

基于多层注意力机制的安检危险品识别算法

王文¹, 周亚同^{1*}, 史宝军², 何昊¹, 张建伟¹

¹河北工业大学电子信息工程学院, 天津 300401;

²河北工业大学机械工程学院, 天津 300401

摘要 针对现有识别算法使用的数据集过于简单、对真实场景中安检图像危险品的识别准确率低且容易导致误检和漏检的问题, 基于渗透假设的类平衡分层求精提出了一种结合多层通道注意力机制和空间注意力机制的算法。首先, 在安检图像分层建模的基础上在特征图中加入通道注意力机制, 对不同通道特征赋予不同的权重。然后, 加入空间注意力机制, 在空间上对安检图像特有的颜色特征赋予不同的权重。最后, 利用残差网络在安检图像不同层分别加入双注意力机制进行消融实验。实验结果表明, 对固定两层同时添加双注意力机制后, 网络对安检危险品的识别精度有显著提高, 验证了多层注意力机制算法的有效性和鲁棒性。

关键词 图像处理; 安检图像; 类平衡分层求精; 通道注意力机制; 空间注意力机制; 多层

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202259.0210011

Recognition Algorithm of Dangerous Goods in Security Inspection Based on Multi-Layer Attention Mechanism

Wang Wen¹, Zhou Yatong^{1*}, Shi Baojun², He Hao¹, Zhang Jianwei¹

¹School of Electronic Information Engineering, Hebei University of Technology, Tianjin 300401, China;

²School of Mechanical Engineering, Hebei University of Technology, Tianjin 300401, China

Abstract Aiming at the problems that the data set used by the existing recognition algorithms is too simple, the recognition accuracy of dangerous goods in security inspection images in real scenes is low, and it is easy to lead to false detection and missed detection, we propose a class-balanced hierarchical refinement algorithm based on penetration hypothesis, which combines multi-layer channel attention mechanism and space attention mechanism. First, based on the hierarchical modeling of security image, channel attention mechanism is added to the feature map to give different weight to different channel features. Then, spatial attention mechanism is added to give different weight to the unique color features of security image in space. Finally, the residual network is used to add double attention mechanism to different layers of security image for ablation experiment. The experimental results show that after adding double attention mechanism to the fixed two layers at the same time, the network can significantly improve the identification accuracy of dangerous goods in security inspection, and verify the effectiveness and robustness of the multi-layer attention mechanism algorithm.

Key words image processing; security inspection image; class-balanced hierarchical refinement; channel attention mechanism; spatial attention mechanism; multi-layer

收稿日期: 2020-12-14; 修回日期: 2021-01-13; 录用日期: 2021-03-16

基金项目: 国家重点研发计划智能机器人专项子课题(2019YFB1312102)、河北省自然科学基金(F2019202364)

通信作者: *zhouyatong.zw@126.com

1 引言

行李中隐藏的危险品极大威胁着旅客的人身安全, X光成像技术在行李安检中的应用已经比较成熟, 但物体在行李箱中随意摆放、杂乱重叠会导致危险品在图像中的比例和观察视点存在差异, 使现有安检危险品识别算法在真实场景中的识别率较低。车站和机场等地大多先用安检机拍摄出 X光图像后再进行人工检查, 这种方式会耗费大量人力物力, 特别是在客流高峰期, 长时间高强度的工作会使安检员极度疲劳, 导致危险品的误检和漏检, 引发公共安全事故。

传统的安检图像危险品识别作为图像分类的一个分支, 和传统图像分类算法一致, 主要在图像颜色分割^[1]、边缘检测^[2]、图像增强^[3-5]、图像去噪^[6]等处理后通过人工进行特征提取, 再利用决策树^[7]、支持向量机(SVM)^[8]等对安检图像危险品进行识别。Lowe^[9]提出用尺度不变特征变换(SIFT)算法提取关键点特征, 得到了不错的识别效果。Bastan等^[10]提出的词袋(BOW)算法先利用SIFT提取每张图像的特征, 然后用K-means算法对特征进行聚类, 构建BOW视觉词序列, 最后用SVM进行识别。Mery等^[11]提出的自适应稀疏表示算法先从图像中提取若干随机块构建自适应性字典, 然后对该字典进行迭代优化, 并根据稀疏表示算法进行分类识别。自LeCun等^[12]提出深度学习网络以来, 各种不同结构层次的神经网络被广泛应用到人脸识别^[13]、交通标志检测^[14]、人体姿态识别^[15-16]等领域, 但神经网络在安检领域的应用较少。Akçay等^[17]通过迁移学习将深度卷积神经网络应用到X光安检图像的识别中, 该算法的特征

