



安检CT图像目标物识别方法研究进展

孙运达, 魏云超, 张 丽

Research Progress in Target Recognition Methods for Security Computed Tomography Images

SUN Yunda, WEI Yunchao, and ZHANG Li

在线阅读 View online: <https://doi.org/10.15953/j.ctta.2023.152>

您可能感兴趣的其他文章

Articles you may be interested in

基于CT安检技术的机场智能旅检通道新探索

New Exploration of Airport Intelligent Passenger Inspection Channel Based on CT Security Screening Technology

CT理论与应用研究. 2020, 29(4): 447–455

物理模型超声CT图像代数法重建和结构识别

Algebraic Reconstruction and Structure Recognition for Ultrasound CT Image on Physical Model

CT理论与应用研究. 2019, 28(2): 195–204

CT在评价肝硬化病程方面的研究进展

The Research Progress of CT in Evaluating the Course of Cirrhosis

CT理论与应用研究. 2019, 28(6): 753–761

结直肠癌能谱CT的研究进展

Research Progress of Energy Spectrum CT of Colorectal Cancer

CT理论与应用研究. 2020, 29(6): 751–758

基于特征学习的低剂量CT成像算法研究进展

Research Progress of Low Dose CT Imaging Based on Feature Learning

CT理论与应用研究. 2019, 28(3): 393–406

基于人工智能的胸部CT肺结节检出及良恶性诊断效能评估

Evaluation of Detection and Diagnostic Efficiency of Pulmonary Nodules by Chest CT Based on Artificial Intelligence

CT理论与应用研究. 2021, 30(6): 709–715



关注微信公众号, 获得更多资讯信息

孙运达, 魏云超, 张丽. 安检 CT 图像目标物识别方法研究进展[J]. CT 理论与应用研究 (中英文), 2024, 33(2): 263-271. DOI:10.15953/j.ctta.2023.152.

SUN Y D, WEI Y C, ZHANG L. Research Progress in Target Recognition Methods for Security Computed Tomography Images[J]. CT Theory and Applications, 2024, 33(2): 263-271. DOI:10.15953/j.ctta.2023.152. (in Chinese).

安检 CT 图像目标物识别方法研究进展

孙运达¹, 魏云超², 张丽^{3✉}

1. 同方威视技术股份有限公司, 北京 100084
2. 北京交通大学计算机与信息技术学院, 北京 100044
3. 清华大学工程物理系, 北京 100084

摘要: 基于先进的安检 CT 扫描设备, 研究三维 CT 图像中的目标物识别方法, 对于提升安检质量的可靠性和稳定性具有十分重要的意义。本文系统评述安检 CT 图像目标物识别的最新研究进展和发展趋势, 不仅对近年来本领域的代表性学术研究成果进行总结回顾, 还首次将工业界的相关专利也纳入分析范围; 在文献归类环节, 除沿用常见的技术路线和任务类型两种归类法, 还根据计算空间维度的差异提出一种新的归类法, 并对检索到的代表性文献进行多标签归类; 通过总结分析本领域研究成果, 发现 3 个重要特点, 分析其主要成因, 最后对本领域的研究方向提出几点建议。

关键词: 安检 CT; 目标物识别; 智能识别

DOI:10.15953/j.ctta.2023.152

中图分类号: TP 391.41; 0 212

文献标识码: A

安全检查是货物运输和旅客出行安全保障不可或缺的重要环节, 安检技术也在持续的发展进步。目前, 能够对货物和行李包裹实施非侵入式查验的 X 光机已经广泛的应用于各种类型的安检场景, 经过培训的工作人员判读 X 光机生成的二维扫描图像后, 给出安全放行或者拦截开检的结论。由于二维图像中可能存在物体之间相互遮挡以及复杂多样的背景物体的干扰, 机场、海关等场所正在逐步增加更先进的安检 CT 扫描设备的部署应用^[1]。三维 CT 图像为安检员准确辨识禁止运输或携带的目标物提供了更多的信息, 使得改变不同的视角检查图像成为了现实, 并为基于密度和等效原子序数等特征区分不同的材料提供了可能性。研究表明, 虽然目前的三维图像质量与二维 X 光机图像相比视觉感受仍有不足, 三维成像能够使安检员更好的辨识威胁^[2]。

无论是判读二维 X 光机图像还是三维 CT 图像, 人类安检员的表现存在主观性, 且严重依赖于经验, 疲劳、工作单调乏味等因素导致注意力不集中, 影响了安检质量的可靠性和稳定性。为了缓解上述问题, 实际场景中也采取了基于危险品图像插入 (threat image projection, TIP) 的培训和考核、更短的换班时间等一系列积极的措施, 成本也不断增加, 却并不能有效消除安检漏洞, 毕竟安检员需要在极为有限的时间内从混乱的背景物体中辨识出包裹中存在的可能威胁, 保证正常的安检通过率, 这是一个难度极大的挑战性任务。因此, 任何自动化、半自动化或辅助完成识别任务、减轻安检工作人员压力的尝试都是极有意义的。

