

基于多型卷积混合的高光谱非烟物质检测方法

黄振军^{1*}, 陈晋², 谭格¹, 符再德³, 刘波兰⁴, 陈实^{5#}, 许文武⁶, 刘承钧⁶

¹湖南中烟工业有限责任公司, 湖南 长沙

²南京理工大学计算机科学与工程学院, 江苏 南京

³湘西鹤盛原烟发展有限责任公司, 湖南 长沙

⁴浏阳天福打叶复烤有限责任公司, 湖南 长沙

⁵贵州中烟工业有限责任公司毕节卷烟厂, 贵州 毕节

⁶南京焦耳科技有限责任公司, 江苏 南京

收稿日期: 2023年12月15日; 录用日期: 2024年1月17日; 发布日期: 2024年1月24日

摘要

有效减少烟叶中的非烟物质, 特别是减少在加工过程不慎混入的次生非烟物质一直是烟草工业中的挑战性问题。本文提出了一种基于MCFnet (Multiple Convolution Fusion U-NET)模型的高光谱非烟物质检测方法。它通过在U-NET模型的下采样段、中间段和上采样段分别使用中心差分卷积残差模块, 快速傅立叶卷积残差模块和常规卷积层的特定混合, 实现高质量的非烟杂物分割结果, 完成非烟物质的定位与类型识别。模型的训练使用加权交叉熵和Dice损失函数的组合以保证模型经充分训练后能得到较好的性能。测试结果表明: 对面积不小于10 mm × 10 mm的8类典型非烟杂物的检测, 虚警率为1.12%, 漏检率为0.45%, 对不考虑杂物类别, 即把所有非烟杂物作为一类, 且包含了部分训练集中未出现的杂物类别的检测, 虚警率为3%, 漏检率为2.4%。与现有非烟杂物检测方法比较, 性能有显著提升。此外, 已在湖南中烟下属复烤厂完成了系统的实际部署与模型性能验证。

关键词

高光谱, 非烟杂物检测, 多型卷积混合

Multi-Convolution Fusion Based Hyperspectral Non-Tobacco-Leaf Materials Detection Method

Zhengjun Huang^{1*}, Jin Chen², Ge Tan¹, Zaide Fu³, Bolan Liu⁴, Shi Chen^{5#}, Wenwu Xu⁶, Chenjun Liu⁶

*第一作者。

#通讯作者。

文章引用: 黄振军, 陈晋, 谭格, 符再德, 刘波兰, 陈实, 许文武, 刘承钧. 基于多型卷积混合的高光谱非烟物质检测方法[J]. 图像与信号处理, 2024, 13(1): 76-91. DOI: 10.12677/jisp.2024.131008

¹Chinese Tobacco Industry Hunan Co. Ltd., Changsha Hunan

²School of Computer Science and Engineering, Nanjing University of Science & Technology, Nanjing Jiangsu

³Xiangxi Henshen Raw Tobacco Development Co. Ltd., Changsha Hunan

⁴Liuyang Tianfu Tobacco Leaf Redrying Co. Ltd., Changsha Hunan

⁵Bijie Tobacco Factory, Chinese Tobacco Industry Guizhou Co. Ltd., Bijie Guizhou

⁶Nanjing Jiao'er Technology Co. Ltd., Nanjing Jiangsu

Received: Dec. 15th, 2023; accepted: Jan. 17th, 2024; published: Jan. 24th, 2024

Abstract

A challenging problem in tobacco industry is to efficiently decrease non-tobacco-leaf materials and especially the secondary produced ones during tobacco production. In this work, we present a non-tobacco-leaf materials detection method by using hyper-spectral images and the proposed Multiple Convolution Fusion U-NET. The method integrates center differential convolution residue modual, fast Fourier transformation convolution residue modual and standard convolution into the U-NET model and respectively deploys each of them to the downsampling stage, center stage and upsampling stage of U-NET. To this end, the quality of impurity segmentation can be guaranteed, which further achieves accurate localiation and recogtion. We combine weighted cross entropy loss and Dice loss to train our model as we find it effectively and efficiently for sufficient training. The overall method achieved a false-alarm rate of 1.12% and a missed-detection rate of 0.45% for eight typical non-tobacco materials with the size less than 10 mm × 10 mm. When test in the case of not considering the types of non-tobacco materials, and a number of new non-tobacco materials being involved, the proposed method achieved 3% false-alarm rate and 2.4% missed-detection rate. A significant improvement is also found when compared with existing non-tobacco-leaf materials detection method. Moreover, system deployment and model performance validation have been finished in a tobacco factory of Chinese Tobacco Industry Hunan Co. Ltd.