提取、特征表示和识别均用卷积神经网络完成, 但使用的图像相对简单, 与真实场景不符。Krizhevsky等^[18]提出了一种基于深度卷积特征的图像分类算法, 通过计算机自主学习得到比较显著的特征, 识别准确率较高。Mery等^[19]对比了传统视觉词袋、稀疏变换、经典模式识别和深度学习算法的识别精度, 结果表明, 深度学习算法的识别准确率最高。Miao等^[20]提出了一种基于类平衡分层求精(CHR)的算法, 该算法结合深度卷积神经网络和渗透假设原理对安检图像进行分层建模, 在处理真实场景下的大量不平衡数据时取得了不错的效果, 但识别精度较低。

综上所述, 传统图像分类算法均利用人工进行特征提取, 提取出的特征有限, 仅适应于简单场景下的安检图像, 不能有效识别真实场景图像中的危险品。而深度学习算法使用的图像数据简单, 识别效率差、识别精度较低, 容易造成误检和漏检, 难以应用到实际问题中。因此, 本文利用深度卷积神经网络, 基于CHR并结合多通道和空间注意力机制^[21-24], 提出了一种基于多层注意力机制的安检危险品识别算法。

2 基本原理

基于多层注意力机制的安检危险品识别算法的结构如图1所示。首先, 基于渗透假设对图像分层建模的思想, 利用特征提取网络分层提取特征图, 对图像进行分层处理; 然后, 分别对分层的特征图像加入通道注意力机制和空间注意力机制, 再利用分层求精方法逐步对特征图求精, 并对得到的特征图进行分类; 最后, 用赋予不同权重的交叉熵损失函数^[25]衡量算法预测结果的好坏。

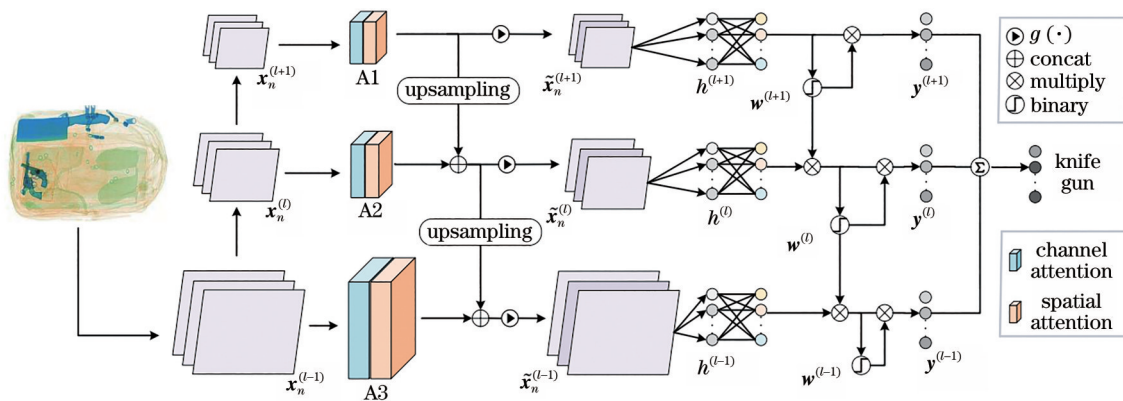


图1 基于多层注意力机制的安检危险品算法的结构

Fig. 1 Structure of the dangerous goods security inspection algorithm based on multi-layer attention mechanism

2.1 基于渗透假设的类平衡分层求精原理

在普通的分类、识别、检测任务中,经常会出现物体被遮挡的问题,导致算法仅能识别到图像中的部分目标物,可通过增大样本数量对算法进行充分训练解决该问题。不同于物体的遮挡情况,由于X射线是可穿透的,安检图像中可以看到目标物及其前后的对象,这种现象也被称为渗透假设。因此,用混合模型描述每张图像,然后对其进行分层求精,最后再进行分类。假设数据集 D 中包含 C 类 N 张图像,其中,危险品有 C' 类。对于每张输入图像 \mathbf{x}_n ,目标是得到一个 C 维向量 \mathbf{y}_n ,每个维度 $y_{n,c}$ 的模值为 0 或 1 表示图像中存在该类对象,0 则表示不存在;标签 \mathbf{y}_n^* 只存在前 C' 维中。假设 \mathbf{x}_n 由 C 个维度的子图像 $\mathbf{x}_{n,c}$ 组成,每个子图像包含一个确切的类别,从条件分布 $P_c = P(\mathbf{x}|c)$ 中采样,则每张输入图像可表示为

$$\mathbf{x}_n = \sum_c \mathbf{y}_{n,c}^* \cdot \mathbf{x}_{n,c}, \mathbf{x}_{n,c} \sim P_c. \quad (1)$$