由于二维 X 光机图像模态更接近研究广泛的可见光图像且存在一批公开数据集, 对于二维安检图像中目标物自动识别方法的研究较为深入^[3-4]。然而对于三维 CT 图像, 简单的将二维方法拓展到三维并不可行, 一方面由于维度的变化使得三维数据量增长了 2~3 个数量级, 对于识别效率的要求却并未降低; 另一方面由于算力和内存限制, 难以利用与二维方法中同样复杂的深度神经网络结构。

此外, 与三维图像中目标物识别相关的研究来自于医学和自动驾驶等领域。相比安检 CT 领域仅有公开数据集 ALERT 包含 188 个包裹 421 个目标物, 医学领域已有 DeepLesion、NLST、TCIA、LIDC-IDRI、MICCAI、ISBI、BraTS、LiTS、WORD 等一批公开三维数据集, 其中 DeepLesion 包含约 1 万组

收稿日期: 2023-07-31。

基金项目: 国家重点研发计划 (针对民航违禁品的复杂背景下安检三维图像智能识别技术研究 (2022YFC3310200))。

CT 扫描图像的超过 3.2 万个病灶标注, NLST 收集了超过 7.5 万组肺部 CT 扫描图像的超过 17 万处异常的信息。自动驾驶领域也有 KITTI、Waymo、ApolloScape、BLVD、Synthia、Oxford RobotCar、nuScenes、Argoverse、A2D2 等 10 余个公开三维数据集, 其中 KITTI 针对 3D 目标检测任务提供了 1.5 万张图像超过 8 万个标记对象, Waymo 含有超过 20 万帧图像的约 1200 万个 3D 标注框。3D CNN 模型已被用于医学 CT/MRI 图像^[5-6]、激光雷达点云数据^[7-8]、RGBD 数据^[9-10]的物体分类和检测。

Hu 等^[5]对基于医学图像的肿瘤检测和诊断研究工作进行了综述, 其中多数方法应用了 3D CNN 的技术; Monkam 等^[6]回顾了基于 3D CNN 的 CT 图像中肺结节检测和分类工作进展。医学 CT/MRI 图像具有亚毫米级近各向同性的分辨率, 自动识别研究一般限定于特定的器官或疾病, 可以利用任务有关的先验信息, 器官数量和布局是固定已知的, 任何偏离将被视为异常。然而在安检场景中由于高通过率的要求, CT 图像普遍具有毫米级各项异性的分辨率, 图像质量相对较低且包含较多的噪声和金属伪影; 扫描对象也是无限制的, 包裹中可能包含不确定数量和种类的物品, 即便已知物品的种类和数量, 其组合和相对位置也有无数种可能, 因此可利用的先验信息很少。

在自动驾驶领域, 感兴趣目标如车辆行人等有相对固定的尺寸和朝向, 行人的 3D 包围框在不同的场景中有类似的尺寸比例, 行人应当是站立的姿态, 这些先验知识可以用于生成候选框和减少误报, 然而这种先验知识在行包目标检测中不成立, 行包中的违禁品可以有任意的尺寸、朝向和位置。自动驾驶数据集如 KITTI^[11]中一组典型的点云包含约 1 万个点, 点云中复杂的背景物体通常与目标物是分离的, 相对容易处理, 然而一组典型安检 CT 扫描数据通常包含 1 亿个以上的体素, 箱包中的目标物往往与其它物品距离很近且交叠在一起。总之, 安检 CT 图像的目标物识别方法也不能机械借鉴医学或自动驾驶领域, 仍需要兼顾准确率和效率针对性的开展研究。

虽然安检 CT 图像自动识别领域已存在几篇综述性文献, 如文献 [4、12-13]。其中 Crawford 等^[12]和 Mouton 等^[13]仅总结了 2015 年以前的工作, 缺少近年来的特别是深度学习技术的应用研究进展; Velayudhan 等^[4]涉及的内容包括了对 X 射线成像系统的介绍和数据集, 但重点在于 2D X 射线图像中的危险目标识别, 对于 3D 图像中的危险目标识别方法的分析不够全面, 且着重于经典图像分析和机器学习方法。本文不仅对于 2015 年以来本领域的代表性学术研究成果进行系统总结, 而且首次将工业界的相关专利也纳入分析范围, 从而使本领域研究人员能够更全面的了解最新的研究进展和发展趋势。

1 归类法

已有的安检图像自动识别综述性文献中常见两种归类法, 一种是根据技术路线划分, 一般包括经典图像分析方法、机器学习方法和深度学习方法; 另一种是根据任务类型划分, 一般包括目标物分类方法、检测方法和分割方法。然而既有研究工作往往会综合使用多种技术路线, 并在多个任务类型上进行实验验证, 使得上述两种归类法的可分性略显不足。

本文除了继续沿用这两种归类法之外, 根据计算空间维度的差异提出一种新的归类法, 即二维识别方法和三维识别方法, 前者在二维空间完成最主要的识别计算, 其中有的方法会将二维识别结论整合以获得三维空间中的识别结论, 后者则直接在三维空间完成所有的识别计算并输出识别结论。

根据上述 3 种归类法, 我们对检索到的代表性学术研究成果和相关专利进行多标签归类, 归类结果列于表 1 中, 每篇文献的主要技术思路简述于本文的第 2 和第 3 部分。文献及其简要描述按论文发表或专利公开时间顺序一览如图 1。为了增强可分性、避免交叉, 在第 2 和第 3 两部分的叙述中我们仅使用计算空间维度归类法。

