

DOI: 10.13382/j.jemi.B2003471

基于小样本学习的 X 光图像违禁物品识别*

邢琛聪¹ 吕淑静² 吕岳²

(1. 华东师范大学 计算机科学与技术系 上海 200062; 2. 华东师范大学 上海市
多维度信息处理重点实验室 上海 200241)

摘要:自动 X 光安检是维护公共安全的重要手段。现有 X 光违禁物品识别的研究仅针对数据集包含的类别,不能直接应用于未参与训练的类别。数据集各类别的数量不平衡也会影响模型的性能。针对以上问题,提出一种基于小样本学习的 X 光图像违禁物品分割方法。模型首先将测试图像与标注的支持图像映射至相同的特征空间,然后度量图像间的像素级相似度与区域级相似度,最后根据特征相似度分割测试图像内违禁物品区域。针对不定数量的支持集图片,采取基于卷积化门控循环单元的操作,将测试图像与不同支持图像的相似度信息融合。实验结果表明,模型在单张标注图像支持集(1-shot)和 5 张标注图像支持集(5-shot)情况下的准确率相比现有最优方法分别提高 2.0% 和 2.2%,进一步证明模型具有扩展至新类别的能力。

关键词: 违禁品检测; X 光图像; 图像分割; 小样本学习; 度量学习

中图分类号: TP183; TN919.81 文献标识码: A 国家标准学科分类代码: 520.2040

Few-shot method for prohibited item inspection in X-ray images

Xing Chencong¹ Lv Shujing² Lv Yue²

(1. School of Computer Science and Technology, East China Normal University, Shanghai 200062, China;

2. Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China)

Abstract: Automatic X-ray security inspection is an important method to maintain public safety. Current research of prohibited item inspection on X-ray images only works on pre-defined classes in the dataset and cannot be generalized to unseen categories. The imbalance problem in the dataset will also affect the performance of models. In order to solve above defects, the paper proposes a segmentation model for prohibited item inspection in X-ray Images based on few-shot learning. The model first embeds the test image and annotated support images to a common space, then measures the spatial pixel-wise similarity and regional similarity, finally segments out suspected areas in the test image. To deal with uncertain numbers of support images, a fusion method based on the ConvGRU is proposed to integrate the similarity information for the test image and different support images. Experiments show that the proposed model improves 2.0% and 2.2% meanIoU compared to the state-of-the-art methods under 1-shot task and 5-shot task, which demonstrates the ability to recognize unseen categories.

Keywords: prohibited item inspection; X-ray images; image segmentation; few-shot learning; metric learning

0 引言

X 光图像违禁物品筛查是维护公共安全的重要手段,现已广泛应用于交通出行、物流运输等行业。X 光图像违禁物品筛查主要依赖人工检查,操作人员必须在短时间内

观察并判断图像中是否包含各类违禁物品。随着筛查效率的要求提高,自动化 X 光图像分析技术是必不可少的。

目前,X 光图像违禁物品识别的研究主要基于图像分类、目标检测、图像分割等方法。图像分类方法能给出违禁品的类别信息。传统图像分类算法通过手工构造的特征,如加速鲁棒特征(SURF)、尺度不变特征变换

(SIFT)、方向梯度直方图(HOG)等,以及视觉词袋模型和支持向量机(SVM)^[1]来判断图像中违禁物品的类别。随着深度学习的发展,卷积神经网络(CNN)也被应用到本领域。Kundegorski等^[2]使用迁移学习方法显著提高了分类的准确率。相比图像分类,目标检测方法能进一步给出违禁物品的矩形位置框。Franzel等^[3]使用滑动窗口分类器将目标检测过渡为图像分类。Akca等^[4]微调经典深度学习目标检测模型^[5]并将其应用于违禁品检测。作为图像识别的关键步骤之一,图像分割能给出像素级分类结果。Gaus等^[6]采用基于DualCNN的深度学习模型实现X光图像违禁物品识别。

现有研究虽然推动了X光图像违禁物品识别的发展,但仍存在一些共同的缺陷,阻碍模型在实际场景中的应用。1)模型在数据集上完成训练后无法识别新的类别。违禁品的定义常随不同场景变化。当需要检测新的类别时,模型必须重新训练。2)不同类别的违禁品出现概率不同,数据集常包含类别数量不平衡的问题。如含有枪支、毒品的包裹在真实场景中极少出现,无法收集大量数据;包含打火机、电池的包裹却很常见。即使采用数据增强^[7]或弱监督学习^[8]也无法在此情况下训练模型。