Keywords

Hyperspectral, Non-Tobacco Material Detection, Multi-Convolution Fusion

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

烟叶在采收、运输和加工过程中，不可避免地会混入非烟物质。烟叶中的非烟物质会极大地影响烟草原料的使用价值和卷烟产品的质量。非烟杂质是指非烟叶以外的各种物质，一般分为两大类：第一类为原生杂物，即在采摘运输过程中混入的非烟物质，包括杂草、石头、昆虫、麻绳等；第二类为次生杂物，即在生产过程中混入的非烟物质，包括金属、皮带、塑料、纸片等[1]。非烟物质的混入不仅会影响烟叶脱粒、复烤等设备，燃烧产生的有害物质也会严重影响人体健康[2]。为了保证卷烟的吸食安全，烟草行业对烟叶中的杂质含量提出了严格的控制要求[3]。

在卷烟生产加工线上，非烟物质的去除通常采用风力、磁力等方法[4]，但风力、磁力去除方法都依赖自动的非烟物质检测算法，基于可见光摄像机的视觉非烟物质检测研究已受到较多的关注，很多视觉检测模型被提出[5][6]，但这些模型的识别性能不尽如人意，表现在虚警率和漏检率均较高。随着烟草行

业对除杂系统准确性、可靠性要求的提高。近年来,研究者发现可见光谱段有限是制约性能提高的一个重要因素。随着高光谱成像技术的飞速发展,他们开始逐步关注多谱段的引入。现阶段,高光谱相关成像器件的性能已有很大提高,比如,线性推扫成像的高光谱成像仪可水平获取光谱信息,显示光谱像素的位置,提供物体的空间分布信息,实现不同被测物体的实时可视化成像。此外,高光谱像机的谱段不断丰富,分辨率和采集速度也得到了保证。因此,高光谱成像仪在烟草领域越来越受到重视[7] [8] [9] [10],比如烟叶水分含量的鉴定[5]、烟草花叶病毒的鉴定[11]、烟叶成熟度的判断[12]等。但是,利用高光谱技术对烟草脱粒和复干过程中的烟叶和杂质进行识别和分类的研究较少,代表性的工作主要有文献[12]提出的一种基于随机森林的方法,该方法首先对高光谱图像进行预处理,然后对烟叶和杂质的感兴趣区域进行建模和分析。最后,利用随机森林算法对烟叶和杂质进行分类,实现了非烟物质的检测,但基于高光谱的非烟物质检测和识别模型的准确率仍有很大提升空间。

本文采用了基于 U-NET 深度网络模型来实现基于高光谱非烟物质检测,通过模型改进获得高质量的分割结果实现非烟物质的空间定位与类型识别。U-NET 模型被广泛运用到了各种目标分割和语义分割任务中,原因是其编码器结合解码器的两阶段架构可以很好把握图像的本质表征,而且编码器先逐次下采样、再由解码器渐进上采样的全卷积方式实现了多尺度的特征提取,特征提取的更全面。但是,很多工作在使用 U-NET 对高光谱数据进行处理时,也集中于调整编码器结合解码器架构,以及对下采样、上采样方式做一些调整,却忽略了一个问题,即卷积方式的多样化,依然沿用标准卷积,多集中在调整卷积核的尺寸,多少和改进计算方式等。本文发现,使用 U-NET 处理高光谱数据时,给架构引入多种不同类型的卷积操作可以显著提升模型的性能。在这一观察指导下,我们在 U-NET 模型的下采样段,中间段和上采样段分别使用中心差分卷积,快速傅立叶卷积和标准卷积的特定混合,让模型更适合提取非烟物质区域,抑制正常烟叶区域的虚警。

既然模型的改进引入了不同类型的卷积去适配 U-NET 网络的不同阶段的特征提取,那么为了与改进模型相得益彰,训练改进模型的损失函数是否需要混合,结果发现,使用加权交叉熵和 Dice 损失的组合比直接使用交叉熵损失、加权交叉熵损失、Focal 损失, Dice 损失更能使模型得到充分稳定的训练,达到最好的性能。

最后,基于以上模型我们设计了一个高光谱非烟物质检测系统,并在湖南中烟下属打叶复烤厂完成了实际部署与验证。验证表明:该系统在片烟中检测和识别面尺寸不小于 $10\text{ mm} \times 10\text{ mm}$ 的次生非烟杂物检测的虚警率为 1.12%,漏检率为 0.45%,对不考虑杂物类别,即把所有非烟杂物作为一类,且包含了部分训练集中未出现的杂物类别的检测,虚警率为 3%,漏检率为 2.4%。较好的支持非烟物质的预警和剔除,有效减少复烤过程次生非烟物质进入片烟造成的质量隐患与风险。

2. 多型卷积融合网络模型构建

本节我们给出了所提出的 MCFnet 模型的架构图,并对架构图进行了简要介绍;分析了在 U-NET 模型基础上引入多类型卷积的目的,详细介绍了新引入的两类卷积模块,即中心差分卷积残差模块和快速傅里叶卷积残差模块。

