使用迁移学习的纺织品分析回收自动化 & 零次基 础模型

Yannis Spyridis Department of Computer Science Kingston University, London, UK y.spyridis@kingston.ac.uk

Abstract—自动化分拣对于提高纺织品回收的效率和可扩展 性至关重要,但从传感器数据中准确识别材料成分和检测污染物 仍然充满挑战。本文研究了使用标准 RGB 图像(一种具有成本 效益的传感模式)在自动化系统中的关键预处理任务。我们设计 了用于传送带设置的计算机视觉组件,以执行(a)四种常见纺 织类型的分类以及(b)非纺织特征如组扣和拉链的分割。在分 类方面,评估了几种预训练架构,通过迁移学习和交叉验证进行 测试,其中 EfficientNetB0 在验证集上达到了 81.25 % 准确 率的最佳性能。对于特征分割,采用了结合 Grounding DINO 开放词汇检测器和 Segment Anything Model (SAM)的零 样本方法,生成的掩码与真实值相比表现出优异的性能,mIoU 达到了 0.90。这项研究展示了使用 RGB 图像结合现代深度学 习技术的可行性,包括用于分类的迁移学习和用于零样本分割的 基础模型,以实现自动化纺织品回收流程中的基本分析步骤。

Index Terms—computer vision, deep learning, transfer learning, foundation models, textile recycling

I. 介绍

全球纺织行业面临重大可持续性挑战,每年都有大量的 消费者后纺织品进入垃圾填埋场或被焚烧 [1]。这种线性 模式消耗了大量资源,包括水、能源和原材料,同时导致 污染和温室气体排放 [2]。向循环经济转型,即纺织品的 再利用和回收,是减轻环境影响并从废物流中释放经济价 值的关键。有效的回收需要根据材料成分对纺织品进行准 确分类,这一过程仍是一个主要瓶颈。实现这一分类过程 的自动化对于提高纺织品回收项目在全球范围内的效率、 可扩展性和经济可行性至关重要 [3]。

当前的纺织品分拣主要是手工进行的,这一过程劳动力 密集,费时,而且在处理混合面料或复杂服装时,往往容 易出现不一致和错误 [4], [5]。手工分拣的速度限制了回 收设施的处理量,同时相关的劳动力成本可能会阻碍盈利。 此外,手工分拣可能无法始终可靠地识别出纽扣和拉链这 样的非纺织成分,在机械或化学回收过程之前必须移除这 些成分,以避免污染回收材料或损坏设备 [6]。因此,迫 切需要能够快速准确识别纺织材料和检测问题非纺织特征 的自动化系统。

本文通过提出一个为回收流程中的自动纺织品分析设计 的计算机视觉系统来解决这些挑战。该系统利用安装在传 送带上方的摄像机捕捉的标准 RGB 图像,模拟了一个真 实的工业分拣环境。当纺织品在传送带上暂停时,图像被 获取并由两个专门的深度学习模型进行处理。第一个是分 类模型,训练用于识别对回收流至关重要的四种常见纺织 品类别:棉、涤纶、棉涤混纺和粘胶涤纶混纺。第二个是 分割模型,旨在精确定位非纺织特征,特别是按钮和拉链, 从而可以实现其自动化移除。重要的是,这项工作完全依 Vasileios Argyriou Department of Networks and Digital Media Kingston University , London, UK vasileios.argyriou@kingston.ac.uk

赖于通过标准传感器捕获的真实世界图像数据,以反映实际的操作条件。

本论文的主要贡献如下:首先,我们展示了使用经济有 效的 RGB 图像在分类特定纺织品混纺材料和分割小型、 可能低对比度特征如纽扣和拉链的艰难任务中的可行性。 其次,我们提出了专门设计的分类和分割模型的设计和实 现,这些模型集成到自动化前处理阶段的工作流程中,以 用于纺织品回收。第三,我们使用在真实条件下收集的数 据集验证了这些模型的性能,为这些特定任务提供了基准。 重要的是,这项工作提供了关键的计算机视觉能力,可以 支持下游自动化过程,例如机器人分拣或去除污染物的精 准激光切割。

本文的其余部分组织如下:第 II 部分回顾了纺织品分 类和计算机视觉应用的相关文献。第 III 部分详细介绍了 我们的数据采集、数据集特征、预处理步骤、模型架构和 评估指标。第 IV 部分展示了分类和分割任务的实验结果。 最后,第 V 部分总结了本文。