为了解决以上的问题,本文从小样本学习角度分析X光图像违禁品识别任务,提出一个新的小样本图像分割模型。小样本学习在完成训练后仅需有限的标注数据就可以泛化到新类别。图1所示为模型的总体结构。本文提出的模型主要包含稠密特征编码模块、双路比较模块、多分辨率解码模块。X光图像常具有重叠透视问题和物体视角变化问题,在稠密特征提取模块中,本文修改DRN^[9]的结构,加强不同层次特征的融合。在非局部操作^[10](non-local operation)的启发下,本文设计了一个双路比较模块,并将其作为相似度度量函数。该模块从细粒度的像素角度和粗粒度的区域角度比较测试图像与所有支持图像的相似度。由于支持集包含一张或多张标注图像,本文使用一种基于卷积门控循环单元^[11](ConvGRU)的特征融合方法,将测试图像与不同支持图像的相似度特征融合。最后,双路比较特征被输入至多分辨率解码模块,以生成分割结果。

本文在SIXray^[12]数据集上进行实验并验证模型性能。结果表明,本文的模型明显优于现有最优的小样本图像分割模型。以单张标注图像为支持集(1-shot),本文的模型取得48.2%的平均交并比(MeanIoU),相比最优模型提升2.0%。以5张标注图像为支持集(5-shot),模型取得高于最优模型2.2%的50.0%平均交并比。

1 模型设计

小样本图像分割可视为单样本图像分割的扩展。

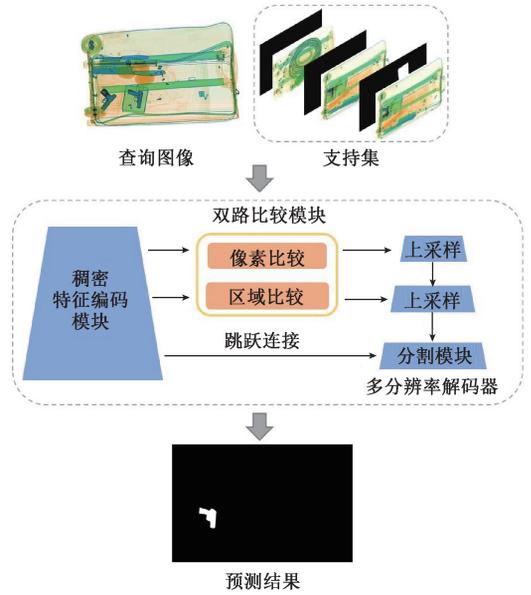


图1 模型总体结构

Fig. 1 Overall structure of the model

1.1 小样本学习任务

小样本学习(few-shot learning)任务旨在获取可迁移的知识,使学习模型仅需少量标注样本就可以预测未参与训练的新类别。在小样本任务中,测试图像一般被称为查询图像(query image),支持集(support set)由标注图像组成。支持集仅包含1张标注图像时,小样本学习等同于单样本学习(one-shot learning)。

小样本分割模型以查询图像和支持集图像为输入,通过先验知识和支持集的标注信息分割查询图像。相似度信息、元知识等都是常用的先验知识。

1.2 单样本分割模型

针对X光违禁品识别问题,本文提出了一个新的小样本分割模型。该模型首先使用稠密特征编码模块将查询图像和支持图像嵌入到相同的特征空间。由像素比较子模块和区域比较子模块组成的双路比较模块被同时应用在block 3和block 5的特征。通过扩展非局部操作^[10],像素比较子模块能快速度量查询特征和支持特征逐像素对的相似度。区域比较子模块利用平均掩膜池化^[13](mask average pooling, MAP)去除不相关信息,并从区域角度比较查询特征和支持特征的相似度。最后,多分辨率解码器逐级融合不同尺度的相似度特征与查询图像特征,进一步生成分割结果。

稠密特征编码模块旨在提取X光图像的特征。X光图像包含重叠、视角变化等问题,传统特征提取网络并不能有效提取X光图像特征。本文使用修改的DRN网络^[9]提取不同层次的特征。稠密特征编码模块的结构如图2所示。