2.1. 模型架构

图 1 给出了 MCFnet 模型的网络结构图, MCFnet 模型以 U-NET 作为基础骨架模型。U-NET 是一种常用于图像语义分割任务的深度学习模型。U-NET 的结构特点在于它的 U 形设计,具有编码器(下采样路径)和解码器(上采样路径)之间的对称结构,能够捕捉不同尺度的特征并保留高分辨率的信息。编码器(下采样路径)由多个卷积层和池化层构成,用于逐渐减小输入图像的尺寸和通道数。每次通过卷积操作,

都可以提取出图像的高层次特征。通过池化操作,可以将图像的尺寸降低,从而加速计算并扩大感受野。编码器的设计有助于捕捉图像的全局特征。解码器(上采样路径)部分也由多个卷积层和上采样(反卷积或转置卷积)层构成,用于将编码器得到的特征进行逐步恢复到原始图像尺寸。解码器的设计有助于精确定位分割目标的边界和细节。跳跃连接(Skip Connections)用于在特征传递时,将编码器的某一层的特征直接连接到解码器相应层,这样可以将低层次的细节特征传递到解码器,帮助精细化分割。

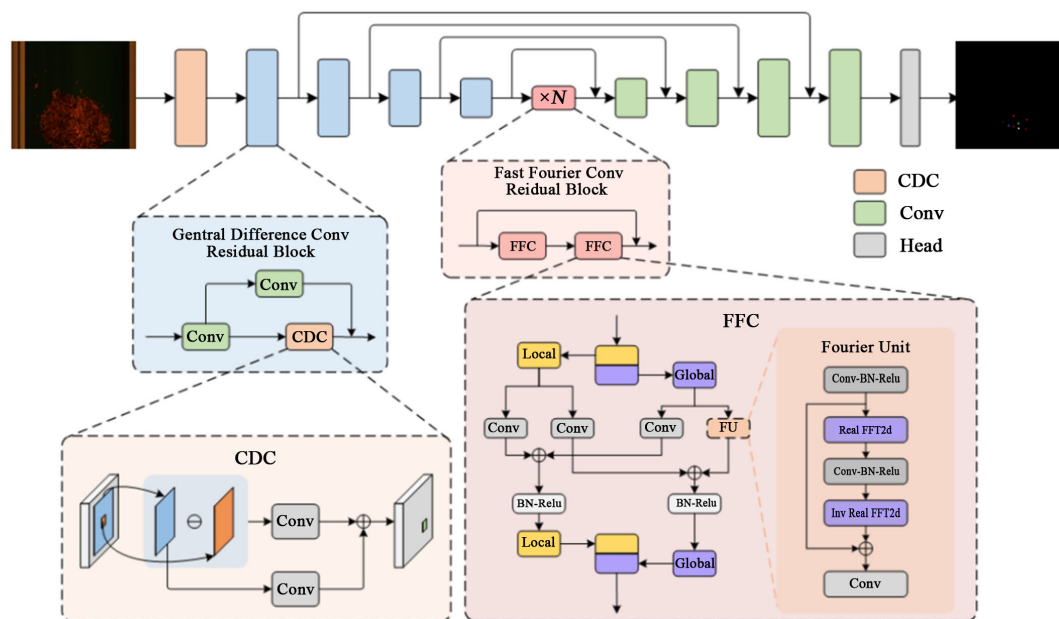


Figure 1. The overall architecture diagram of MCFnet model
图 1. MCFnet 模型完整架构图

本文将 U-NET 作为基础骨架模型的原因有三个: (1) U-NET 对高光谱特征的提取性能较好, (2) 考虑到本项目杂物目标都比较小, U-NET 被证明是一种适合检测弱小目标的骨架模型; (3) U-NET 是全卷积网络, 参数量小于其它主流骨架模型范式, 推理速度快。

我们针对高光谱数据的通道维度高的特性, 对 U-NET 网络结构在通道维度上进行了拓展, 使其更适合于高光谱数据的分割问题。其次, 引入多种不同类型的卷积操作来改进 U-NET 网络结构, 这种结构可以使模型对杂物鉴别特征的提取能力更强。首先使用一个中心差分卷积操作做特征提取, 让模型在浅层更关注提取非烟杂物有烟叶的差异特征。编码器部分采用 4 个中心差分卷积残差模块, 它是中心差分卷积与标准卷积的混合, 并结合了残差连接; 其次, 对得到的特征图采用傅立叶卷积残差模块, 它是快速傅立叶卷积与标准卷积的混合, 也结合了残差连接。解码器阶段使用 4 个标准卷积模块, 最后是语义分割输出头, 输出三维概率图, 其宽、高与输入图像尺寸一样, 通道数与杂物类别个数相同。