II. 相关工作

纺织材料的自动化分析对于推动纺织行业中的高效分类 和回收过程至关重要 [7]。传统的人工分类在速度、成本 和一致性方面存在局限性,促使开发仪器化技术。光谱法, 尤其是近红外光谱 (NIR),结合化学计量学和机器学习分 类器,如支持向量机 (SVMs),已被广泛研究并证明对于 识别纤维成分有效 [8]-[13]。其他方法可能包括提供详细 表面信息的高光谱成像和光度技术 [14]。虽然功能强大, 这些方法往往需要专业的、可能昂贵的设备,并可能在传 感器校准和数据融合方面面临挑战,如果使用多种模态。

同时,使用标准 RGB 摄像头的计算机视觉(CV)提供了一种可能更低成本且更易于访问的替代方案来分析纺织品。CV 已被应用于纺织行业的各种任务中,例如质量控制和自动缺陷检测[15]-[19]。然而,仅通过 RGB 图像识别纺织材料成分具有挑战性,因为材料之间的细微视觉差异以及纹理、颜色和光照的变化。一些研究探索了基于 RGB 的特定织物颜色的分类[20],[21]。卷积神经网络(CNNs),包括诸如 VGG [22]和 EfficientNet [23]等架构,已成为图像分类任务的标准,通常通过对大数据集(如 ImageNet [24],[25])上预训练模型的迁移学习而有效利用。有效应用这些技术来区分常见服装材料,包括视觉上相似的混合物,仍然是一个活跃的研究领域。

除了材料识别之外,准备纺织品进行回收往往需要检测 和定位如纽扣和拉链等非纺织污染物。这属于计算机视 觉中的目标检测和语义分割领域。常用的方法通常依赖 于在大数据集上的带有特定物体注释的训练来训练基于 CNN 的模型,如 Faster R-CNN、YOLO、Mask R-CNN 或 U-Net。最近,一种新的模式出现了,即基础模型,它 们是预先在大量数据集上训练的大模型,可以适应各种下 游任务,基本不需要或不需要特定于任务的训练数据。对 于分割,Segment Anything Model (SAM) 提供了强大的 可提示分割能力。SAM 可以根据各种提示生成蒙版,例如 点、边界框或先前的蒙版。最近,像 Grounding DINO 的 模型已经被引入,可以接受文本提示来执行开放词汇的目 标检测,能够在不需要对这些类别进行显式训练的情况下, 根据任意描述来定位物体。这些文本引导的检测器可以提 供用作分割模型如 SAM 的提示的边界框。在工业环境中 特别是对于小物体如衣物扣件的零样本检测和分割管道的 应用,可以为传统的监督方法提供一种新颖的替代方案。

无论是通过监督训练还是应用预训练系统,为这些任务 开发和严格评估模型都需要相关的数据集。一个显著的挑 战是缺乏针对基于 RGB 的纺织材料识别和与回收相关的 污染物检测而特别编制的大型公开可用数据集。虽然存在 用于织物缺陷检测的数据集 [26], [27],但它们没有解决分 类材料类型或为分拣预处理所需的特征分割。这种公共数 据的缺乏阻碍了稳健算法的开发和基准测试。

III. 方法论

本节详细介绍了本研究中采用的实验方法。我们描述了 获取图像数据的过程、所得数据集的特性、应用的预处理 步骤、分类和分割模型的架构以及训练过程。为了在数据 集大小的情况下保证评估的稳健性,采用了5折交叉验证 策略进行性能评估,并使用标准指标来量化模型性能。

A. 数据采集

本研究中使用的图像数据是模拟自动化纺织品分拣设施 的条件而收集的,该设施利用了传送带系统。纺织品平放 在传送带表面。当传送带处于静止状态时触发图像捕捉, 以确保尽量减少运动模糊。采用了一部现代智能手机作为 采集设备,其具有 1/1.33 英寸传感器,f/1.8 光圈,并设 置为以 12 百万像素的分辨率拍摄图像。这一选择反映了 在实际系统中使用可获得组件的潜力。为了保持数据集的 一致性,所有图像都是在相似的室内环境光照条件下拍摄 的。重要的是,拍摄后对原始图像没有进行后期处理或数 字增强,以保持数据对真实传感器输出的忠实性。