2 论文方法简述

2.1 二维识别方法

文献 [14] 基于箱包中的大多数液体包括危险液体都有一个水平表面的特点, 提出了一种两阶段的几何拟合方法以自动定位 CT 数据中的液体表面, 而无需样本训练过程。第 1 阶段使用 RANSAC 技

术在连续多个水平切片中拟合一个三维平面方程的参数，第 2 阶段在第 1 阶段结果的基础上拟合一个二维椭圆方程的参数。测试数据使用 Reveal 公司的 CT-80XL 采集并插值为各项同性的 CT 数据，测试集为 100 个包裹，其中 15 个没有包含液体。第 1 阶段液体表面所在平面的检出率达到 98%，两个阶段后液体表面具体位置的检出率为 85%、误报率为 12%。

表 1 3 种归类法及文献归类结果
Table 1 Three taxonomies and literature classification results

归类法	技术路线			任务类型			计算空间维度	
	经典图像分析方法	机器学习方法	深度学习 方法	分类方法	检测方法	分割方法	二维识别方法	三维识别方法
文献多 标签归 类结果	[14][19]	[17][18]	[15][16] [23][24]	[16][17]	[14][15] [16][19]	[15][16]	[14][15]	[17][18]
	[20][22]	[20][21]	[25][26]	[20][21]	[23][24]	[18][19]	[16][26]	[19][20]
	[26][29]	[22]	[27][28]	[22][23]	[25][26]	[20][21]	[27][28]	[21][22]
	[30][34]		[29][30]		[27][28]	[22][25]	[29][30]	[23][24]
			[31][32] [33]		[29][30] [31][33]	[26][29] [30][32]	[31][32] [33]	[25][34]

注：黑色字体文献为学术论文，蓝色字体文献为专利申请。

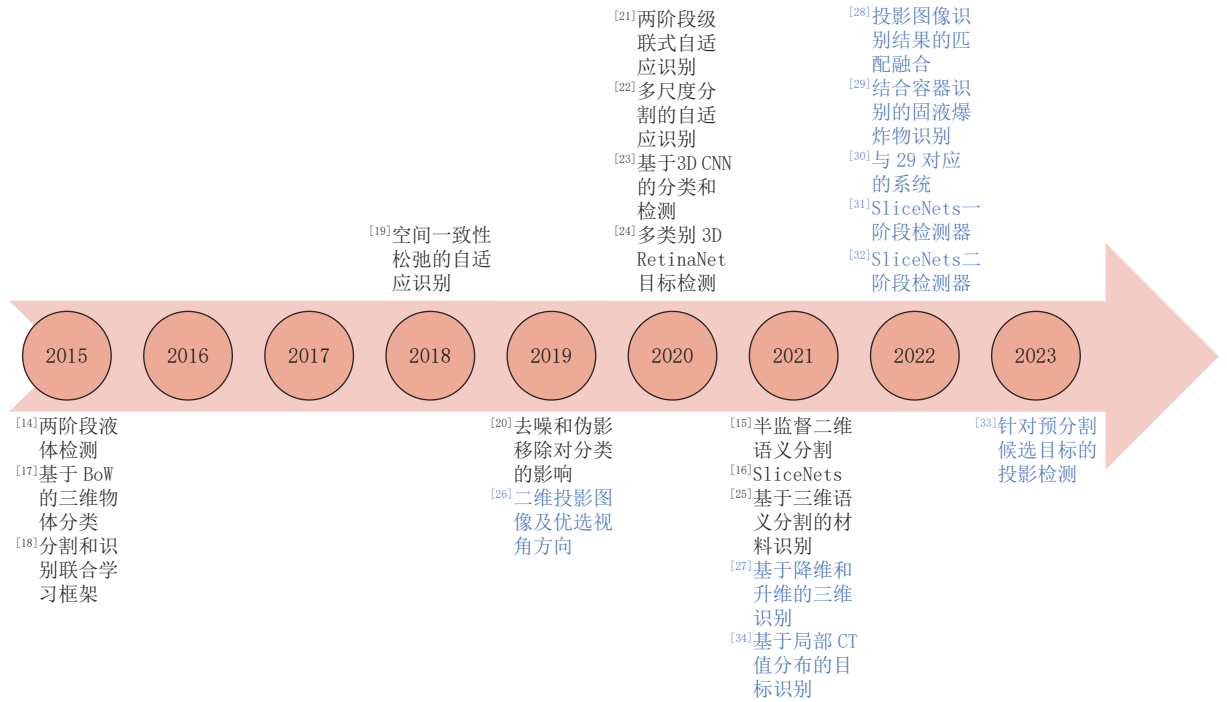


图 1 文献按论文发表或专利公开时间顺序一览图

Fig.1 Chronological overview of relevant literature

文献 [15] 提出了一个基于伪标签的半监督学习框架，使用有限的切片标注数据训练二维语义分割网络，识别无固定形状特征的违禁品物质。以少量人工标注的切片数据训练一个语义分割模型，使用该模型对其它不带标注的切片数据进行伪标注，然后用 3 个方向的所有切片数据及其标签重新训练一个语义分割模型。预测时 3 个方向的分割结果可以使用投票方式融合获得更准确的分割结果。对于分割结果进行形态学运算和联通区域分析等后处理，获得目标检测结果。在 ALERT 数据集 [35] 上实验，结果表明即便使用有限的切片标注数据，半监督学习方法能够获得与全监督方法相近的违禁品物质识别性能，各类别综合的精度/召回达到 91.2% 和 87.5%，检出率/误报率达到 90% 和 3%。