2.2. 多型卷积混合

通过改变卷积核大小实现通道适配, U-NET 网络模型可以直接处理高光谱图像, 实现非烟杂物的检测。但是, 由于高光谱空间尺寸不高, 而且部分非烟杂物本身就很小(如螺丝), 或者被烟叶部分遮挡, 导致高光谱图像中这部分杂物面积非常小, 成为较难解决的小目标检测问题, 必须特别考虑如何更好的突出杂物与烟叶差异, 否则会在模型处理过程中容易丢失; 同时成像系统在长时间运行后会出现一些负面因素(如光源均匀性, 噪声影响等), 导致部分杂物像素之间的光谱差异变大, 不利于非烟杂物的检测。如

图 2(b)显示了一些成像系统运行良好时采集的正常烟叶与非烟杂物的光谱曲线,此时,非烟物质与正常烟叶的曲线可分性较明显。

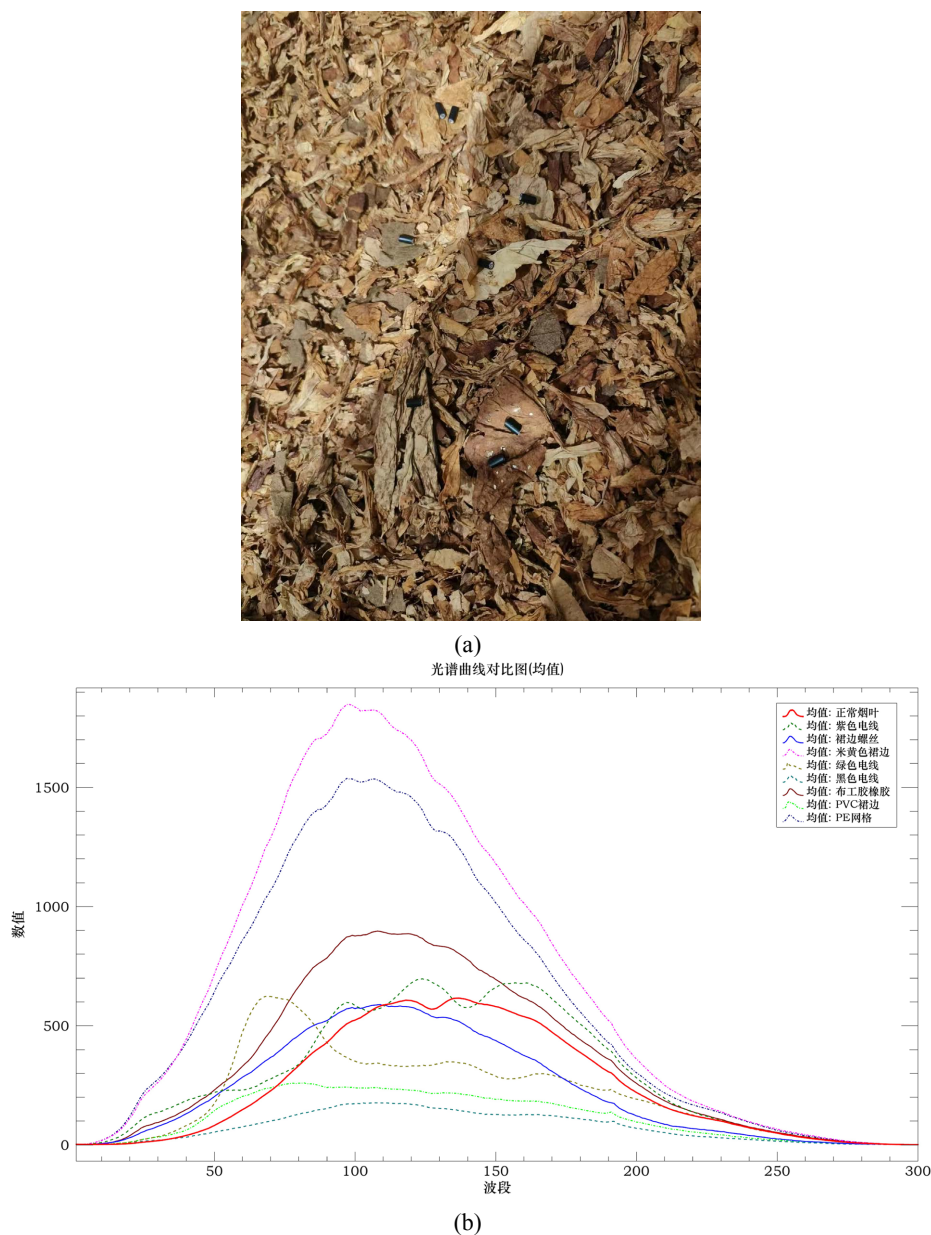


Figure 2. The real picture and spectral curves of normal tobacco leaves and typical non-tobacco materials (a) The real picture of normal tobacco leaves and typical non-tobacco materials; (b) The spectral curves of normal tobacco leaves and typical non-tobacco materials