通过判别函数 $\mathbf{y}_n = f(\mathbf{x}_n, \theta)$ 预测图像的标签,其中, θ 为函数 f 的参数。物品能以不同的尺度在不同层出现,因此,通常对多级视觉信息进行处理,以更好地识别危险品。从第 l 层提取特征图像 $\mathbf{x}_n^{(l)}$ 后,对每一层训练一个分类器 $h^{(l)}$,则每层的输出 $\mathbf{y}_n^{(l)}$ 可表示为

$$\mathbf{y}_n^{(l)} = h^{(l)}[\mathbf{x}_n^{(l)}; \xi^{(l)}], \quad (2)$$

式中, $\xi^{(l)}$ 为 $h^{(l)}$ 的参数。用标签 \mathbf{y}^* 作为监督,对每个 $\mathbf{y}_n^{(l)}$ 进行求和平均,得到输出的类别置信度,可表示为

$$\mathbf{y}_n = \sum_{l=1}^L \mathbf{y}_n^{(l)}, \quad (3)$$

式中, L 为总分层数。为了忽略图像中的非危险物

品,用函数 $g^{(l)}$ 对每一层进行精确求精,求精后的特征图像 $\tilde{\mathbf{x}}_n^{(l)}$ 可表示为

$$\tilde{\mathbf{x}}_n^{(l)} = g^{(l)}[\mathbf{x}_n^{(l)}; \mathbf{y}_n; \tau^{(l)}], \quad (4)$$

式中, $\tau^{(l)}$ 为 $g^{(l)}$ 的参数。判别函数可表示为

$$\mathbf{y}_n = \frac{1}{L} \cdot \sum_{l=1}^L h^{(l)}[\tilde{\mathbf{x}}_n^{(l)}; \xi^{(l)}]. \quad (5)$$

考虑到危险品识别的本质是多标签分类问题,常用交叉熵损失作为损失函数,可表示为

$$E[\mathbf{y}_n^*, \mathbf{y}_n^{(l)}] = -\left\{ \mathbf{y}_n^* \odot \log \mathbf{y}_n^{(l)} + (\mathbf{1} - \mathbf{y}_n^*) \odot \log [\mathbf{1} - \mathbf{y}_n^{(l)}] \right\}. \quad (6)$$

针对分层原理和类别不平衡的问题,设计了先对图像进行建模、提取分层特征并求精,最后计算损失函数的 CHR 方法。该方法的损失函数加入了交叉权重 $\mathbf{w}^{(l)}$,可表示为

$$L_n^{(l)} = \mathbf{w}^{(l)\top} \cdot E[\mathbf{y}_n^*, \mathbf{y}_n^{(l)}]. \quad (7)$$

2.2 通道注意力机制

深度学习中注意力(权重)机制的本质与人类的选择性机制类似,核心目标是从众多信息中选择出比当前任务目标更关键的信息。由于安检仪扫描出来的安检图像是 X 光图像,特征提取时会在不同通道提取出不同特征,因此,加入通道注意力机制为重要特征赋予更大的权重。本算法中的通道注意力机制原理如图 2 所示。首先,将输入的特征图 (Feature) 分别经过基于宽、高的全局最大池化 (Maxpool) 和全局平均池化 (Avgpool) 操作;然后,将特征图分别经过两个全连接层 (FC) 并与全连接层输出的权重向量进行加和操作;最后,将特征图经 Sigmoid 函数激活,生成最终的通道注意力 (Channel attention) 权值。将权值与输入的特征图 F 相乘,得到添加通道注意力机制的特征图 F' 。

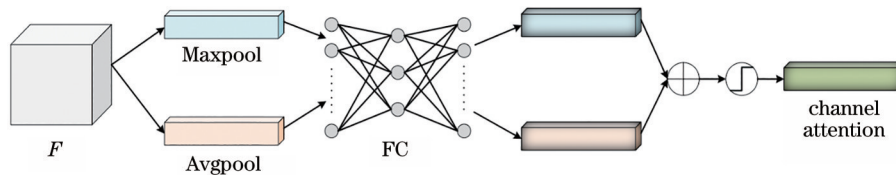


图 2 通道注意力机制的原理

Fig. 2 Principle of the channel attention mechanism

2.3 空间注意力机制

空间注意力机制指对图像中的空间域信息进行空间变换,进而提取出关键信息。考虑到 X 光图像颜色的特殊性(不同材料具有不同的颜色,且大部分危险品的颜色相同),在通道注意力机制后添

加了空间注意力机制,增大危险品颜色特征的权重,从而提高算法对危险品的识别精度,本算法中的空间注意力机制原理如图 3 所示。对于添加通道注意力机制后的特征图 F' ,首先,基于通道进行全局最大池化和全局平均池化处理,并对两个结果在