B. 数据集

获取的图像构成用于开发和评估计算机视觉模型的数据 集的基础,该数据集完全由如上所述获取的真实世界 RGB 图像组成。整个数据集由 80 张独特的图像组成。在纺织品 分类任务中,该数据集支持在与回收相关的四种关键材料 类型中的识别:棉、涤纶、棉涤混纺和粘纤涤纶混纺。对 于这四个类别中的每一个,提供了 20 张不同的图像,从 而形成了一个用于训练和验证的平衡数据集。

在分割任务中,使用了包含纽扣和拉链的图像。这些特 征在图像中被手动标注,以创建真实值掩膜,从而能够评 估像素级定位的分割性能。数据集中所有图像都针对这些 特征进行了手动标注,创建了像素级的真实值掩膜,以作 为评估分割结果的参考。

鉴于整个数据集的规模有限,采用了五折交叉验证策略 进行分类模型的训练和验证,最大限度地利用可用数据并 提供更可靠的泛化性能估计。80 张图像被随机分为 5 个 相互独立的折,每个折包含 16 张图像,并在各折内保持 类别平衡。在每次折迭代中,80 % 的数据用于训练,20 % 用于验证,所有折轮流进行。这种方法确保了每张图像在 5 个折中恰好被用于一次验证。

在分类模型的最终测试阶段,为棉、聚酯和棉-聚酯类别 分别构建了一个单独的测试集,每个类别包含6张图像。 对于粘胶-聚酯类别,我们加入了更多样本,使测试集总共 包含14个样本。这就形成了一个32张图像的最终测试 集,该测试集完全独立于训练和验证过程,以确保模型泛 化的无偏评估。第IV节中报告的分类任务的最终性能指 标代表了在此专用测试集上的结果,并辅以在模型开发期 间通过5折交叉验证获得的见解。

C. 数据预处理

将输入图像系统地预处理,以确保其适合用于分类任务 中所采用的不同神经网络架构。首先加载原始图像,然后 进行空间归一化,以匹配每个模型所需的特定输入维度, 如表 I 所详述。此过程包括从图像中提取一个中心裁剪, 尺寸根据原始维度与目标维度的最小值决定,然后使用双 三次插值将该裁剪调整为准确的目标分辨率。

在进行空间调整和颜色空间转换后,像素值被线性缩放 到范围 [-1, 1]。这一步归一化操作确保输入数据符合所用 模型的期望值范围。

为了增强模型的泛化能力和减轻过拟合问题,尤其是在 给定数据集大小的情况下,数据增强仅在分类模型的训练 阶段应用。在每个批次的训练图像上动态应用了一组随机 变换。具体的增强技术及其对应的参数列在表 II 中。

对于纺织品分类任务,我们评估了五种深度卷积神经网络架构的性能,这些架构利用从在 ImageNet 数据集上预训练的模型进行迁移学习。选定的架构包括 VGG16、VGG19、 EfficientNetB0、EfficientNetV2-S 和 EfficientNetV2-M。这些模型提供了一系列计算复杂性,并在各种视觉任务中展示了强大的性能。在预处理中,使用了每种模型所需的具体输入图像尺寸,详细信息如??中所述。

对于每种架构,卷积基被初始化为其 ImageNet 预训练 权重。原始的分类层被移除,并用一个适合我们四类纺织 品识别问题的新头替代,该头由一个全局平均池化 2D 层, 后面接一个具有 Softmax 激活的 Dense 层组成。

在 5 折交叉验证框架内采用了两阶段的训练策略,以有 效地将预训练模型适应于我们的特定数据集:

阶段 1:特征提取:最初,仅训练新添加的分类头的权重,而预训练的卷积基的权重保持冻结。此阶段的目的是使分类器适应基础模型提取的特征。此阶段使用 Adam 优化器运行最多 15 个周期,学习率为 1e – 3。

TABLE I 分类模型架构和相应的输入维度

Model Architecture	Input Dimensions (pixels)
VGG16	224×224
VGG19	224×224
EfficientNetB0	224×224
EfficientNetV2-S	384×384
EfficientNetV2-M	480×480