图 2. 正常烟叶与典型非烟物质实物图与光谱曲线(a) 正常烟叶与典型非烟杂物实物图;(b) 成像系统运行良好时采集的正常烟叶与非烟杂物的光谱曲线

但是, 图 3(a)所示的黑色电线, 在采集系统长时间运行后, 观察拍摄出来的高光谱图像中的几个像素点的光谱曲线(图 3(b))就会发现, 这些像素的光谱曲线一致性变差。此外, 受到了较大的系统噪声的影响, 光谱峰值也在较大范围变动, 容易造成模型对黑色电线像素的漏检和虚警(如图 4)。因此, 在高光谱 U-NET 骨架基础上, 我们增加了中心差分卷积残差模块和快速傅里叶卷积残差模块, 以期更加突出非烟

杂物的同时，减少这些因素的影响。

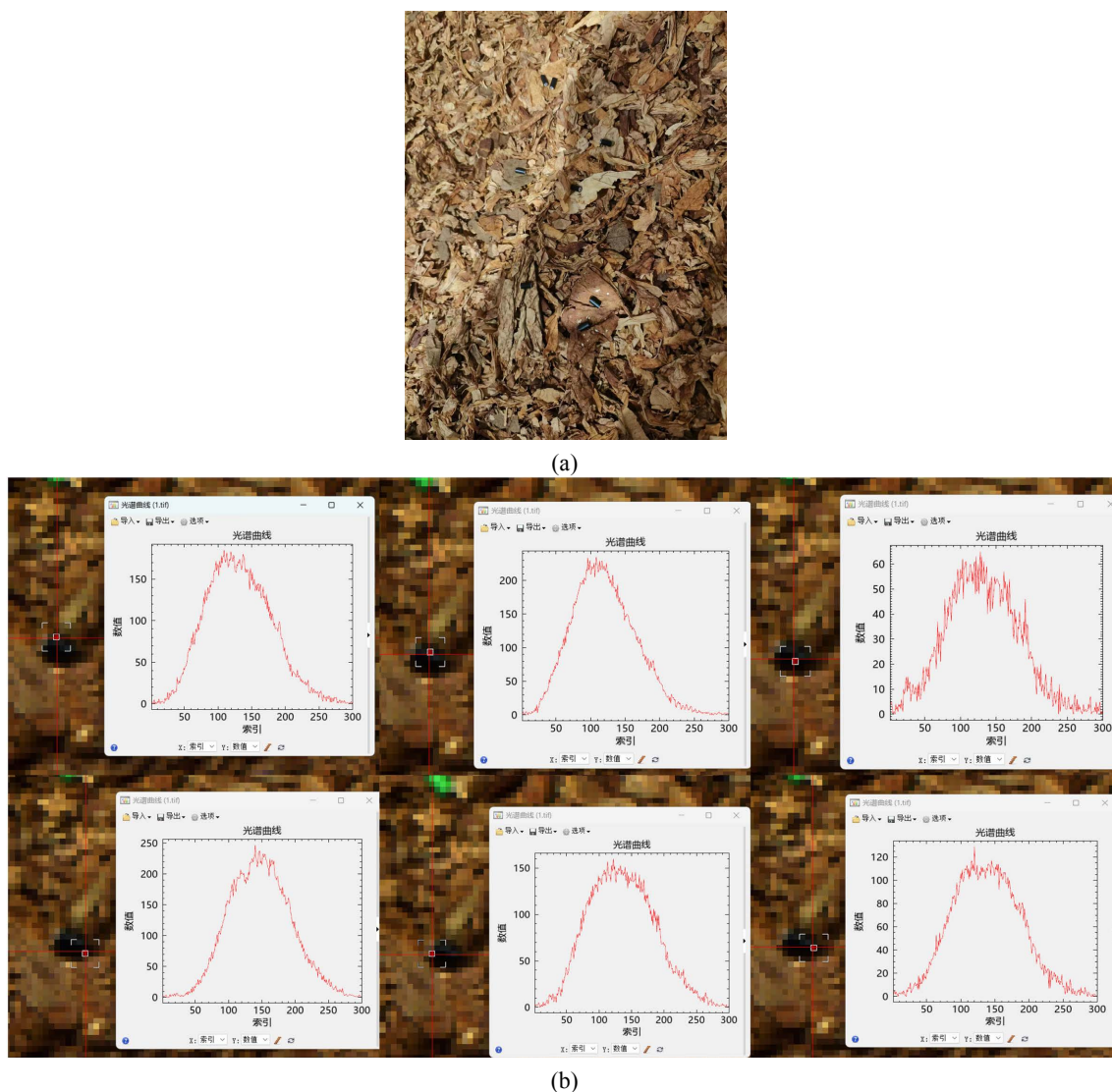


Figure 3. The discrimination between non-tobacco materials and normal tobacco with aspect to their spectral curves and the heterogeneity of some non-tobacco materials. (a) The real picture of black wires; (b) The significant difference of spectral curves at six pixels of black wires