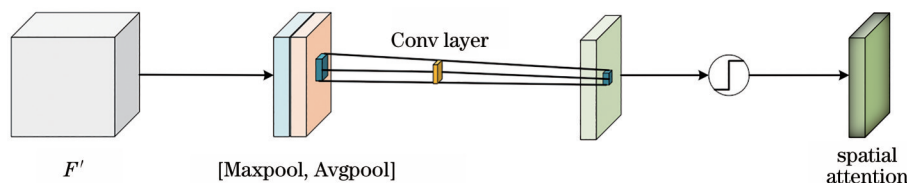


图 3 空间注意力机制的原理

Fig. 3 Principle of the spatial attention mechanism

通道上进行拼接;然后,经过一个卷积层(Conv layer)进行卷积操作,将特征图降维到一个通道上;最后,通过 Sigmoid 激活函数生成空间注意力(Spatial attention)权值。将其与添加通道注意力机制后的特征图 F' 相乘,得到最终添加通道注意力机制和空间注意力机制的特征图。

2.4 多层注意力机制的设计

安检危险品识别的本质是多标签分类问题,常用的网络有 VGG (Visual geometry group)^[26]、Inception^[27]、残差网络(ResNet)^[28]、深度卷积网络(DenseNet)^[29]。将特征提取网络与基于渗透假设的 CHR 方法相结合,在特征提取网络进行特征提取的过程中,随着网络层次的加深,不同层的特征图呈现不同的特点。浅层特征图往往会保留更多的边缘信息,深层特征图则会保留更多的语义信息。在每一层特征图上分别加入通道注意力机制和空间注意力机制,分别注意浅层特征图、中层特征图、深层特征图的关键特征,为通道和空间颜色特征赋予更高的权重,从而提升网络的特征提取能力,提高安检危险品的识别准确率。

选取 ResNet101 作为特征提取网络,该网络由不同数量的 Bottleneck 块组成四个不同的层模块 layer 1、layer 2、layer 3、layer 4,每层 Bottleneck 块的数量分别为 3、4、23、3。Bottleneck 块先通过 1×1 的卷积核对通道进行降维操作,然后用 3×3 的卷积核进行特征提取,再用 1×1 的卷积核将其恢复到原来的维度,最后将输入与提取后的输出相加。ResNet101 的结构如图 4 所示,首先,用卷积和池化层将通道数变为适应 Bottleneck 块的输入通道数;然后,通过 33 个 Bottleneck 块进行特征提取,再经过线性整流单元(ReLU)激活函数和平均池化层处理;最后,用线性层进行分类。用该特征提取网络结合基于渗透假设的 CHR 方法,对 layer 2、layer 3 和 layer 4 的输出特征分别加入通道注意力机制和空间注意力机制 A1、A2、A3。再将输出的双注意力机制特征图进行

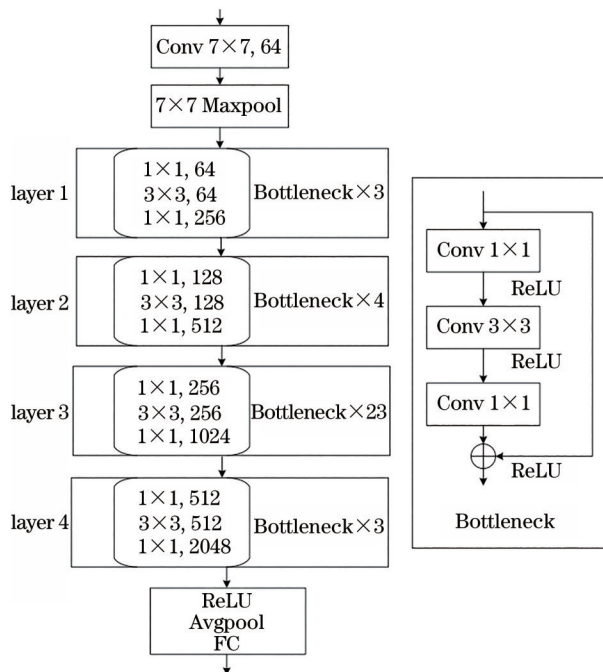


图 4 ResNet101 的结构

Fig. 4 Structure of the ResNet101

分层求精,得到 \tilde{x}_n ,最后经过全局平均池化和全连接层进行分类。

3 实验结果及分析

3.1 实验平台

实验在 Ubuntu16.04 环境下完成,计算机内存为 16 G,CPU 为 Inter(R) Xeon(R) CPU E5-2640 v4@ 2.40 GHz,GPU 为 GTX1080ti,显存为 11 GB。在 Python3.5.1 上安装了 CUDA9.0 和 CuDNN7.0.5,以支持 GPU 的使用,利用 Pytorch0.3.1 框架搭建神经网络。