TABLE II 在训练期间应用的数据增强技术

Augmentation technique	Parameter range
Random Rotation	Up to 90 degrees
Random Width Shift	Up to 30 % of image width
Random Height Shift	Up to 30 % of image height
Random Zoom	Up to 30 % (zoom in or out)
Random Horizontal Flip	Applied with 50 % probability

TABLE III 模型训练配置和超参数设置

Hyperparameter	Value / Setting	
Pre-trained Weights	ImageNet	
Classification Head	Global Average Pooling 2D +	
	Dense (4 units, Softmax)	
Training Phases	1: Feature Extraction (head	
	only), 2: Fine-tuning (head $+$	
	top 30 base layers (4-6 for VGG)	
Optimiser	Adam	
Learning Rate (Phase 1)	1e - 3	
Learning Rate (Phase 2)	1e - 5	
Loss Function	Categorical Cross-Entropy	
Batch Size	4	
Max Epochs (Phase 1)	15	
Max Epochs (Phase 2)	50	
Early Stopping	Monitor: val_loss, Patience: 15,	
	Restore Best Weights: True	
Reduce LR on Plateau	Monitor: val_loss, Patience: 10	
	Factor: 0.2, Min LR: $1e - 6$	
Model Checkpoint	Monitor: val_accuracy, Save	
	Best Only: True	
Evaluation Framework	5-Fold Cross-Validation	

阶段 2: 微调:在初始阶段之后,解冻部分预训练的基础模型层以允许与分类头一同进行微调,使模型能够更好地适应纺织品数据集的特征。具体来说,对于 EfficientNet 模型,解冻了顶部的 30 层,而对于 VGG 模型,分别解冻了顶部的 4 层 (VGG16)和顶部的 6 层 (VGG19)。训练继续进行最多 50 个额外的周期,使用 Adam 优化器,并将学习率显著降低至 1e-5。在此微调阶段,模型能够更好地将提取的特征适应于纺织品数据集的特性。在两个训练阶段中,均最小化了分类交叉熵损失函数。训练使用批量大小为 4 进行。在每个折叠的训练过程中,采用了几种回调函数来管理培训过程:

为了优化训练并防止过拟合,使用了几个回调函数:

- 早停:用15个周期的耐心监控验证损失。如果没有观察到改善,则停止训练,并恢复验证损失最低的周期的权重。
- 平台上的学习率下降:监测验证损失,在连续 10 个 周期内没有改善的情况下,以 0.2 的因子降低学习率, 最低阈值为 1×10⁻⁶。
- 模型检查点:保存了每个交叉验证折叠中两个训练阶段中验证准确率最高时对应的权重。

第 IV 节中报告的所有结果都基于每个模型表现最好的 折叠。用于训练分类模型的关键超参数汇总在表 III 中。

对于识别和定位服装图像中非纺织特征的任务,采用了 一种利用最新预训练基础模型的方法,避免了在此特定数 据集上训练专门的分割网络的需要。这种方法结合了用于 文本提示的开放词汇对象检测的 Grounding DINO [28] 和 用于高保真提示分割的 SAM 模型 [29] 。

操作流程如下:每张输入图像会与特定的文本提示一起 处理,这些提示对应关注的特征,主要是"纽扣"和"拉 链"。Grounding DINO 模型会解释这些提示以执行开放 集检测,识别出图像中可能对应文本描述的候选区域。此 阶段的输出由边界框组成,划定了检测到的纽扣和拉链实 例。可以应用与检测器相关的标准置信度阈值来过滤检测 结果,不过默认参数通常已经足够应对此应用。

随后,由 Grounding DINO 生成的边界框作为输入提示 传递给 SAM 模型。SAM 在原始图像和这些空间边界框提 示的条件下,生成了精确的像素级分割掩码和对应对象的 轮廓。这个输出精确地划分了属于检测到的扣子或拉链的 边界和像素区域。

这一协同管道利用了 Grounding DINO 的零样本能力, 根据语言描述定位对象,并结合 SAM 强大的分割能力, 通过空间线索进行指导,从而无需特定任务的训练即可实 现目标特征的分割。该方法的主要优势在于能够利用大规 模的预训练,可能减少通常需要的广泛标注数据,这些数 据在有监督的分割模型训练中是必需的,尽管标注数据仍 然是评估所必需的。