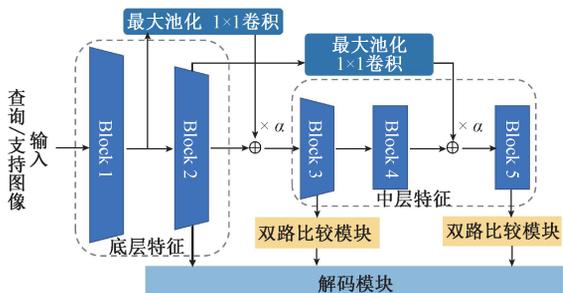


图 2 稠密特征编码模块

Fig. 2 Dense feature encoder

稠密特征编码模块通过融合底层特征与中层特征有效提升 X 光图像的特征质量。卷积神经网络的多层结构能逐步提取不同层次的特征。底层特征主要包含的边缘和颜色信息。在 X 光图像中,不同颜色表示不同的物质。底层特征包含的颜色、边缘信息有助于从透视重叠的图像背景中分割出不同的目标。中层特征包含的属性信息是解决小样本分割问题的关键。不同类别的违禁物品虽然高层特征不同,但会共享相同的属性信息,这有助于模型识别在训练集中未出现的新类别。本文稠密特征编码模块将底层特征通过最大池化、卷积等操作与中层特征融合。具体过程定义为:

$$F_{output} = F_{Middle} + \alpha \cdot MaxPool(F_{Lower}) \quad (1)$$

式中: α 是融合参数,控制底层特征的比例,本文将 α 置为 0.5; F_{output} 是融合后的特征输出。查询图像和支持图像使用相同的稠密特征编码模块提取特征。

双路比较模块是模型的核心部分,包含两个独立的子模块,为像素比较子模块和区域比较子模块。本文提出的模型包含两个双路比较模块,分别置于稠密特征编码模块 block3 和 block5 后。每个双路比较模块以相应层级的查询特征和支持特征为输入,度量特征相似度。

像素比较子模块旨在捕捉查询图像特征和支持图像特征在逐像素的相似度。像素比较子模块的结构如图 3 所示。本文利用非局部操作计算查询图像每个像素与支持图像所有位置的相似度。逐像素比较操作可定义为:

$$S_i = \frac{1}{C(x^q)} \sum_{v_j} f(x_i^q, x_j^s) \quad (2)$$

其中:

$$f(x_i^q, x_j^s) = e^{\theta(x_i^q)^T [\phi(x_j^s) \cdot mask]} \quad (3)$$

式中: i 指示查询特征上的一个位置; j 枚举支持特征的所有位置; S 是相似度矩阵; S_i 是查询特征 x_i^q , 支持特征 x_j^s 计算出的相似度值; f 是度量特征像素对 x_i^q, x_j^s 的相似度函数。相似度矩阵 S 使用 $C(x)$ 函数归一化。

在实现过程中, f 使用 2 个 1×1 的卷积模块作为 $\theta(\cdot), \phi(\cdot)$, 将特征通道数减半以降低计算复杂度。支持特征需要与其二值掩膜相乘,去除不相关区域的信息。

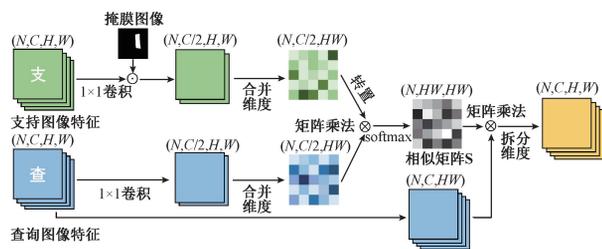


图 3 像素比较子模块结构

Fig. 3 Pixel comparison submodule

随后,将查询特征和支持特征的后两维合并,特征大小从 $(N, C/2, H, W)$ 转化为 $(N, C/2, HW)$, 并计算两者的矩阵乘积作为相似度矩阵 S 。最后, f 使用 $softmax$ 函数归一化相似度矩阵 S 。像素比较子模块的输出是查询特征与相似度矩阵的乘积。

从空间角度分析,像素比较子模块相比直接拼接查询和支持特征更加精细,也更具有可解释性。相似度矩阵 S 不仅提供逐像素对的相似值,同时也包含模型将标注信息从支持图像迁移至查询图像的过程。