图 3. 非烟物质与正常烟叶光谱曲线可分性与部分非烟物质光谱的不一致性, (a) 黑色电线杂物实物图; (b) 黑色电线上六个不同像素光谱曲线差异(光谱峰值在 60 到 250 范围变动, 噪声强度差异大)

2.2.1. 中心差分卷积残差模块

标准卷积[13]通过考虑邻域像素特征的加权组合可以有效地提取二维图像的颜色、形状、纹理等信息, 能够较好的实现图像的识别。但是对非烟杂物检测任务来说, 标准卷积的特征提取性能受限。由于绝大多数非烟杂质尺寸比较小, 其特征很容易在 U-NET 的下采样过程中被淹没, 导致无法检测到非烟杂质。为了避免过度下采样导致的漏检, 我们使用中心差分卷积来突显非烟物质与正常烟叶的差异信息。中心差分卷积先计算卷积核中心像素点和周围像素点的特征差异, 并在此基础上进行卷积, 从而引导网络模型学习差异特征, 由于非烟物质与正常烟叶特征差异大于正常烟叶之间的特征差异, 一定程度上提高了非烟物质的对比度, 中心差分卷积的公式描述如下:

$$f(x, y) = \sum_{(i,j)} w_{(i,j)} (F_{(x+i,y+i)} - F_{(x,y)}) \quad (1)$$

其中 x, y 为当前中心, R 为卷积窗口大小。为了保留标准卷积的特性, 一般将中心差分卷积与标准卷积进行加权融合, 描述如下:

$$CDC(x, y) = \theta \times \sum_{(i,j)} w_{(i,j)} (F_{(x+i,y+i)} - F_{(x,y)}) + (1-\theta) \sum_{(i,j)} w_{(i,j)} F_{(x+i,y+i)} \quad (2)$$

其中 $\theta \in (0,1)$ 为中心差分卷积和标准卷积的比例因子, 其值越大, 则中心差分卷积所占比例越高。用于计算差异的窗口大小 R 一般设置为卷积核大小, 中心差分卷积模块的感受野会随着网络前向传播获得多层次的对比度信息, 帮助网络识别不同大小的非烟杂物。

尽管中心差分卷积层补充了多尺度的对比度特征, 但学习上比标准卷积要复杂, 为了易于学习, 最大程度发挥其性能, 我们在中心差分卷积基础上加入的残差模块, 设计了中心差分卷积残差模块。如图 1 中的蓝色方块所示, 它包含一个残差连接, 在中心差分卷积前和残差连接上各引入一个标准卷积层。

2.2.2. 快速傅里叶变换卷积残差模块

基于高光谱图像的非烟物质检测的最显著优势在于采集较多精细的光谱通道, 从而提供了更多辨识非烟物质的有效信息, 但这些有效信息的提取对模型的信息挖掘或特征提取能力提出了更高的要求。尽管我们在 U-NET 模型基础上使用中心差分卷积, 但整个模型与普通的只有三个通道的可见光图像相比, 并没有特别考虑空-谱两维度上的特征提取。为此, 我们采用了快速傅里叶卷积弥补这一不足, 用其把经过编码器输出的特征图分成局部特征提取和全局特征提取两个分支, 在局部特征提取分支中使用标准卷积进行空-谱通维度的细节提取, 在全局特征提取分支中, 使用傅里叶单元实现空-谱维度上的全局特征提取, 并在两个分支间, 使用捷径链接进行信息交流。

进一步的, 在傅里叶单元中我们先将输入特征通过快速傅里叶变换从空间域转换到频域, 由于频域上的每个点总是跟空间域的所有点相关的, 在频域上只需采用小尺寸核进行卷积操作就能影响原空-谱域上的所有点, 因此我们可以获得跟特征图尺寸同大小的感受野, 最后再通过傅里叶逆变换得到包含全局信息的特征图。傅里叶单元的流程如下:

- (1) 使用二维快速傅里叶变换将输入特征图从空间域转换到频域:

$$R_1^{(H \times W \times C)} \rightarrow C_1^{(H \times \frac{W}{2} \times C)} \quad (3)$$

- (2) 并在通道维度上拼接复数的实数部分和虚数部分, 当做两个通道的实数来处理:

$$C_1^{(H \times \frac{W}{2} \times C)} \rightarrow R_2^{(H \times \frac{W}{2} \times 2C)} \quad (4)$$

- (3) 在频域上应用卷积、归一化和激活函数操作:

$$R_2^{(H \times \frac{W}{2} \times 2C)} \rightarrow R_3^{(H \times \frac{W}{2} \times 2C)} \quad (5)$$