文献 [16] 提出了一种先切片后融合的识别框架，并设计了两种检测器，即一步法的 Retinal-SliceNet 和两步法的 U-SliceNet。切片操作沿着 3 个正交坐标轴方向进行，通过特定厚度切片的最

大池化, 将体数据转换成若干幅 2D 图像。对于 Retinal-SliceNet 检测器, 获得每幅 2D 图像中的目标包围框和置信度, 将每个坐标轴方向上的 2D 预测结果累积成 3D 预测结果, 对 3 组 3D 预测结果逐体素平均池化完成融合操作。对于 U-SliceNet 检测器, 获得每幅 2D 图像的像素级类别标记, 通过上述累积和融合操作得到 3D 预测结果, 基于 3D 预测结果设定若干个锚框, 使用枪支 3D 分类网络和锐器 3D 分类网络分析每个锚框, 输出整个包裹的分类结论, 或者生成目标物更精细的语义分割结果。实验使用 IDSS 公司 DETECT 1000 采集的含枪支和锐器的 1001 组真实数据和 15700 组合成数据, 以及 11400 组无目标物数据, 得出了 Retinal-SliceNet 在箱包分类和目标检测任务、U-SliceNet 在语义分割任务上的定量实验结果, 单包单卡推理计算时间约为 5~20 秒。

2.2 三维识别方法

文献 [17] 实验对比研究了基于词袋模型 (bag of words, BoW) 从人工分割的包裹 CT 图像中识别危险品的适用性, 对比 4 种 3D 描述子, 即密度直方图、密度梯度直方图、3D SIFT、3D RIFT, 还对比了 3 种码书赋值方式, 从而决定了最优的模型参数组合。实验数据为使用 Reveal 公司 CT-80DR 采集的 CT 数据并经过各项同性插值。其中手枪正负样本分别为 284/971 组, 瓶子正负样本分别为 534/1170 组。在人工分割的手枪和瓶子的分类实验中, 3D SIFT 关键点位置采样的密度直方图描述子具有最高的识别准确率, 瓶子的检出率 89.3% 误报率 3.0%, 手枪的检出率 97.3% 误报率 1.8%, 同时基于 3D SIFT 和 3D RIFT 描述子的识别性能明显降低, 体现了光学反射成像和透视成像两种模态的差异。

文献 [18] 没有采用将物体分割和识别作为两个独立步骤的常规做法, 提出了一种分割和识别任务联合学习框架。该框架定义了多标签 (感兴趣目标、金属、背景) 离散优化目标函数并用图割算法求解。目标函数有两项, 一是基于 P-KNN 算法^[36]得到每个体素属于某种标签的似然函数值, 其中靠近金属区域体素的权重将被降低, 金属区域由二值 (金属、非金属) 离散优化目标函数估计; 二是基于边界场计算相邻体素标签的空间平滑函数值。实验仅用到了 ALERT 数据集中两个包裹的切片, 一个包裹用于训练, 另一个包裹用于测试, 虽然体素标签估计的定性和定量结果表明了方法的有效性, 但未经性能优化的算法 CPU 处理时间长达 2 小时。

文献 [19] 着重探讨了从 CT 图像中自适应识别违禁物品的方法, 给定拟识别的物体需求描述 (object requirement specification, ORS), 包括感兴趣目标的材料组成 (如密度或线性衰减系数的范围)、物理特征 (如质量和厚度)、性能指标 (检出率和误报率), 系统不必重新训练分类器, 即可适用于变化的需求。为此作者提出一种空间一致性松弛方法, 此方法首先利用 ORS 生成的条件概率密度函数为每个体素计算最可能的材料组成, 将 CT 图像转换为 ID 图像, 从而将 CT 图像分割为若干子区域, 为每个体素和每个子区域计算分类统计值, 将分类统计值与决策阈值进行比较以判定正确检出或误报的体素和子区域。所谓松弛体现在每个体素的分类统计值是在其空间邻域内计算的条件概率密度值的平均值, 一致性松弛参数即邻域大小, 参数越大, 分割的子区域越完整, 但过大的参数可能会导致欠分割。文中提出了 3 种性能评估方法, 分别是基于子区域的性能评估 (包括材料组成和物理特征)、基于体素的性能评估 (仅包括材料组成)、基于体素的性能评估 (包括材料组成和物理特征), 并分别在 ALERT 数据集上进行实验, 给出了不同决策阈值下检出率和误报率的 ROC 曲线。文中还讨论了结合 ORS 中的性能指标需求和 ROC 曲线自动估计不同违禁物品决策阈值的方法。本方法使用 CPU 单核处理一幅 $512 \times 512 \times 400$ 的单能 CT 图像约需 15~30 秒。