区域比较子模块聚焦查询图像和支持图像的局部区域相似度。由于使用相同的稠密特征编码器映射查询图像和支持图像,相似目标的特征表示必定相似。区域比较子模块通过拼接和卷积操作隐式得计算特征相似度。区域比较子模块的结构如图 4 所示。不同于像素比较子模块,该操作可以避免细粒度的噪声干扰,生成区域连续的特征响应。

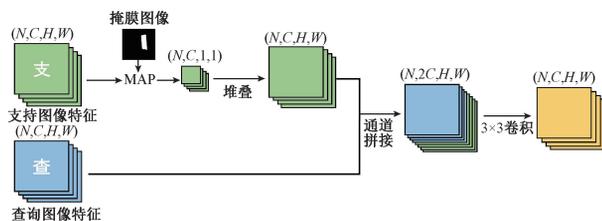


图 4 区域比较子模块结构

Fig. 4 Regional comparison submodule

在实现过程中,首先使用 $MAP^{[13]}$ 去除不相关信息,并将特征图压缩成特征向量。平均掩膜池化的定义如下:

$$MAP = \frac{\sum_{v_i} x_i \cdot \{Mask_i = 1\}}{\sum_{v_j} 1 \{Mask_j = 1\}} \quad (4)$$

式中: i 指代特征图上得一个具体位置, $1 \{ \cdot \}$ 是狄利克雷函数,当内部表达式为真时,值为 1,反之则为 0。相比分割任务中常用的全局平均池化 (global average pooling, GAP),平均掩膜池化利用标注的信息将池化操作限制在目标出现的区域而非整张特征图,此操作可以降低物体

尺度变化带来的影响。随后模块将查询特征与堆叠到相同大小的支持向量拼接,并使用 3×3 的卷积隐式得进行特征比较。

双路比较模块将像素比较子模块和区域比较子模块输出的特征按通道拼接,并用 1×1 的卷积压缩通道数,最终生成的双路比较特征送入解码器进一步处理。

多分辨率解码器逐级放大特征图,并生成最终的分割结果。由于稠密特征编码模块使用最大池化操作提升感受野,此操作会造成信息损失。为了降低信息损失对分割结果的影响,多分辨率解码器以 block3、block5 的双路比较特征及稠密特征编码模块的 block2 特征为输入,逐级放大融合并生成二值化分割结果。

多分辨率解码器主要由两个上采样模块和 1 个分割模块组成。上采样模块如图 5 所示。模块先将输入的低分辨率特征图放大 1 倍,并与相同尺度的高分辨率特征图拼接。空间膨胀金字塔池化^[14](ASPP)使用不同膨胀率的空洞卷积(atrous convolution)处理不同尺度的目标。分割模块是多分辨率解码模块的最后部分,用于压缩特征图的通道数和生成分割结果。本文选择使用残差模块^[15]压缩通道数, Sigmoid 函数激活并生成分割结果。

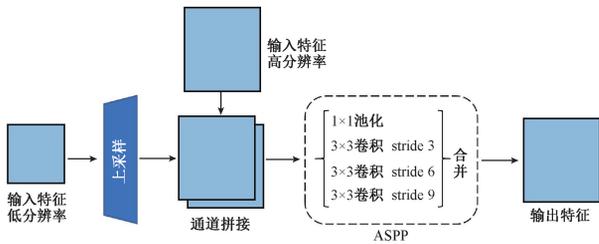


图 5 上采样模块
Fig. 5 Upsample block

1.3 扩展至小样本模型

本文设计了一个 X 光图像单样本分割模型,但在小样本问题中,支持集常包含多张相同类别的标注图片。在 OSLSM^[16]和 CANet^[17]中,模型通常使用加权平均和注意力机制融合多张图像的相似度特征。此类方法效果不稳定,不适合当支持集图像数量不确定的场景。