3.2 数据集与评价指标

实验使用的数据集为中国科学院最新公开的安检图像数据集 SIXray^[20],该数据集为地铁安检站真实场景下的 X 射线扫描行李图像,如图 5 所示。数据集 SIXray 共包含 1059231 张图像,包括 1050302 张无危险品的行李扫描图像,8929 张含有危险品的行李扫描图像。

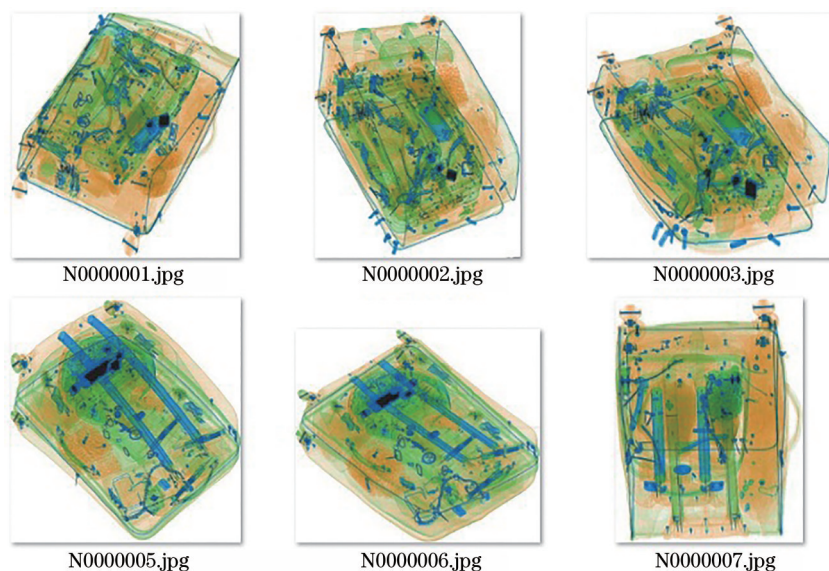


图 5 安检图像数据集中的图像

Fig. 5 Images in the security image data set

数据集 SIXray 中的危险品可分为五类,包括枪 (Gun)、刀 (Knife)、扳子 (Wrench)、钳子 (Pliers) 和剪刀 (Scissors), 对应的图像数量分别为 3131、1943、2199、3961 和 983 张, 如图 6 所示。根据不含危险品和含有危险品行李扫描图像的比例将数据集分成 SIXray10、SIXray100、SIXray1000 三个子集, 选取子集 SIXray10 作为数据集, 该子集将 8929 张含有危险品的行李扫描图像按 1:10 的比例添加无危险品的行李扫描图像, 共包含 88372 张行李扫描图像, 选取 74960 张作为训练集, 13412 张作为测试集。

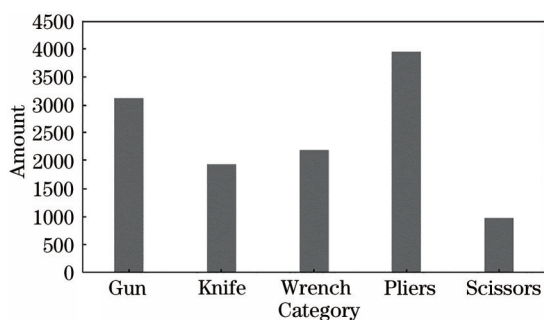


图 6 五类危险品的数量

Fig. 6 Number of five types of dangerous goods

取每类危险品的平均识别准确率 (mAP) 作为评价指标, 通过单独计算每一类别识别正确的数量占该类危险品总数量的比率, 得到每类危险品的平均识别准确率, 可表示为

$$X_{\text{mAP}} = \frac{1}{C'} \sum_{i=1}^{C'} \left(\frac{X_i^{\text{TP}}}{X_i^{\text{TP}} + X_i^{\text{FP}}} \right), \quad (8)$$

式中, C' 为危险品的类别, 实验中为 5, X_i^{TP} 为第 i 类危险品识别正确的数量, X_i^{FP} 为行李扫描图像中含有第 i 类危险品但没有被识别出来的危险品数量。