通过将预测的分割掩码(或由轮廓隐含的区域)与手动 创建的像素级真实值注释(在第 3.2 节中描述)进行比较, 对该特征分割管道的性能进行了定量评估。使用第 3.6 节 中详细说明的标准分割指标,对包含这些注释特征的数据 集中的全部 80 张图像进行了评估。

为了定量评估开发的纺织品分类模型和特征分割方法的 性能,采用了标准且成熟的评估指标。针对多类分类和像 素级分割任务使用了不同的适用指标。

1) 分类指标: 纺织品分类模型的性能通过交叉验证程序的5个折叠结果的汇总所产生的指标进行评估。这些指标依赖于每个类别基础上考虑的真正例(TP)、真负例(TN)、假正例(FP)和假负例(FN)的计数。

准确率: 表示所有类别中正确分类实例的整体比例。

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \qquad (1)$$

精确率: 衡量被预测为属于类别 *i* 的实例中实际属于类别 *i* 的比例。它反映了预测的准确性。

$$\operatorname{Precision}_{i} = \frac{TP_{i}}{TP_{i} + FP_{i}} \tag{2}$$

召回率:衡量模型正确识别出分类 i 的实际实例的比例。 它反映了分类的完整性或覆盖率。

$$\operatorname{Recall}_{i} = \frac{TP_{i}}{TP_{i} + FN_{i}} \tag{3}$$

F1-分数:精确率与召回率的调和平均,为平衡这两个指标提供了一个单一的评分。当类别分布不均匀时,它特别有用。

$$F1_i = 2 \times \frac{\operatorname{Precision}_i \times \operatorname{Recall}_i}{\operatorname{Precision}_i + \operatorname{Recall}_i}$$
(4)

2) 分割指标:特征分割流程的性能通过将按钮和拉链的预测像素级别的掩码与数据集中所有相关图像的人工标注进行比较来评估。分割方法的评估在像素级别进行,并依赖于某些指标来衡量预测的分割掩码与真实掩码在每像素基础上的匹配程度。采用了以下标准指标:

TABLE IV 总体分类性能。结果是针对每种架构表现最好的分折(检查点)报告的。

Model Architecture	Accuracy	Weighted F1-Score
VGG16 (Fold 5)	0.4688	0.4646
VGG19 (Fold 4)	0.4375	0.4483
EfficientNetB0 (Fold 3)	0.8125	0.8012
EfficientNetV2-S (Fold 3)	0.8125	0.7927
EfficientNetV2-M (Fold	0.6562	0.6062
2)		

TABLE V 测试数据的每类 F1 分数。结果是针对每个架构表现最佳的折叠报告的。

Model	Cotton	Polyester	Cotton- Polyester	Viscose- Polyester
VGG16 VGG19 EfficientNetB0 EfficientNetV2-S EfficientNetV2-M	$\begin{array}{c} 0.4545 \\ 0.4444 \\ 0.8571 \\ 0.7500 \\ 0.6000 \end{array}$	$\begin{array}{c} 0.2857 \\ 0.4615 \\ 0.7692 \\ 0.8000 \\ 0.8000 \end{array}$	$\begin{array}{c} 0.1818 \\ 0.0000 \\ 0.4000 \\ 0.5000 \\ 0.0000 \end{array}$	$\begin{array}{c} 0.6667\\ 0.6364\\ 0.9630\\ 0.9333\\ 0.7857\end{array}$

交并比 (IoU): IoU 测量预测分割掩码和真实掩码之间 的重叠。它被定义为预测像素集和实际像素集的交集与它 们的并集之比:

$$IoU_i = \frac{TP_i}{TP_i + FP_i + FN_i} \tag{5}$$

IV. 结果与实验

本节详细介绍了在纺织品分类和特征分割任务中获得的 实验结果。我们基于分类的5折交叉验证中的最佳折叠, 呈现定量性能指标,随后是特征分割分析的结果。

根据第 III 节中描述的 5 折交叉验证训练和选择过程, 评估了五种架构中表现最好的折的检查点在未见测试数据 上的表现。这些模型在测试集上的整体性能总结在表 IV 中,比较了它们的准确率和加权平均 F1 分数。