为了解决这个问题,本文使用 ConvGRU^[11]的特征融合方法。从单样本学习扩展为小样本学习方法。在融合过程中,查询图像与不同支持图像的双路比较特征被视为序列数据依次输入到 ConvGRU,融合后的结果被输出至多分辨率解码器。本文选择第一组查询-支持图像对的特征作为 ConvGRU 的初始隐藏状态。不同于加权平均或者注意力机制,ConvGRU 提供的记忆机制允许模型可以先记住目标的全局语义,并逐渐从不同的查询-支持图像对添加细节信息,此过程旨在模仿人类认知时,先学习整体概念,再逐步发现细节的能力。

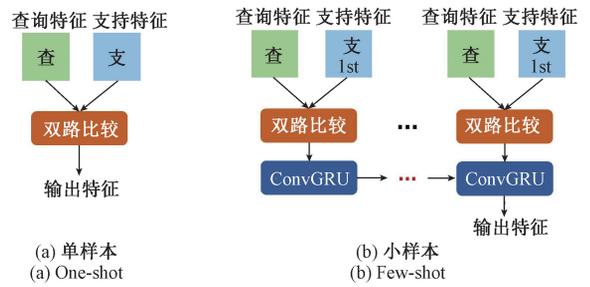


图 6 从单样本学习扩展到小样本学习的方法
Fig. 6 Extension methods for one-shot model

2 实验验证

2.1 实验设置

为了验证模型的性能,本文在 SIXray^[12]数据集上进行实验。SIXray 数据集包含 8 929 张含有违禁物品的 X 光图像,1 050 302 张正常样本图像。违禁品共 6 类,包含枪支、刀具、扳手、钳子、剪刀、锤子。在实验过程中,依次选择 1 类违禁品的图像组成测试集,其余 5 类的图像作为训练集建立模型。为了确保测试结果的代表性以及避免不同类别图像数量带来的影响,本文在测试时随机选择 500 对查询-支持图像对作为输入,衡量模型的性能。实验使用 MeanIoU 描述模型的性能。交并比度量了模型的分割结果与真实掩膜的重叠百分比。

模型使用 Pytorch 框架实现,在 Nvidia 2080Ti 上训练。模型的初始学习率为 2.5×10^{-3} ,并用随机梯度下降方式优化模型。由于不同违禁物品的尺度差异较大,本文使用 dice loss 作为损失函数来处理不平衡的目标和背景像素数量。在模型消融实验部分,本文将进一步证明 dice loss 相比其他常用的分割损失函数更适合小样本违禁物品分割任务。

2.2 实验结果

SIXray 数据集的实验结果如表 1 所示。本文使用支持集包含 5-shot 的结果代表模型在小样本情况下的分割性能。所有对比模型均采用原始论文中的参数设置,并在相同设备上训练、测试。为了训练非小样本模型 DualCNN^[6],本文将 500 对查询-支持图像对的查询图像作为测试集,支持图像作为训练集。表 1 的结果表明,本文提出的模型优于现有小样本分割模型。在 1-shot 任务下,模型取得了 48.2% 的平均交并比,高于最优模型 2.0%。在 5-shot 的任务下,模型也取得了高于最优模型 2.2% 的平均交并比。此结果体现出本文提出的模型比现有小样本分割模型更适合 X 光图像。与非小样本模型 DualCNN 相比,实验结果证明现有小样本模型在未见类别上的性能与非小样本模型直接在这些类别上训练的性能仍然存在差距,但小样本模型可以预测未见类别的能力是无法替代的。

表 1 不同模型在 SIXray 数据集的分割结果

Table 1 Segmentation results of different models on the SIXray dataset

实验模型	1-shot							5-shot						
	手枪	扳手	钳子	刀具	剪刀	锤子	MeanIoU	手枪	扳手	钳子	刀具	剪刀	锤子	MeanIoU
OSLSM ^[16]	33.2	29.4	32.5	33.4	10.3	24.2	27.2	34.7	31.4	34.2	34.1	11.9	25.0	28.6
PANet ^[13]	40.2	45.1	35.2	50.4	11.6	31.4	35.7	43.5	46.8	39.7	52.0	13.7	32.6	38.1
CANet ^[17]	43.4	61.5	59.8	52.1	22.1	38.5	46.2	45.6	63.4	61.3	54.3	22.5	39.9	47.8
本文	45.3	64.7	61.9	55.2	21.9	38.9	48.2	47.0	67.5	64.6	56.7	23.0	39.7	50.0
DualCNN ^[6]	60.6	71.4	68.5	66.3	52.0	57.6	62.7	63.2	75.9	71.7	69.4	55.8	59.1	65.9