文献 [20] 提出一种 3D 物体分割和分类方法, 其中无监督分割基于材料属性 (即双能索引值), 首先使用多段阈值化和连通成分分析获得粗分割结果, 训练一个随机森林评估器对粗分割得到的每个物体计算质量评价分数, 分数过低的物体需要细分割, 细分割通过寻找并移除物体表面体素中的“热点”结合形态学操作实现。在 3D 物体分类阶段, 原始特征为密集采样的 3D 关键点位置的密度直方图, 每个物体的特征向量由 ERC 随机森林^[37]对原始特征进行编码得到, 分类器为使用 RBF 核函数

的多类支持向量机。实验数据为 Reveal CT-80DR 采集的箱包数据, 手枪正/负样本分别有 208/150 组, 瓶子正/负样本分别有 146/190 组。该工作对比了 4 种去噪或金属伪影移除 (metal artefact reduction, MAR) 预处理方法, 即简单的阈值操作、非局部平均^[38]、基于线性插值的 MAR、距离驱动的 MAR, 发现两种 MAR 方法对分类准确率提升有限但会导致运行时间的严重增加。综合性能最优的非局部平均预处理方法, 手枪的检出率 TPR/误报率 FPR 分别为 96.63%/1.33%, 瓶子的 TPR/FPR 分别为 100%/2.11%。

文献 [21] 提出了两阶段的方法解决自适应目标识别问题, 即在第 1 阶段争取尽量高的召回率且无需针对变化的识别需求重新调整, 在第 2 阶段依据新需求对系统整体性能进行精调。其中第 1 阶段使用基于图的算法分割得到候选物体集合, 第 2 阶段基于归一化密度直方图等特征训练随机森林分类器。识别需求变化时, 使用高斯加权函数对训练样本的权重进行动态调整后重新训练第 2 阶段的分类器。ALERT 数据集上的实验表明, 该自适应方法的识别性能不差于用传统方法从头训练一套分类器, 成本则显著降低。

文献 [22] 研究了爆炸物等主要依靠材料属性识别的危险品自动识别 (automatic threat recognition, ATR) 技术, 包括自适应识别问题。该项工作首先设计多尺度形态学算子, 通过反复迭代进行基于形状特征的 3D 图像分割, 然后进行基于强度特征的分割, 以处理贴合紧密但材料不同的物体; 随后训练 4 类 SVM 分类器以识别 ALERT 数据集中的盐水、橡胶、粘土和其它物质等 4 个模拟物类别。该项工作还通过加上一个根据检出率需求确定的偏移值调整分类器输出概率, 实现了能够自适应危险品类型和检出率等可变需求的 ATR。

文献 [23] 探索了 3D CNN 和主流目标检测框架在安检 CT 图像违禁品分类和检测中的应用, 其中 ResNet 和 VoxelNet-ResNet 被用于 3D 违禁品分类的实验, Faster R-CNN 和 RetinaNet 被用于 3D 违禁品检测的实验。实验数据来自于 Reveal CT-80 设备采集的瓶子和手枪的图像。用于分类实验的人工分割的瓶子数据集 1704 个样本, 手枪数据集有 1255 个样本, 评估指标为检出率 TPR 和误报率 FPR, 两类目标物的 TPR 约 98%, FPR 约 1.5%。用于检测实验的瓶子数据集有 305 个箱包 526 个实例, 手枪数据集有 267 个箱包 282 个实例, 检测实验的评估指标为平均精度 (mean average precision, MAP), 瓶子和手枪的 mAP 分别约为 76% 和 88%。

文献 [24] 在文献 [23] 的基础上尝试用单一模型解决多类别 3D 目标检测问题, 基于多种不同的骨干网络设计了 3D RetinaNet。由瓶子、手枪、望远镜、Glock 套筒座、iPods 5 个类别构成的数据集包含 478 组真实 CT 图像和 287 组合成 CT 图像。5 类最优的召回 Recall 分别为 74%、91%、87%、54% 和 57%, 对应的精度 Precision 分别为 81%、75%、78%、64% 和 33%, 其中 Glock 套筒座和 iPods 两类的误报较多导致召回率和精准度较低。实验发现合适的数据增强技术有利于提升检测器的性能。

文献 [25] 研究了从 CT 体数据中检测炸药、毒品等无特定形状但由特定材料构成的物质, 把这一问题定义为三维语义分割问题。该工作尝试两条不同的技术路线, 一是应用 3D U-Net 及其变种从三维体数据中提取特征并给出体素级的语义分割结果; 二是先通过体素筛选和采样将三维体数据转换成点云数据, 再应用 pointnet^[39]/pointnet++^[40] 从点云数据中提取特征并给出体素级的语义分割结果。两条技术路线都需要对分割结果进行后处理运算将分割结果转化为目标检测结果。在 ALERT 数据集上测试对比两种技术路线的性能, 指出第 1 条技术路线的分割和检测效果明显优于第 2 条技术路线。此外, 该研究仅分析了网络推理步骤的计算复杂度, 缺少整个检测过程的运算效率。

3 专利方法简述

3.1 二维识别方法

文献 [26] 基于 CT 断层图像判断物体中是否有燃爆物, 给出燃爆物报警; 基于物体的二维投影

图像完成燃爆物之外的违禁品识别, 给出违禁品报警; 二维投影图像可能是扫描直接得到, 或者从物体的原始投影数据重组而来, 或者从三维断层图像投影而来。其中, 将 CT 断层数据沿 Z 向 (即物体运动方向) 累加后进行 360° 正投影形成正弦图, 正弦图上宽度最大的行所对应的视角方向设定为优选的二维投影视角方向。