表格清晰地显示了 EfficientNet 模型在测试数据上显著 优于 VGG 架构。EfficientNetB0 和 EfficientNetV2S 达到 了最高的准确率 (81.25 %),其中 EfficientNetB0 还获得了 最高的加权 F1 分数 (0.8012),这表明它在测试集的各类 中提供了最佳的精准率与召回率的平衡。EfficientNetV2M 表现稍弱,而两个 VGG 模型甚至未达到 50 % 的准确率 或 F1 分数。为了提供各类纺织品类别中性能的更细致视 图,表格 V 展示了每个模型在测试集上的每类别 F1 分 数。

表 V 中的每类 F1-分数揭示了特定的优势和劣势。表现 最好的 EfficientNetB0 模型在棉(0.8571)和最常见的类 粘胶-聚酯(0.9630)上表现出色,但在棉-聚酯(0.4000)上 显示出中等难度。EfficientNetV2S 虽然整体准确率相同, 但在聚酯(0.8000)和棉-聚酯(0.5000)上的 F1-分数比 EfficientNetB0 更高,但在棉和粘胶-聚酯上略有滞后。值 得注意的是,EfficientNetV2M和 VGG19都无法正确识 别测试集中的任何棉-聚酯样本,突显了这种混合物是一个 显著的挑战。VGG 模型在大多数类别上普遍表现出较差 的 F1-分数。

EfficientNetB0 模型的混淆矩阵如图 1 所示,并以可视 化方式总结了最佳模型在测试集上的分类性能。它证实了 对于棉质样本有出色的结果,所有样本均被正确分类;对 于粘胶-聚酯也有 13/14 样本被正确分类,而且后者展示 了完美的精度,因为没有其他类别被错误分类为粘胶-聚 酯。主要难点是棉-聚酯混合物,其中只有 6 个实例中的 2 个被正确识别,2 个被错误分类为棉,2 个被错误分类为 聚酯。还出现了一个小误差,即1 个聚酯实例被错误分类 为棉-聚酯。这突出了在数据集中可靠地区分棉-聚酯混合 物与纯棉和聚酯的挑战。

A. 分类模型训练分析

为了提供对在测试集评估中表现最佳模型学习过程的深入了解,图2和3分别展示了EfficientNetB0折叠3检查 点的训练和验证准确率以及损失曲线。这些曲线展示了模型在特定交叉验证折叠中的两阶段训练过程中的行为。图



Fig. 1. EfficientNetB0 (折叠 3) 的混淆矩阵



Fig. 2. EfficientNetB0(第3折)的训练和验证准确率



Fig. 3. EfficientNetB0 (第3折) 的训练和验证损失

 Original Image
 With Masks
 With Masks & Boxes

Fig. 4. 示例分割结果

TABLE VI 特征分割流程的性能指标。

Metric	Bounding Box Prediction	Mask Prediction
Precision	1.00	1.00
Recall	1.00	1.00
F1-Score	1.00	1.00
Mean IoU	0.92	0.90

形显示模型在训练过程中的成功收敛,表现为准确率的上 升趋势和损失的下降趋势。尽管在训练和验证性能之间表 现出预期的差距(针对这个数据集的大小),验证损失趋于 稳定,这表明通过早停进行了适当的正则化。

虽然第 3 折的检查点在最终测试集上表现最佳,但值得 注意的是,第 4 折在交叉验证过程中达到了更高的验证准 确率峰值。这说明了在训练中观察到的验证性能与在单独 的测试集上最终泛化性能之间的预期差异,尤其是在我们 的小数据集中。

B. 特征分割性能

特征分割流程利用 Grounding DINO 进行文本提示检测和 SAM 生成掩码,在数据集中识别按钮和拉链的任务 上进行了评估。定量评估重点关注生成的边界框和分割掩 码相对于人工注释的质量。表格 VI 列出了结果。

结果表明,该方法在收集的数据集上表现非常有效。衡量预测区域与真实区域重叠程度的平均交并比(Mean Intersection over Union)被计算出来。对于最终的分割掩码,该流程实现了0.90的平均IoU。中间的边界框预测也显示出强大的定位能力,平均IoU为0.92。此外,对象级检测指标,包括精度、召回率和F1分数,无论是掩码还是边界框预测,都被报告为1.00,这表明基本上所有目标特征均被正确检测到且没有假阳性。