在实际场景中,模型的预测速度也是重要的评价指标。表 2 为不同模型在 1-shot 和 5-shot 任务下的预测单张查询图像的速度。从表 2 结果可知,在 1-shot 任务下,本文提出的模型比现有最快模型慢 3 ms;在 5-shot 的任务下,本文提出的模型比现有最快模型慢 4 ms。综合表 1 的结果,本文的模型实现了准确率与预测速度的平衡,即以较快的速度实现更高的准确率。

模型分割的结果如图 7 所示。图 7(a) 为 1-shot 分割的结果,图 7(b) 为 5-shot 分割的结果。红色阴影代表违禁物品。通过观察可以发现,具有明显形状、纹理特征的

违禁物品如钳子、手枪等,更容易被识别。

表 2 不同模型的预测速度比较

Table 2 Comparison of inference speed of different models (ms)

实验模型	1-shot 预测速度	5-shot 预测速度
OSLSM ^[16]	26	37
PANet ^[13]	28	38
CANet ^[17]	31	43
本文	30	41

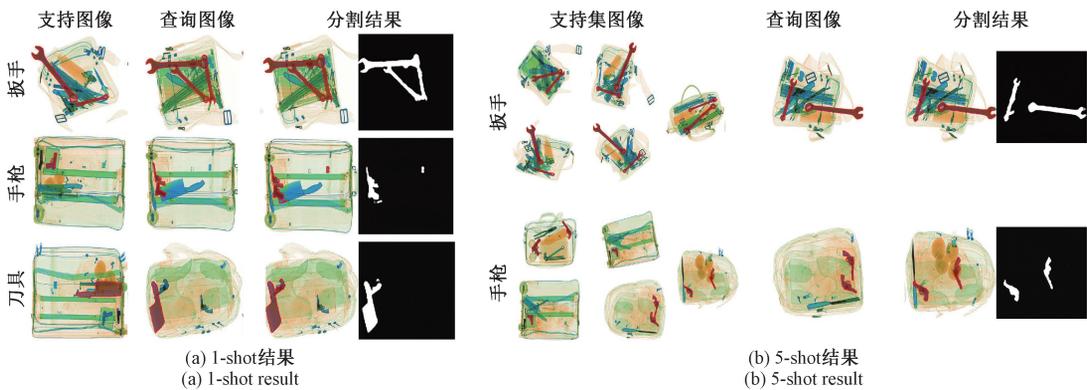


图 7 模型在 SIXray 数据集的分割结果

Fig. 7 Segmentation results on the SIXray dataset

2.3 消融实验

首先,本文测试了不同的特征提取模块。本文选择目前常用的 VGG16^[18], Resnet50^[15], DRN-C^[9], InceptionV3^[19] 特征提取网络与稠密特征提取模块进行对比实验。所有特征提取网络均使用 block3 和 block5 (由于层数限制, VGG16 网络采用 block4) 的特征放入双路比较模块。实验结果如图 8 所示。实验表明,针对 X 光图像,本文提出的稠密特征提取模块优于传统特征提取网络。通过对比 DRN 网络的结果,也可证明本文采取将底层特征与中层特征融合的操作是有效的。

针对双路比较模块,本文测试了在 1-shot 情况下不使用双路比较模块、仅使用像素比较子模块、仅使用区域比较子模块、使用完整双路比较模块时模型的结果。实验结果如表 3 所示,本文提出的像素比较子模块与区域

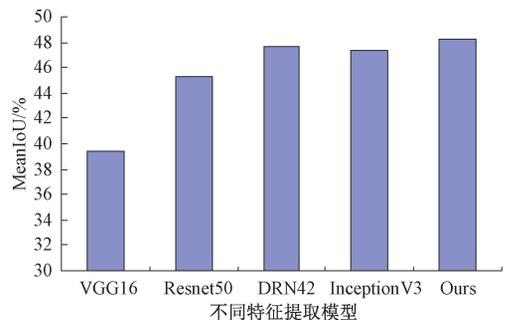


图 8 不同特征提取模块的实验结果

Fig. 8 Results of different feature encoders

比较子模块均可大幅提升模型的性能。双路比较模块在同时采用 2 个子模块时,模型的结果达到最优。

表 3 双路比较模块消融实验结果

Table 3 Ablation experiments on dual comparison module

像素比较子模块	区域比较子模块	MeanIoU/%
		31.9
✓		45.7
	✓	46.3
✓	✓	48.2

为了验证 dice Loss 损失函数带来的性能提升,本文尝试了 3 种常用的图像分割损失函数, Cross entropy Loss, Focal Loss 和 Dice Loss。实验结果如表 4 所示,在本文提出的任务上,使用 Dice Loss 约束的模型最终的平均交并比最高,证明了 Dice Loss 更适合本文提出的任务和模型。