文献 [27] 可对三维 CT 数据设定任意多个方向 (不限于安检过程中物体行进方向的正交方向), 按照所述多个方向进行投影或渲染, 生成多个二维降维视图; 针对任意多个二维视图进行目标物识别, 获得目标物的二维语义描述集合; 将所述二维语义描述集合利用反投影技术映射到三维空间, 获得语义特征矩阵, 并将语义特征矩阵压缩为三维概率图; 对三维概率图进行特征提取, 例如连通区域分析、主成分分析等运算, 获得三维图像语义描述集合以及目标物的三维识别结果。

文献 [28] 围绕 X 轴、 Y 轴或 Z 轴对三维数据旋转不同角度, 并对每个角度下的三维几何结构体数据进行投影形成投影图像, 根据图像识别方法对不同角度的投影图像进行目标物识别, 获得每一幅投影图像中的若干个包围框。在沿着同一坐标轴旋转得到的若干幅投影图像中, 同一个物体的包围框的二维坐标中总有一维近似相同。基于这一特点, 设计坐标匹配算法判断某角度的投影图像中的包围框与另一角度的投影图像中的包围框的差异指标是否小于预设阈值, 若小于预设阈值则认为表征了同一个物体。最后同一个物体在所有投影图像中的包围框在三维空间中进行切割融合, 得到某一目标物的三维包围框。

文献 [29] 和文献 [30] 在投影图像中检测瓶、罐等容器的边界框, 根据不同角度的识别结果进行交集运算, 得到容器三维识别的边界框; 从每个二维断层中分割出若干个二维物体, 并根据二维物体特征相似性的综合评分, 将具有连通性的二维物体组合形成三维物体, 完成三维物体分割; 将每个三维物体的边界框与识别出的容器的边界框进行匹配, 判断三维物体是否被一个容器包裹着; 如果是, 则该三维物体进入液体和粉末状爆炸物识别库进行匹配; 如果不是, 则该三维物体进入固体爆炸物识别库进行匹配。

文献 [31] 和文献 [32] 两份专利与其他学术机构早期发表的 SliceNets 论文^[16]中的方法完全相同, 其中文献 [31] 等价于 SliceNets 中的一阶段检测器, 文献 [32] 等价于 SliceNets 中的二阶段检测器。

文献 [33] 将待检测目标的三维体数据转化为二维切片数据, 转化方法与 SliceNets^[16]雷同; 然后对二维切片数据利用 RPN (region proposal network) 等网络进行预分割得到多个单独的待检测子目标; 将每个待检测子目标进行多个方向的投影得到多个待检测子目标的二维投影图像, 利用预训练的目标检测模型, 对每个所述待检测子目标的二维投影图像进行目标检测识别, 获得对应的检测结果; 将所述检测结果按照投影时的方向反映射回 CT 重建图像中, 得到各个物品的三维包围框信息, 实现对待检测目标的三维目标检测。

3.2 三维识别方法

文献 [34] 基于密度、等效原子序数、形状等特征识别物品之间的边缘, 将三维图像分割为一系列的物品; 从物品中识别电子产品等容易藏匿炸药等违禁品的藏匿容器; 对所有物品或者藏匿容器, 通过沿着物品主轴分层等方式将物品划分为若干个子区域, 依据子区域中 CT 值 (密度和/或等效原子序数等) 的分布将每个子区域分类为安全或有嫌疑, 如果一个子区域被分类为有嫌疑, 则该物品乃至整个包裹即可判定为可能藏匿有违禁品。

4 结论

通过总结分析 2015 年以来安检 CT 图像目标物识别方法的代表性文献, 不难看出有以下几个特点。

(1) 本领域研究成果近 3 年来呈现明显上升态势, 特别是工业界出现了较多的专利申请: 主要原因在于部分国家地区或者行业出台了引导性的政策, 使得安检 CT 扫描设备部署应用规模显著增长,

从而吸引了更多的研究资源和企业投资。

(2) 学术界和工业界的研究路径和关注重点存在明显差异。学术界的技术路线和任务类型呈现多元化特点, 兴趣广泛, 同时计算空间维度以三维识别为主, 偏好技术难度更高、更富有挑战性的目标。工业界的技术路线则结合使用跨代的经典图像分析方法和深度学习方法, 任务类型聚焦于目标物检测, 计算空间维度以相对成熟的二维识别为主, 实用化色彩浓厚, 选择研究路径以成本低和能够满足现场应用需要为出发点。

(3) 本领域研究的广泛性和影响力仍然有较大的上升空间。从公开发表的文献数量和频度来看, 本领域与可见光图像中目标物识别领域的研究参与度不具有可比性, 即使与二维 X 光机图像中目标物识别领域相比研究力度也有一定的差距, 这一现象是与安检行业原始数据的敏感性以及安检 CT 公开数据集的稀缺性密切相关的, ALERT^[35]是本领域目前仅有的半公开数据集, 且并非实际应用现场的真实数据。

