指标都有着显著提高性能，在部分指标中甚至拥有接近全监督对应物的效果。

通过对标签演化过程中的标签抽检可以发现，在标签演化的过程中，生成的目标掩码相较于人工标注的目标掩码尺寸更不规则，且范围不稳定，存在标注框过大或者目标标注不完全的现象，这会导致有错误的特征信息被使用或者特征信息提取不完全的现象，这可能是使用标签演化网络后的检测效果不如全监督下的检测效果的原因。

表 2 不同方法实验结果对比

Table 2 Comparison of experimental results of different methods

Method	弱监督			全监督		
	IoU/ $\times 10^2$	Pd/ $\times 10^2$	Fa/ $\times 10^6$	IoU/ $\times 10^2$	Pd/ $\times 10^2$	Fa/ $\times 10^6$
DNA-Net	2.92	48.31	0.19			
DNA-Net +	57.86	92.17	21.74	71.45	95.18	22.36

表 3 不同方法实验结果对比

Comparison of experimental results of different methods

Method	实现	评价指标		
		IoU/ $\times 10^2$	Pd/ $\times 10^2$	Fa/ $\times 10^6$
ACM	全监督	62.31	94.18	65.53
	弱监督	39.72	87.23	35.24
ALCNet	全监督	60.76	92.29	55.66
	弱监督	32.13	88.52	81.91
DNA-Net	全监督	71.45	95.18	22.36

弱监督	44.11	84.24	19.60
全监督（上采样）	72.83	96.06	12.97
弱监督（上采样）	46.18	89.56	17.84

5.2 与传统红外目标检测方法的比较

我们将动态上采样技术应用于 DNA-net 中与使用标签演化框架的几种基于 CNN 的最先进方法进行对比实验，包括 ACM [19]、ALCNet [13]和 DNA-Net[27]比较。具体对比结果如表 3 所示。为了公平比较，我们在相同的设置下重新训练了所有模型。通过对比，可以看出在使用了标签演化框架之后的粗点监督，基本可以达到全监督过程的 71-75%的 IoU 结果，同时检测概率达到全监督过程的 80%以上，并且误报率的增加也在可以接受的范围内，说明了标签演化框架对于弱监督下模型性能的提升具有普遍性。而使用动态上采样方法后的 DNA-Net 把在全监督还是弱监督的情况下都有不同程度的提升，特别是对于弱监督的模型，提升效果更多，并且评测指标优于未使用动态上采样方法的其他模型。

5.3 标签演化的普适性

我们将不同模型使用标签演化框架训练，不同模型在弱监督的 IoU 值随训练 epoch 的变化趋势如下图 8，根据分析模型的 IoU 值以及包括损失值、PD 值的变化趋势可以看出，评价指标在 epoch 为 30-50 时趋于最大值，并随着继续训练，检测模型逐渐趋于过拟合状态，导致交并比开始缓慢减低。这个现象也很好的证明了，通过利用映射退化现象的标签演化框架对使弱监督下的卷积神经网络优化的能力具有普适性，可以使其在弱监督下获得接近于全监督下训练的结果。

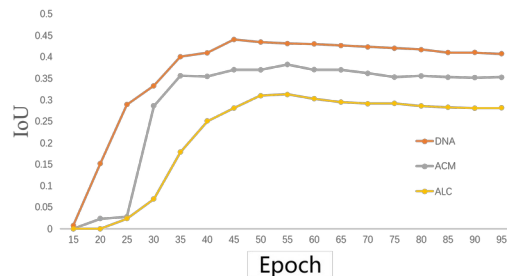


图 8 不同模型基于弱监督的 IoU 值的变化

Fig. 8 The variation of IoU values for different models based on weak supervision.