表 4 不同损失函数的实验结果

Table 4 Ablation experiments with different loss functions

损失函数	1-shot MeanIoU	5-shot MeanIoU
Cross entropy Loss	47.6	49.5
Focal Loss ^[20]	47.9	49.6
Dice Loss	48.2	50.0

本文还测试了支持集包含不同数量的标注图像时模型的性能。实验对比基于 ConvGRU 的特征融合方法、OSLSM^[16]的特征平均方法和 CANet^[17]的基于注意力机制方法。支持集图像数量从 1~16 逐渐递增,结果如图 9 所示。实验结果表明,当支持集样本数量少于 5 张时,基于 ConvGRU 的方法优势不明显。随着图像数量不断增加,基于 ConvGRU 的方法的优势逐渐体现。

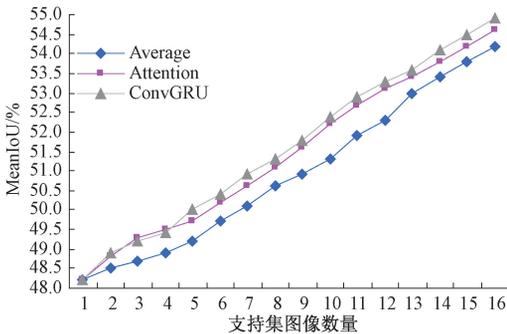


图 9 特征融合模块对比实验结果

Fig. 9 Comparison of different feature integration methods

3 结论

本文从小样本学习角度分析 X 光图像违禁物品识别问题,并提出一个全新的小样本分割网络,模型在完成训

练后仅需要一张或多张标注图像就能预测新的类别。

本文的模型由稠密特征编码模块、双路比较模块、多分辨率解码器组成。稠密特征编码模块能有效提取 X 光图像特征。双路比较模块从像素角度及区域角度同时度量输入查询-支持图像对的特征相似度。多分辨率解码器逐级放大、融合特征并生成最终的分割结果。在 SIXray 数据集上的各项实验结果充分体现本文模型的优势性。在 1-shot 的情况下,模型取得 48.2% 的平均交并比,在 5-shot 的情况下,模型取得 50.0% 的评价交并比,分别高于现有最优模型 2.0%、2.2%。

参考文献

[1] 郝勇,耿佩,温钦华,等. 滚动轴承保持架缺陷的图像处理及模式识别方法研究[J]. 仪器仪表学报, 2019, 40(9): 165-172.

HAO Y, GENG P, WEN Q H, et al. Research on image processing and pattern recognition of skew defect of antifriction bearing cage [J]. Chinese Journal of Scientific Instrument, 2019, 40(9): 165-172.

[2] KUNDEGORSKI M E, AKCAY S, DEVEREUX M, MOUTON A, et al. On using feature descriptors as visual words for object detection within x-ray baggage security screening[C]. IEEE International Conference on Crime Detection Prevention, 2016: 12-18.

[3] FRANZEL T, SCHMIDT U, ROTH S. Object detection in multi-view X-ray images [C]. Joint German Association for Pattern Recognition and OAGM Symposium, 2012: 144-154.

[4] AKCAY S, KUNDEGORSKI M E, DEVEREUX M. Transfer learning using convolutional neural networks for object classification within x-ray baggage security imagery [C]. IEEE International Conference on Image Processing, 2016: 1057-1061.

[5] 周晓彦,王珂,李凌燕. 基于深度学习的目标检测算法综述[J]. 电子测量技术, 2017, 40(11): 89-93.

ZHOU X X, WANG K, LI L L. Review of object detection based on deep learning [J] Electronic Measurement Technology, 2017, 40(11): 89-93.