在系统回顾现有文献研究成果的基础上, 并深入思考本领域痛点和趋势, 我们针对新的研究方向提出如下 3 点建议。

(1) 研究巨大类内差异下的学习和表征。有的目标物类别中不同实例存在形状、结构或材料的显著变化, 如生活刀具、自制爆炸装置等, 依靠训练数据集覆盖所有变化或者数据增强是不现实的。

(2) 研究仅学习正常样本的目标物检测。本领域普遍存在无安全风险的正常样本, 包含违禁品的样本稀缺而且不同类别的样本数量也严重不均衡, 通过学习正常样本的模式发掘新样本中的异常势在必行。

(3) 研究机器智能与安检工作人员的最佳协同机制。机器完全代替人工存在较大安全风险时, 如何设计机器智能结论因人而异的展示时机和方式, 协助安检员在有限的判图时间内做出可靠决策, 将是安检 CT 图像目标物识别技术成功运用的关键一环。

参考文献

- [1] 李新斌, 张丽, 陈志强, 等. 行包货物实时验放 CT 智能解决方案[J]. *CT 理论与应用研究*, 2022, 31(5): 597-615. DOI:10.15953/j.ctta.2022.124.
LI X B, ZHANG L, CHEN Z Q, et al. CT intelligent solution for real-time inspection and release of baggage and cargo[J]. *CT Theory and Applications*, 2022, 31(5): 597-615. DOI:10.15953/j.ctta.2022.124. (in Chinese).
- [2] HATTENSWILER N, MENDES M, SCHWANINGER A. Detecting bombs in X-ray images of hold baggage: 2D versus 3D imaging[J]. *Human Factors*, 2019, 61(2): 305-321. DOI:10.1177/0018720818799215.
- [3] AKCAY S, BRECKON T. Towards automatic threat detection: A survey of advances in deep learning within X-ray security imaging[J]. *Pattern Recognition*, 2022, 122: 108245-108256. DOI:10.1016/j.patcog.2021.108245.
- [4] VELAYUDHAN D, HASSAN T, DAMIANI E, et al. Recent advances in baggage threat detection: A comprehensive and systematic survey[J]. *ACM Computing Surveys*, 2023, 55(8): 1-38.
- [5] HU Z, TANG J, WANG Z, et al. Deep learning for image-based cancer detection and diagnosis: A survey[J]. *Pattern Recognition*, 2018, 83: 134-149. DOI:10.1016/j.patcog.2018.05.014.
- [6] MONKAM P, QI S, MA H, et al. Detection and classification of pulmonary nodules using convolutional neural networks: A survey[J]. *IEEE Access*, 2019, 7: 78075-78091. DOI:10.1109/ACCESS.2019.2920980.
- [7] ZHOU Y, TUZEL O. Voxnet: End-to-end learning for point-cloud-based 3D object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4490-4499.
- [8] CHEN X, MA H, WAN J, et al. Multi-view 3D object detection network for autonomous driving[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6526-6534.
- [9] MATURANA D, SCHERER S. Voxnet: A 3D convolutional neural network for real-time object recognition[C]//International Conference on Intelligent Robots and Systems, 2015: 922-928.
- [10] QI C R, LIU W, WU C, et al. Frustum pointnets for 3D object detection from rgb-d data[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2018: 918-927.
- [11] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving? The KITTI vision benchmark suite[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2012: 3354-3361.