图 4 展示了说明流程输出的定性示例。这些示例显示了 原始纺织品图像,以及为检测到的按钮和拉链生成的分割 掩码和边界框,直观地证实了该方法所达到的精确定位。

V. 结论

本文针对使用标准 RGB 图像对纺织材料进行自动化分 析以支持高效纺织品回收的需求进行了研究。我们开发并 评估了计算机视觉组件,以完成两个关键任务:分类常见 纺织材料(包括复杂的混纺)和分割需要移除的非纺织污 染物。在分类方面,我们研究了几种预训练架构,通过迁 移学习和自定义数据集的5折交叉验证进行分析,结果 表明 EfficientNetB0 在未见过的测试集上表现最佳,达到 了 81.25%的准确率。对于特征分割,采用了一种结合 Grounding DINO 开放词汇检测器和 Segment Anything 模型的零次方法,生成的掩膜达到了 0.9 的平均 IoU,效果非常出色。

这项工作的贡献在于展示了在纺织品回收领域中,使用 易于获取的 RGB 摄像机数据和现代深度学习技术进行特 定实际任务的可行性。我们为包括混合材料在内的常见材 料分类建立了性能基准,并成功应用了一个最先进的基础 模型管道,在无需特定任务训练的情况下实现了精准的污 染物分割。这些开发的视觉能力代表了实现关键预处理步 骤自动化的必要技术,例如材料分类和污染物去除,从而 为更具规模性和成本效益的自动化纺织品回收系统铺平了 道路。

VI.

致谢 该工作由英国政府的 Innovate UK Smart Grants 资助的英国研究与创新 (UKRI) 资助 [拨款编号 10111591 (ReFibres)]。

References

- I. Wojnowska-Baryła, K. Bernat, M. Zaborowska, and D. Kulikowska, "The growing problem of textile waste generation—the current state of textile waste management," *Energies*, vol. 17, no. 7, p. 1528, 2024.
- [2] W. Leal Filho, M. A. P. Dinis, O. Liakh, A. Paço, K. Dennis, F. Shollo, and H. Sidsaph, "Reducing the carbon footprint of the textile sector: an overview of impacts and solutions," *Textile Research Journal*, vol. 94, no. 15-16, pp. 1798–1814, 2024.
- [3] C. Alpert, M. Turkowski, and T. Tasneem, "Scalability solutions for automated textile sorting: a case study on how dynamic capabilities can overcome scalability challenges," 2021.
- [4] R. Ben Amor, K. T. W. Ng, T. T. Sithi, and T. S. Mahmud, "Opportunities and challenges for the sorting of post-consumer textile waste," in *Canadian Society of Civil Engineering Annual Conference*. Springer, 2023, pp. 89–99.
- [5] J. Clark, G. Johnson, O. Duran, and V. Argyriou, "Fabric composition classification using hyper-spectral imaging," in 2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT), 2023, pp. 347–353.
- [6] T. Behrendt and E. Eppinger, "Contamination threshold values for textile recycling," in *ITC-ICEE*. Springer, 2023, pp. 468– 479.
- [7] M. D. Stanescu, "State of the art of post-consumer textile waste upcycling to reach the zero waste milestone," *Environmental Science and Pollution Research*, vol. 28, no. 12, pp. 14253– 14270, 2021.
- [8] J. Zhou, L. Yu, Q. Ding, and R. Wang, "Textile fiber identification using near-infrared spectroscopy and pattern recognition," *Autex Research Journal*, vol. 19, no. 2, pp. 201–209, 2019.
- [9] P. Peets, I. Leito, J. Pelt, and S. Vahur, "Identification and classification of textile fibres using atr-ft-ir spectroscopy with chemometric methods," *Spectrochimica Acta Part A: Molecular* and Biomolecular Spectroscopy, vol. 173, pp. 175–181, 2017.