- (4) 将两个实数通道组合成一个通道的复数逆变换到空间域:

$$R_3^{(H \times \frac{W}{2} \times 2C)} \rightarrow C_2^{(H \times \frac{W}{2} \times C)} \quad (6)$$

- (5) 快速傅立叶逆变换:

$$C_2^{(H \times \frac{W}{2} \times C)} \rightarrow R_4^{(H \times W \times C)} \quad (7)$$

最后, 在通道维度上将局部分支与全局分支进行拼接, 得到包含局部信息与全局信息的特征图。

同样，快速傅立叶卷积引入增加了模型的学习负担，为此，我们对快速傅立叶卷积进一步引入残差连接，转变成快速傅立叶残差模块，这样不仅可以易于模型的学习，实验发现还能进一步提升模型的性能。此外，实验还发现对整个编码器处理完的高光谱特征进行快速傅立叶残差卷积层进行处理最有效，因此，我们把快速傅立叶残差卷积模块放在整个骨架的中间阶段。

2.3. 损失函数

我们首先使用交叉熵损失函数用于模型训练，根据模型的任务用到了二分类和多分类。

(1) 二分类

在二分类的情况下，模型最后需要预测的结果只有两种情况，即正常烟叶和非烟物质。对于每个类别我们的预测得到的概率为 p 和 $1-p$ ，此时表达式为：(log 以 e 为底)：

$$L_{CE} = \frac{1}{N} \sum_i L_i = \frac{1}{N} \sum_i -[y_i \log(p_i) + (1-y_i) \log(1-p_i)] \quad (8)$$

其中： y_i 表示样本 i 的类别标签，正常烟叶为 1，非烟物质为 0， p_i 表示样本 i 预测为正常烟叶概率。

(2) 多分类

多分类的情况就是对二分类的扩展，公式为：

$$L_{CE} = \frac{1}{N} \sum_i L_i = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (9)$$

其中， M 为所有非烟物质类别加正常烟叶类的总类别数(实验中为 9 类)， y_{ic} 为符号函数，它表示如果样本 i 的真实类别等于类别 c 取 1，否则取 0， p_{ic} 为观测样本 i 属于类别 c 的预测概率。

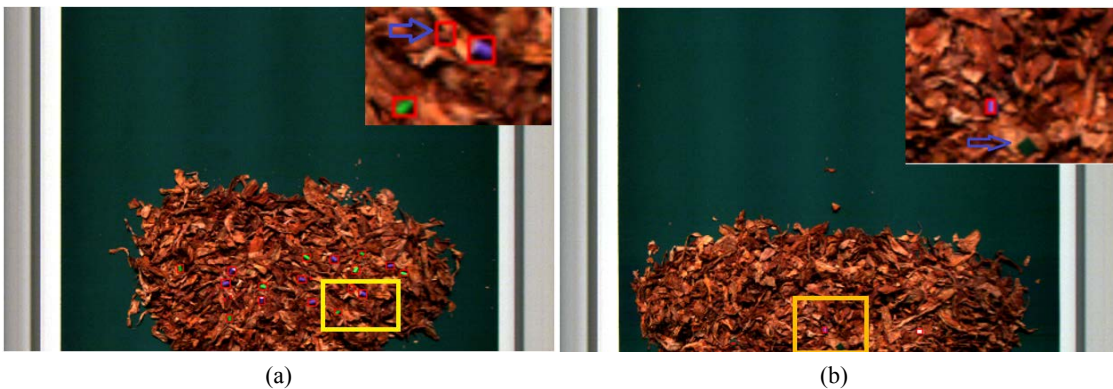


Figure 4. Typical cases of false alarms and missed detection (a) A false alarm case; (b) A missed detection case
图 4. 典型虚警和漏检实例图(a) 虚警实例; (b) 漏检实例

由于非烟杂物和背景(烟叶和传送带)空间占比悬殊，直接使用二分类或多分类交叉熵损失，会造成模型关注空间占比大的背景像素的错分损失，而忽略了非烟杂物的错分损失，导致非烟杂物漏检(如图 4(b))。为此，将交叉熵损失公式中的求和项按像素类别分成非烟杂物和背景两部分，对背景像素的损失求和后的损失值赋予一个较低的权重，而对非烟杂物像素求和后的损失值赋予一个较大的权重，这称为加权交叉熵损失 L_{WCE} 。此外，我们尝试了常用的 Focal 损失 L_{Focal} 和 Dice 损失 L_{Dice} 两种常用于样本比例不均衡的损失函数，并观察了这些损失与加权交叉熵损失联合形成组合损失的训练效果。实验发加权交叉熵损失与 Dice 损失等比例加权效果最好，即 $L = L_{WCE} + L_{Dice}$ ，而单独使用这些损失函数中一种或其他不同的组合都达不到上述组合的效果。这里 Dice 损失定义为：

$$L_{Dice} = 1 - \frac{2TP}{2TP+FP+FN} \quad (10)$$

其中 TP 表示模型预测为正例，实际是正例；FP 表示模型预测为正例，实际是反例；FN 表示模型预测为反例，实际是正例；TN 表示模型预测为反例，实际是反例。 L_{Dice} 取值范围在 0 到 1 之间。