- [12] CRAWFORD C. Advances in automatic target recognition (ATR) for CT-based object detection systems[R]. Northeastern University, 2015.
- [13] MOUTON A, BRECKON T P. A review of automated image understanding within 3D baggage-computed tomography security screening[J]. *Journal of X-ray Science and Technology*, 2015, 23(5): 531–555. DOI:10.3233/XST-150508.
- [14] CHERMAK L, BRECKON T P, FLITTON G T, et al. Geometric approach for the automatic detection of liquid surfaces in 3D computed tomography baggage imagery[J]. *The Imaging Science Journal*, 2015, 63(8): 447–457. DOI:10.1179/1743131X15Y.0000000019.
- [15] WANG Q, BRECKON T P. On the evaluation of semi-supervised 2D segmentation for volumetric 3D computed tomography baggage security screening[C]//International Joint Conference on Neural Networks, 2021: 1–8.
- [16] YANG A, PAN F, SARAGADAM V, et al. SliceNets: A scalable approach for object detection in 3D CT scans[C]//IEEE Winter Conference on Applications of Computer Vision, 2021: 335–344.
- [17] FLITTON G, A. MOUTON A, BRECKON T P. Object classification in 3d baggage security computed tomography imagery using visual codebooks[J]. *Pattern Recognition*, 2015, 48(8): 2489–2499. DOI:10.1016/j.patcog.2015.02.006.
- [18] MARTIN L, TUYSUZOGU A, KARL W C, et al. Learning-based object identification and segmentation using dual-energy CT images for security[J]. *IEEE Transaction on Image Processing*, 2015, 24(11): 4069–4081. DOI:10.1109/TIP.2015.2456507.
- [19] PAGLIERONI D W, CHANDRASEKARAN H, PECHARD C, et al. Consensus relaxation on materials of interest for adaptive ATR in CT images of baggage[C]//Anomaly Detection and Imaging with X-rays (ADIX) III, 2018: 10632.
- [20] MOUTON A, BRECKON T P. On the relevance of denoising and artefact reduction in 3D segmentation and classification within complex computed tomography imagery[J]. *Journal of X-ray Science and Technology*, 2019, 27(1): 51–72. DOI:10.3233/XST-180411.
- [21] MANERIKAR A, PRAKASH T, KAK A C. Adaptive target recognition: A case study involving airport baggage screening[C]//Anomaly Detection and Imaging with X-rays (ADIX) V, 2020: 11404.
- [22] WANG Q, ISMAIL K N, BRECKON T P. An approach for adaptive automatic threat recognition within 3D computed tomography images for baggage security screening[J]. *Journal of X-ray Science and Technology*, 2020, 28(1): 35–58. DOI:10.3233/XST-190531.
- [23] WANG Q, BHOWMIK N, BRECKON T P. On the evaluation of prohibited item classification and detection in volumetric 3D computed tomography baggage security screening imagery[C]//International Joint Conference on Neural Networks, 2020: 1–8.
- [24] WANG Q, BHOWMIK N, BRECKON T P. Multi-class 3D object detection within volumetric 3D computed tomography baggage security screening imagery[C]//IEEE International Conference on Machine Learning and Applications, 2020: 13–18.
- [25] WANG Q, BRECKON T P. Contraband materials detection within volumetric 3D computed tomography baggage security screening imagery[C]//IEEE International Conference on Machine Learning and Applications, 2021: 75–82.
- [26] 不公告发明人. 一种CT检测方法及装置: 中国, 201910173197.1[P]. 2019-07-05.
- [27] 陈志强, 张丽, 孙运达, 等. 安检CT目标物识别方法和装置: 中国, 202211693747.0[P]. 2021-12-14.
- [28] 罗亮, 张海平, 范美仁, 等. 一种基于计算机断层扫描的三维图像融合方法及装置: 中国, 202111590868.8[P]. 2022-07-05.
- [29] 李保磊, 牛素莹, 魏增辉, 等. 一种爆炸物识别方法: 中国, 202211215419.X[P]. 2022-12-23.
- [30] 李保磊, 牛素莹, 魏增辉, 等. 一种爆炸物识别系统和方法: 中国, 202211214188.0[P]. 2022-12-09.
- [31] 李春阳, 何竞择, 张文杰, 等. 一种用于安检CT三维图像的目标物体检测方法及系统: 中国, 202111604528.6[P]. 2022-04-01.
- [32] 李春阳, 何竞择, 张文杰, 等. 一种用于安检CT三维图像的目标物体分割方法及系统: 中国, 202111603031.2[P]. 2022-03-29.
- [33] 张文杰, 魏增辉, 乔建森, 等. 一种基于深度学习的CT三维目标检测方法: 中国, 202211693747.0[P]. 2023-03-21.
- [34] SIMANOVSKY S, SCHAFER D. Item classification using localized CT value distribution analysis: America, US11087468B2[P]. 2021-08-10.
- [35] CRAWFORD C, MARTZ H, CASTANON D. Adaptive automated threat recognition[R]. Northeastern University, 2018.
- [36] DUDA R O, HART P E, STORK D G. Pattern Classification[M]. 2ed. New York: Wiley, 2000.
- [37] MOOSMANN F, TRIGGS B, JURIE F. Fast discriminative visual codebooks using randomized clustering

- forests[C]//Advances in Neural Information Processing Systems, 2007, 19: 985-992.
- [38] BUADES A, COLL B, MOREL J M. A non-local algorithm for image denoising[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2005: 60-65.
- [39] QI C R, SU H, MO K, et al. Pointnet: Deep learning on point sets for 3D classification and segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017: 652-660.
- [40] QI C R, YI L, SU H, et al. PointNet++: Deep hierarchical feature learning on point sets in a metric space[C]//Advances in Neural Information Processing Systems, 2017: 5099-5108.

Research Progress in Target Recognition Methods for Security Computed Tomography Images

SUN Yunda¹, WEI Yunchao², ZHANG Li^{3✉}

1. Nuctech Company Limited, Beijing 100084, China

2. School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

3. Department of Engineering Physics, Tsinghua University, Beijing 100084, China

Abstract: Researching target recognition methods for 3D computed tomography (CT) images is important for improving the reliability and stability of the security inspection quality in advanced security CT scanning equipment. This article systematically reviews the latest research progress and developmental trends in target recognition of CT images for security purposes. Not only representative academic research achievements of recent years in this field are summarized and reviewed, but also relevant industrial patents are included for the first time. For literature classification, in addition to two common taxonomies, a new taxonomy is proposed, and the retrieved representative literature is classified with multiple labels. In summarizing and analyzing the research achievements in this field, three important characteristics were identified, and the origins were analyzed. Finally, some suggestions are put forward for future research directions in this field.

Keywords: security CT; target recognition; intelligent recognition



作者简介: 孙运达, 男, 信号与信息处理专业博士, 主要从事安检领域人工智能技术的研究, E-mail: sunyunda@nuctech.com; 张丽[✉], 女, 清华大学工程物理系首席研究员、博士生导师, 主要从事辐射成像研究, E-mail: zli@tsinghua.edu.cn。