此外通过对三种模型应用标签演化框架后模型训练达到拟合的速度分析，可以看出经过处理后的模型的训练速度提升幅度是不同的，对不同模型的训练速度的影响与模型本身的网络架构也有关系，其中对于 DNA 网络的提升最明显，效果也最好。

5.4 上采样优化效果

我们将使用动态上采样方法后的目标检测网络与原有网络结构的 DNA 网络在标签演化框架下的 IoU 值进行对比，两者随 epoch 增加的变化趋势如图 9，通过对比图像变化趋势可以看出在使用了动态上采样方法后的网络相较于原模型可以提前 5 个 epoch 左右拟合，调高了模型训练效率，并且训练的模型在交并比和检测率的效果上都有 5-10%的提升，误报率有些许提高，但是因为误报率是经过 $\times 10^6$ 的放大，实际的增加是可以接受的。这个实验结果说明应用了动态上采样的网络模型在时间上和效果上优于原网络模型。

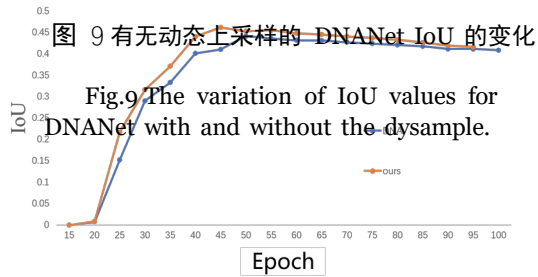


图 9 有无动态上采样的 DNANet IoU 的变化
Fig.9 The variation of IoU values for DNANet with and without the dysample.

6. 结论

在本文中，我们完成了单点监督实现弱监督红外小目标检测结合动态上采样器的工作。在我们的方法中，我们利用了一个名为单点监督标签演化的标签演化框架，以自动实现点到掩模的回归。此外，还加入了一个快速、有效、通用的动态上采样器。将改动应用于 DNANet 网络后，在 IoU 和 PD 的评测指标上分别实现超过 70% 和 90% 的全监督下原网络模型的性能。

参考文献

- [1] ZHANG M J, ZHANG R, YANG Y X, *et al.* ISNet: shape matters for infrared small target detection[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 867-876
- [2] ZHANG M J, YUE K, ZHANG J, *et al.* Exploring feature compensation and cross-level correlation for infrared small target detection[C]//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa: ACM, 2022: 1857-1865
- [3] YING X Y, WANG Y Q, WANG L G, *et al.* Local motion and contrast priors driven deep network for infrared small target superresolution[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2022, **15**: 5480-5495
- [4] ZHANG K, NI S Y, YAN D S, *et al.* Review of dim small target detection algorithms in single-frame infrared images[C]//Proceedings of the 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference. Chongqing: IEEE, 2021: 2115-2120
- [5] SU Zhenhua, LIU Xinying, DING Yibing, *et al.* A detection method for space target based on infrared images[J]. *Spacecraft Environment Engineering*, 2024, **41**(1): 27-33. DOI: 10.12126/sec.2023103(苏振华, 刘新颖, 丁奕冰, 等. 一种基于红外图像的空间目标探测方法[J]. 航天器环境工程, 2024, **41**(1): 27-33. DOI: 10.12126/sec.2023103)
- [6] PU Jingsong, XU Dong, LIU Yijun, *et al.* Rapid detection of infrared dim targets under starry sky background[J]. *Laser & Infrared*, 2017, **47**(4): 513-518. DOI: 10.3969/j.issn.1001-5078.2017.04.024 (蒲静松, 许东, 刘乙君, 等. 星空背景下红外弱小目标的快速检测[J]. 激光与红外, 2017, **47**(4): 513-518. DOI: 10.3969/j.issn.1001-5078.2017.04.024)
- [7] JIAN Yuan, HUANG Zili, WANG Xun. Infrared small target detection based on randomized tensor algorithm[J]. *Laser Technology*, 2024, **48**(1): 127-134 (蹇渊, 黄自力, 王询. 基于随机化张量算法的红外弱小目标检测[J]. 激光技术, 2024, **48**(1): 127-134)
- [8] LI Shouchang, WU Yingyue. Dual-neighborhood local weighted contrast algorithm for infrared small target detection within complex backgrounds[J]. *Semiconductor Optoelectronics*, 2024, **45**(5): 853-860 (李守昌, 吴滢跃. 复杂背景下双邻域局部权重对比度红外弱小目标检测算法[J]. 半导体光电, 2024, **45**(5): 853-860)
- [9] NIE Qingfeng, LIU Yingjie, LIANG Yun. Infrared dim target detection based on neural network model with sparsity constraint[J]. *Electronics Optics & Control*, 2022, **29**(8): 40-44 (聂青凤, 刘应杰, 梁赞. 基于稀疏约束神经网络的红外弱小目标检测技术[J]. 光电与控制, 2022, **29**(8): 40-44)
- [10] RIVEST J F, FORTIN R. Detection of dim targets in digital infrared imagery by morphological image processing[J]. *Optical Engineering*, 1996, **35**(7): 1886-1893
- [11] HAN J H, MORADI S, FARAMARZI I, *et al.* A local contrast method for infrared small-target detection utilizing a tri-layer window[J]. *IEEE Geoscience and Remote Sensing Letters*, 2020, **17**(10): 1822-1826
- [12] SUN Y, YANG J G, AN W. Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, **59**(5): 3737-3752
- [13] DAI Y M, WU Y Q, ZHOU F, *et al.* Attentional local contrast networks for infrared small target detection[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, **59**(11): 9813-9824
- [14] LI B Y, XIAO C, WANG L G, *et al.* Dense nested attention network for infrared small target detection[J]. *IEEE Transactions on Image Processing*, 2023, **32**: 1745-1758