[6] GAUS Y F A, BHOWMIK N, AKCAY S. Evaluation of a dual convolutional neural network architecture for object-wise anomaly detection in cluttered X-ray security imagery [C]. International Joint Conference on Neural Networks, 2019: 1-8.

[7] 吴睿曦,肖秦琨. 基于深度网络和数据增强的多物体图像识别[J]. 国外电子测量技术, 2019, 38(5): 86-90.

WU R X, XIAO Q K. Multi-object images recognition based on deep networks and data augmentation [J].

- Foreign Electronic Measurement Technology, 2019, 38(5): 86-90.
- [8] 谢晓蔚, 史健芳. 弱监督卷积神经网络的多目标图像检测研究[J]. 电子测量与仪器学报, 2019, 33(6): 31-37.
- XIE X W, SHI J F. Research of convolutional neural networks with weakly-supervised learning on multi-object image detection [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(6): 31-37.
- [9] YU F, KOLTUN V, FUNKHOUSER T. Dilated residual networks [C]. IEEE conference on Computer Vision and Pattern Recognition, 2017: 472-480.
- [10] ZHU Z, XU M, BAI S. Asymmetric non-local neural networks for semantic segmentation [C]. IEEE International Conference on Computer Vision, 2019: 593-602.
- [11] SHI X, GAO Z, LAUSEN L. Deep learning for precipitation nowcasting: A benchmark and a new model [C]. Advances in Neural Information Processing Systems, 2017: 5617-5627.
- [12] MIAO C, XIE L, WAN F. Sixray: A large-scale security inspection x-ray benchmark for prohibited item discovery in overlapping images [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2019: 2119-2128.
- [13] WANG K, LIEW J H, ZOU Y. Panet: Few-shot image semantic segmentation with prototype alignment [C]. IEEE International Conference on Computer Vision, 2019: 9197-9206.
- [14] CHEN L C, PAPANDEOU G, KOKKINOS I. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40 (4): 834-848.
- [15] HE K, ZHANG X, REN S. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [16] SHABAN A, BANSAL S, LIU Z, et al. One-shot learning for semantic segmentation [C]. The British Machine Vision Conference, 2017: 167-182.
- [17] ZHANG C, LIN G, LIU F. Canet: Class-agnostic segmentation networks with iterative refinement and attentive few-shot learning [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2019: 5217-5226.
- [18] KAREN S, ANDREW Z. Very deep convolutional

networks for largescale image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1-14.

- [19] SZEGEDY C, VANHOUCKE V, IOFFE S. Rethinking the inception architecture for computer vision [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818-2826.
- [20] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42 (2): 318-327.

作者简介



邢琛聪, 2018 年于上海理工大学获得学士学位, 现为华东师范大学硕士研究生, 主要研究方向为图像处理与小样本学习。

E-mail: 51184506047@stu.ecnu.edu.cn

Xing Chencong received his B. Sc. degree from University of Shanghai for Science and Technology in 2018. Now he is a M. Sc. candidate at East China Normal University. His main research interests include image processing and few-shot learning.



吕淑静, 2000 年于山东师范大学获得学士学位, 2005 年于电信科学技术研究院获得硕士学位, 2014 年于华东师范大学获得博士学位, 现为华东师范大学专任研究员, 主要研究方向为图像处理、模式识别、机器学习。

E-mail: sjlv@cs.ecnu.edu.cn

Lv Shujing received her B. Sc. degree from Shandong Normal University in 2000, M. Sc. degree from China Academy of Telecommunications Technology in 2005, Ph. D. degree from East China Normal University in 2014. Now she is a research professor at East China Normal University. Her main research interests include image processing, pattern recognition and machine learning.



吕岳, 1990 年于浙江大学获得学士学位, 1993 年于浙江大学获得硕士学位, 2000 年于上海交通大学获得博士学位, 现为华东师范大学教授, 主要研究方向为图像处理、模式识别与智能系统。

E-mail: ylu@cee.ecnu.edu.cn

Lv Yue received his B. Sc. degree from Zhejiang University in 1990, M. Sc. degree from Zhejiang University in 1993, Ph. D. degree from Shanghai Jiaotong University in 2000. Now he is professor at East China Normal University. His main research interests include image processing, pattern recognition and intelligent system.