- [10] X. Sun, M. Zhou, and Y. Sun, "Classification of textile fabrics by use of spectroscopy-based pattern recognition methods," *Spectroscopy Letters*, vol. 49, no. 2, pp. 96–102, 2016.
- [11] J.-R. Riba, R. Cantero, T. Canals, and R. Puig, "Circular economy of post-consumer textile waste: Classification through infrared spectroscopy," *Journal of Cleaner Production*, vol. 272, p. 123011, 2020.
- [12] W. Du, J. Zheng, W. Li, Z. Liu, H. Wang, and X. Han, "Efficient recognition and automatic sorting technology of waste textiles based on online near infrared spectroscopy and convolutional neural network," *Resources, Conservation and Recycling*, vol. 180, p. 106157, 2022.
- [13] J.-R. Riba, R. Cantero, P. Riba-Mosoll, and R. Puig, "Postconsumer textile waste classification through near-infrared spectroscopy, using an advanced deep learning approach," *Polymers*, vol. 14, no. 12, p. 2475, 2022.
- [14] V. Argyriou, M. Petrou, and S. Barsky, "Photometric stereo with an arbitrary number of illuminants," *Computer Vision and Image Understanding*, vol. 114, no. 8, pp. 887–900, 2010.
- [15] N. Haleem, M. Bustreo, and A. Del Bue, "A computer vision based online quality control system for textile yarns," *Comput*ers in Industry, vol. 133, p. 103550, 2021.
- [16] S. Dlamini, C.-Y. Kao, S.-L. Su, and C.-F. Jeffrey Kuo, "Development of a real-time machine vision system for functional textile fabric defect detection using a deep yolov4 model," *Textile Research Journal*, vol. 92, no. 5-6, pp. 675–690, 2022.
- [17] V. Li, B. Villarini, J.-C. Nebel, and A. Vasileios, "A modular deep learning framework for scene understanding in augmented reality applications," in 2023 IEEE IAICT. IEEE, 2023, pp. 45–51.
- [18] V. Li, G. Tsoumplekas, I. Siniosoglou, V. Argyriou, A. Lytos, E. Fountoukidis, and P. Sarigiannidis, "A closer look at data augmentation strategies for finetuning-based low/few-shot object detection," in 2024 IEEE 14th International Symposium on Industrial Embedded Systems (SIES). IEEE, 2024, pp. 156– 163.
- [19] V. Li, I. Siniosoglou, T. Karamitsou, A. Lytos, I. D. Moscholios, S. K. Goudos, J. S. Banerjee, P. Sarigiannidis, and V. Argyriou, "Enhancing 3d object detection in autonomous vehicles based on synthetic virtual environment analysis," *IVC*, vol. 154, p. 105385, 2025.
- [20] A. C. da Silva BarrosM, E. Firmeza Ohata, S. P. P. da Silva, J. Silva Almeida, and P. P. Rebouças Filho, "An innovative approach of textile fabrics identification from mobile images using computer vision based on deep transfer learning," in 2020 International Joint Conference on Neural Networks (IJCNN), 2020, pp. 1–8.
- [21] J. Zhou, X. Zou, and W. K. Wong, "Computer vision-based color sorting for waste textile recycling," *International Journal* of *Clothing Science and Technology*, vol. 34, no. 1, pp. 29–40, 2022.
- [22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015. [Online]. Available: https://arxiv.org/abs/1409.1556
 [23] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling
- [23] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," 2020. [Online]. Available: https://arxiv.org/abs/1905.11946
- [24] T. Ridnik, E. Ben-Baruch, A. Noy, and L. Zelnik-Manor, "Imagenet-21k pretraining for the masses," arXiv preprint arXiv:2104.10972, 2021.
- [25] J. Gupta, S. Pathak, and G. Kumar, "Deep learning (cnn) and transfer learning: a review," in *Journal of Physics: Conference Series*, vol. 2273, no. 1. IOP Publishing, 2022, p. 012029.
- [26] C. Li, J. Li, Y. Li, L. He, X. Fu, and J. Chen, "Fabric defect detection in textile manufacturing: a survey of the state of the art," *Security and Communication Networks*, vol. 2021, no. 1, p. 9948808, 2021.
- [27] P. Bergmann, K. Batzner, M. Fauser, D. Sattlegger, and C. Steger, "The mvtec anomaly detection dataset: a comprehensive real-world dataset for unsupervised anomaly detection," *International Journal of Computer Vision*, vol. 129, no. 4, pp. 1038– 1059, 2021.
- [28] S. Liu, Z. Zeng, T. Ren, F. Li, H. Zhang, J. Yang, Q. Jiang, C. Li, J. Yang, H. Su *et al.*, "Grounding dino: Marrying dino

with grounded pre-training for open-set object detection," in *European Conference on Computer Vision*. Springer, 2024, pp. 38–55.

[29] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo et al., "Segment anything," in *Proceedings of the IEEE/CVF* international conference on computer vision, 2023, pp. 4015– 4026.