- [15] ZHANG Jianjun, CHEN Yudan, LIU Yuling, *et al.* Infrared dim target detection method based on improved YOLOv5s model[J]. *Journal of Applied Optics*, 2024, 45(5): 975-981 (张建君, 陈玉丹, 刘玉玲, 等. 基于改进 YOLOv5s 模型的红外弱小目标检测方法[J]. 应用光学, 2024, 45(5): 975-981)
- [16] LI Y W, ZHAO H S, QI X J, *et al.* Fully convolutional networks for panoptic segmentation[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 214-223
- [17] YING X Y, LIU L, WANG Y Q, *et al.* Mapping degeneration meets label evolution: learning infrared small target detection with single point supervision[C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023: 15528-15538
- [18] LIU W Z, LU H, FU H T, *et al.* Learning to upsample by learning to sample[C]//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision. Paris: IEEE, 2023: 6004-6014
- [19] DAI Y M, WU Y Q, ZHOU F, *et al.* Asymmetric contextual modulation for infrared small target detection[C]//Proceedings of 2021 IEEE Winter Conference on Applications of Computer Vision. Waikoloa: IEEE, 2021: 949-958
- [20] WU X, HONG D F, CHANUSSOT J. UIU-Net: U-Net in U-Net for infrared small object detection[J]. *IEEE Transactions on Image Processing*, 2023, 32: 364-376
- [21] CHEN L Y, YANG T, ZHANG X Y, *et al.* Points as queries: weakly semi-supervised object detection by points[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 8819-8828
- [22] CHENG B W, PARKHI O, KIRILLOV A. Pointly-supervised instance segmentation[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 2607-2616
- [23] ZHANG S L, YU Z R, LIU L Y, *et al.* Group R-CNN for weakly semi-supervised object detection with points[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans: IEEE, 2022: 9407-9416
- [24] BEARMAN A, RUSSAKOVSKY O, FERRARI V, *et al.* What's the point: Semantic segmentation with point supervision[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016: 549-565
- [25] QIAN R, WEI Y C, SHI H H, *et al.* Weakly supervised scene parsing with point-based distance metric learning[C]//Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu: AAAI, 2019: 8843-8850
- [26] ZHANG S Y, LIEW J H, WEI Y C, *et al.* Interactive object segmentation with inside-outside guidance[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 12231-12241
- [27] REN D D, LI J B, HAN M, *et al.* DNANet: dense nested attention network for single image dehazing[C]//Proceedings of 2021 IEEE International Conference on Acoustics, Speech and Signal Processing. Toronto: IEEE, 2021: 2035-2039

赵祎鸣 男, 2001 年出生于山东省济南市, 计算机学士, 2024 年于中国海洋大学获得学士学位, 现就读于中国科学院大学国家空间科学中心攻读研究生, 主要从事微弱信号处理等方面研究

E-mail: zhaoyiming24@mails.ucas.ac.cn



郑伟 (通信作者) 男, 工学博士, 副研究员, 硕士生导师, 1996 年于西北工业大学获得学士学位, 2001 年于中国科学院西安光学精密机械研究所获得硕士学位, 2009 年于中国科学院国家空间中心获得博士学位, 主要从事微弱信号处理、非线性信号处理与天基信息处理方面的研究。

E-mail: zhengwei@nssc.ac.cn