3.3 消融实验与结果分析

在神经网络的训练中, 参数设置会直接影响算法的最终检测效果。首先, 用 ResNet101 作为特征提取网络, 并加入 CHR 对安检图像进行识别。然后, 在不同层分别添加通道注意力机制和空间注意力机制, 共完成六组对比实验, 第一组不添加双注意力机制 (ResNet101+CHR), 第二、三、四组分别在第二、三、四层加入双注意力机制 (ResNet101+CHR+A1、ResNet101+CHR+A2、ResNet101+CHR+A3), 第五组在第三、四层同时加入双注意力机制 (ResNet101+CHR+A2A3), 第六组在第二、三、四层同时加入注意力机制 (ResNet101+CHR+A1A2A3)。训练前先加载预训练权重, 采用随机梯度下降 (SGD) 优化器进行优化, 动量 (Momentum) 为 0.9, 学习率采用逐级递减模式, 每 20 个 epoch 后对当前学习率乘以 0.1, 初始学习率为 0.01, batchsize 为 56, 共 40 个 epoch。训练过程中的损失值曲线如图 7 所示。可以发现, 六组实验的损失值下降均较快, 且下降趋势相同。当 epoch 小于 5 时, 损失值迅速下降; 当 epoch 介于 5 和 25 之间时, 损失值下降相对缓慢; 当 epoch 大于 25 时, 损失值趋于平稳。每组实验的损失值最终均稳定在 0.01 附近, 所

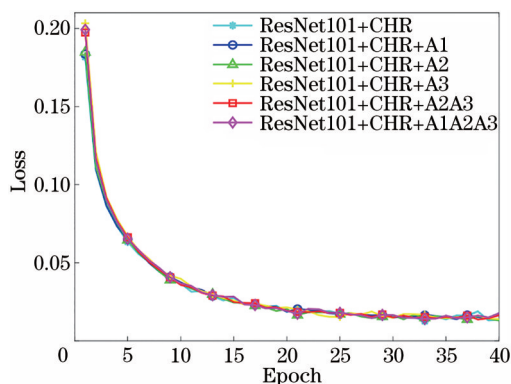


图 7 六种算法训练过程中的损失值

Fig. 7 Loss values in the training process of six algorithms

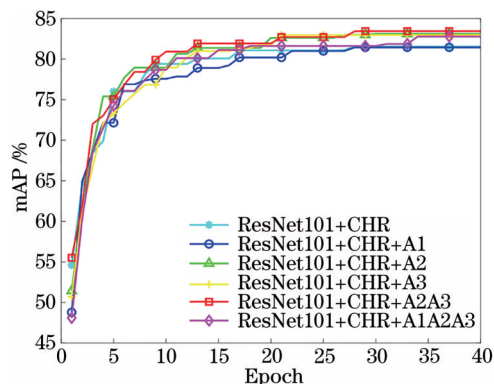


图 8 六种算法的最优准确率

Fig. 8 Optimal accuracy of six algorithms

表 1 消融实验结果

Table 1 Results of ablation experiments

Algorithm	ResNet101	CHR	A1	A2	A3	mAP/%
ResNet101+CHR	✓	✓				81.57
ResNet101+CHR+A1	✓	✓	✓			81.46
ResNet101+CHR+A2	✓	✓		✓		82.99
ResNet101+CHR+A3	✓	✓			✓	82.97
ResNet101+CHR+A2A3	✓	✓		✓	✓	83.26
ResNet101+CHR+A1A2A3	✓	✓	✓	✓	✓	82.79

有算法全部完成收敛,训练结果达到了最优效果。

不同算法的准确率曲线如图 8 所示,消融实验结果如表 1 所示。可以发现,不加入双注意力机制的 ResNet101+CHR 算法准确率为 81.57%, ResNet101+CHR+A1 算法的准确率为 81.46%,下降了 0.11 个百分点。ResNet101+CHR+A2 算法的准确率达到 82.99%, ResNet101+CHR+A3 算法的准确率达到 82.97%,相比 ResNet101+CHR 算法,分别提升了 1.42 和 1.40 个百分点。这表明在浅层对边缘特征加入双注意力机制的效果较差,在中层的融合边缘和语义特征以及在深层的语义特征中加入双注意力机制能够提高危险品的识别准确率。ResNet101+CHR+A2A3 算法的准确率达到 83.26%,相比 ResNet101+CHR 算法,提升了 1.69 个百分点。而在第二、三、四层同时加入双注意力机制后的 ResNet101+CHR+A1A2A3 算法准确率为 82.79%,效果差于 ResNet101+CHR+A2A3 算法。这表明在浅层对边缘特征加入双注意力机制反而会降低安检危险品的识别准确率,基于多层注意力机制的最优算法为在中层和深层加入两层双注意力机制的算法。

在中层和深层加入两层双注意力机制后的特征图可视化结果如图 9 所示,可以发现,加入通道注

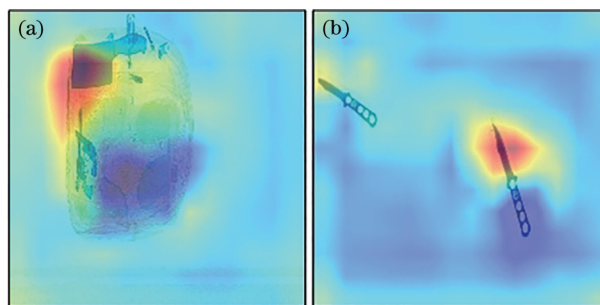


图 9 特征图的可视化结果。(a) 示例 1; (b) 示例 2

Fig. 9 Visualization result of feature map. (a) Example 1; (b) example 2

意力机制和空间注意力机制后危险品所在的位置被赋予了更高的权重。注意力的位置相对准确,验证了加入该多层注意力机制的有效性。此外,图 9 还可以作为危险品位置的定位结果。

3.4 对比实验及结果分析

为了验证最优算法的有效性,将该算法与传统算法进行了对比实验,并在其他算法中固定的两层网络加入双注意力机制,以验证多层注意力机制的鲁棒性,对比实验结果如表 2 所示。可以发现,利用 SIFT 进行特征提取加 SVM 识别的传统算法 (SIFT+SVM) 对复杂场景下的数据集识别准确率仅有 73.65%。利用 ResNet50+CHR 和具有 121 层

表 2 对比实验结果

Table 2 Results of comparative experiment unit: %

Algorithm	mAP
SIFT+SVM	73.65
ResNet50+CHR	80.68
ResNet101+CHR	81.57
DenseNet+CHR	82.06
ResNet50+CHR+A2A3 (ours)	82.34
ResNet101+CHR+A2A3 (ours)	83.26
DenseNet+CHR+A2A3 (ours)	83.89

的 DenseNet+CHR 对危险品识别的准确率分别为 80.68% 和 82.06%。ResNet50+CHR 的准确率不如 ResNet101+CHR, 而具有 121 层的 DenseNet+CHR 的准确率略高于 ResNet101+CHR。这表明网络层数越多, 就能学习到更深层次的语义特征。由于 ResNet50 和 DenseNet 的结构与 ResNet101 类似, 同样对其在固定两层加入双注意力机制。在 ResNet50+CHR 上加入多层注意力机制后得到的准确率为 82.34%, 相比原始网络提升了 1.66 个百分点。在 DenseNet+CHR 上加入多层注意力机制后的准确率达到 83.89%, 相比原始网络提升了 1.83 个百分点, 进一步验证了多层注意力机制对安检危险品识别的有效性。

4 结 论

提出了一种多层注意力机制的安检危险品识别算法, 以渗透假设的 CHR 原理为基础, 对安检图像分层模型同时加入通道和空间注意力机制, 使安检图像中的危险品被赋予更高的权重, 并用 ResNet101 在不同层分别加入双注意力机制。消融实验结果表明, 在分层安检图像中间层和深层加入双注意力机制的算法效果最佳, 相比原始算法, 对危险品的平均识别准确率提升了 1.69 个百分点。与传统算法 SIFT+SVM 和 ResNet50、DenseNet 算法的对比实验结果表明, 本算法对危险品的平均识别准确率最高, 验证了多层注意力机制的有效性和鲁棒性。但该算法仍存在少量误检和漏检图像, 且实际情况中有时需要定位危险品所在的具体位置, 因此, 还需在进一步提高危险品识别准确率的同时对所有危险品进行定位。

参 考 文 献

[1] Lin K Y, Wu J H, Xu L H. A survey on color image segmentation techniques[J]. Journal of Image and

Graphics, 2005, 10(1): 1-10.

林开颜, 吴军辉, 徐立鸿. 彩色图像分割方法综述[J]. 中国图象图形学报, 2005, 10(1): 1-10.

- [2] Duan R L, Li Q X, Li Y H. Summary of image edge detection[J]. Optical Technique, 2005, 31(3): 415-419.
段瑞玲, 李庆祥, 李玉和. 图像边缘检测方法研究综述[J]. 光学技术, 2005, 31(3): 415-419.
- [3] Jia X Y, Li T T, Jiang Z H, et al. Hue preserving low illumination image enhancement based on gene expression programming optimization[J]. Laser & Optoelectronics Progress, 2019, 56(9): 091502.
贾新宇, 李婷婷, 江朝晖, 等. 基因表达式编程优化的色调保持低照度图像增强[J]. 激光与光电子学进展, 2019, 56(9): 091502.
- [4] Yang X G, Yang L R. A method on X-ray security image enhancement[J]. Computerized Tomography Theory and Applications, 2012, 21(4): 705-712.
杨晓刚, 杨立瑞. X 射线安检图像增强方法研究[J]. CT 理论与应用研究, 2012, 21(4): 705-712.
- [5] Zheng L T, Dong Y S, Shi H L. X-ray security inspection image enhancement algorithm based on MSR and fusion theory[J]. Video Engineering, 2015, 39(7): 34-37.
郑林涛, 董永生, 史恒亮. 基于 MSR 和融合理论的 X 射线安检图像增强算法[J]. 电视技术, 2015, 39(7): 34-37.
- [6] Zhao G C, Zhang L, Wu F B. Application of improved median filtering algorithm to image de-noising[J]. Journal of Applied Optics, 2011, 32(4): 678-682.
赵高长, 张磊, 武风波. 改进的中值滤波算法在图像去噪中的应用[J]. 应用光学, 2011, 32(4): 678-682.
- [7] Luan L H, Ji G L. The study on decision tree classification techniques[J]. Computer Engineering, 2004, 30(9): 94-96, 105.
栾丽华, 吉根林. 决策树分类技术研究[J]. 计算机工程, 2004, 30(9): 94-96, 105.
- [8] Ding S F, Qi B J, Tan H Y. An overview on theory and algorithm of support vector machines[J]. Journal of University of Electronic Science and Technology of China, 2011, 40(1): 2-10.
丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报, 2011, 40(1): 2-10.
- [9] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [10] Bastan M, Yousefi M R, Breuel T M. Visual words on baggage X-Ray images[M]//Real P, Diaz-Pernil D, Molina-Abril H, et al. Computer analysis of

- images and patterns. Lecture notes in computer science. Heidelberg: Springer, 2011, 6854: 360-368.
- [11] Mery D, Riffo V, Zscherpel U, et al. GDxray: the database of X-ray images for nondestructive testing[J]. Journal of Nondestructive Evaluation, 2015, 34(4): 1-12.
- [12] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [13] Wang J X, Lei Z C. A convolutional neural network based on feature fusion for face recognition[J]. Laser & Optoelectronics Progress, 2020, 57(10): 101508.
王嘉欣, 雷志春. 一种基于特征融合的卷积神经网络人脸识别算法[J]. 激光与光电子学进展, 2020, 57(10): 101508
- [14] Tong Y, Yang H C. Real-time traffic sign detection method based on improved convolution neural network[J]. Laser & Optoelectronics Progress, 2019, 56(7): 071003.
童英, 杨会成. 基于改进卷积神经网络的实时交通标志检测方法[J]. 激光与光电子学进展, 2019, 56(7): 071003.
- [15] Zhu M K, Lu X L. Human action recognition algorithm based on Bi-LSTM-Attention model[J]. Laser & Optoelectronics Progress, 2019, 56(15): 151503.
朱铭康, 卢先领. 基于Bi-LSTM-Attention算法的人体行为识别算法[J]. 激光与光电子学进展, 2019, 56(15): 151503.
- [16] Huang Y W, Wan C L, Feng H. Multi-feature fusion human behavior recognition algorithm based on convolutional neural network and long short term memory neural network[J]. Laser & Optoelectronics Progress, 2019, 56(7): 071505.
黄友文, 万超伦, 冯恒. 基于卷积神经网络与长短期记忆神经网络的多特征融合人体行为识别算法[J]. 激光与光电子学进展, 2019, 56(7): 071505.
- [17] Akçay S, Kundegorski M E, Devereux M, et al. Transfer learning using convolutional neural networks for object classification within X-ray baggage security imagery [C]//2016 IEEE International Conference on Image Processing (ICIP), September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE Press, 2016: 1057-1061.
- [18] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [19] Mery D, Svec E, Arias M, et al. Modern computer vision techniques for X-ray testing in baggage inspection[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2017, 47(4): 682-692.
- [20] Miao C J, Xie L X, Wan F, et al. SIXray: a large-scale security inspection X-ray benchmark for prohibited item discovery in overlapping images[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 2114-2123.
- [21] Cheng D C, Meng G F, Cheng G L, et al. SeNet: structured edge network for sea-land segmentation [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(2): 247-251.
- [22] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.
- [23] Krause J, Sapp B, Howard A, et al. The unreasonable effectiveness of noisy data for fine-grained recognition [M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9907: 301-320.
- [24] Wang F, Jiang M Q, Qian C, et al. Residual attention network for image classification[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6450-6458.
- [25] Zhang Z, Sabuncu M. Generalized cross entropy loss for training deep neural networks with noisy labels [C]//proceedings of the 32nd International Conference on Neural Information Processing Systems, December 3-8, 2018, Red Hook, NY, USA. [S.l.: s.n.], 2018: 8792-8802.
- [26] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2020-12-10]. <https://arxiv.org/abs/1409.1556>.
- [27] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2818-2826.
- [28] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [29] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//2017 IEEE