3. 完整检测方法

在上一节模型构建和训练方法基础上，进一步设计合理的预处理和后处理步骤以实现完整可靠的杂物检测与识别方法。

(1) 黑白校正

在系统运行过程中，高光谱成像系统容易受到外界影响，如外界光照不均匀、相机镜头暗电流等。因此，为了获得更准确、更稳定的高光谱图像数据，需要在实验前对高光谱图像进行黑白校正。校正公式为：

$$R = \frac{I - B}{W - B} \quad (11)$$

其中， R 是校正后的高光谱图像； I 是原始高光谱图像； B 是反射率为 0 的黑色校准图像； W 是反射率为 99% 的白色校准图像，它由图 5 所示的反射率为 99% 白板获得。



Figure 5. The real picture of the used white board with 99% reflection
图 5. 反射率为 99% 白板实物图

(2) 检测与后处理

利用训练数据集和前面设计的加权损失函数充分训练 MCFnet 模型，通过验证集对模型的超参数进行优化，最终确定的模型用于对测试图像分割并识别出非烟物质，用于性能评估。杂物区域后处理。对于 MCFnet 模型输出图，按非烟杂物的类别做连通域分析，再将所有类别的连通区域收集到一起。这是为了避免两种不同类的非烟杂物靠在一起而误认为是一个非烟杂物的错误。将每个连通域的面积与预先设定的面积阈值比较(测试中设置为 2，相当于 10 mm × 10 mm 实物面积)，滤除面积小于或等于该面积阈值的噪声区域，剩余区域全部作为非烟物质。

4. 实验与分析

本节我们先给出所设计的非烟杂物检测系统的实物部署图和软件界面；介绍了实验数据的构成、模型评价指标以及方法的实现细节；进一步给出了本文方法与当前非烟杂物方法的综合测试与比较实验，最后给出了 MCFnet 模型的消融实验。

4.1. 系统介绍

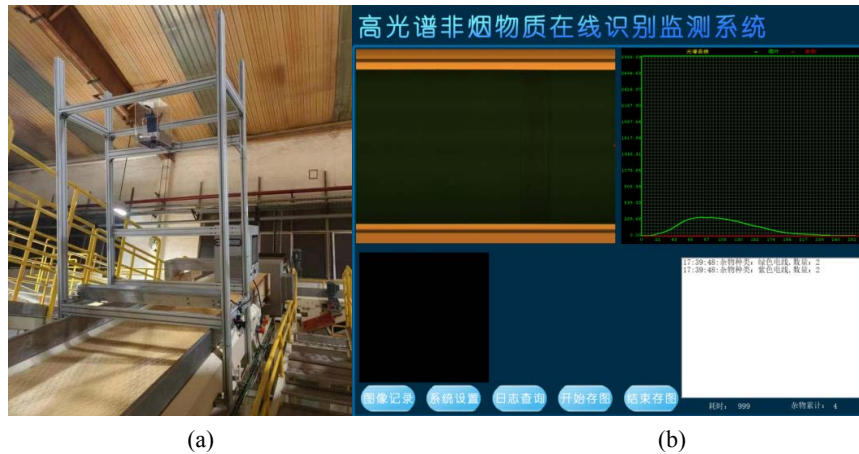


Figure 6. System deployment and software interface (a) System deployment; (b) Software interface
图 6. 系统部署与软件界面(a) 系统部署图; (b) 软件检测界面

如图 6(a)在烟叶传送带上方 1.5 米处架设了高光谱摄像机,与皮带垂直,保证在宽度方向上覆盖整个皮带上的烟叶区域。其中,高光谱摄像机为 LISEN 公司的 iSpecHyper-VS200 系列高光谱相机,光谱范围 400~1000 nm,分辨率优于 2 nm,点列斑直径 < 0.5 像元。系统光源采用覆盖 400 nm~2500 nm 全光谱的卤素灯光源,光学镜头焦距采用 12.5 mm。采用 VS2015 开发了基于深度学习的非烟物质图像后处理系统。使用反射率数据作为采集数据格式,因为反射率主要与材料特性有关,而当光场发生变化,同一材料的反射率保持不变。

4.2. 实验数据

实验数据集共包含 600 幅高光谱烟叶图像。图像空间尺寸为 300 * 480,每个像素有 300 个光谱维度。该数据集包含了电线、捆扎带、工布胶皮、PE、麻绳、PVC、胶片、螺丝共 8 类不同颜色和材质的烟草生产环境中的常见杂物。针对该数据集,我们提供了 tif 格式的高光谱图像文件, png 格式和 xml 格式的带类别标注信息,可用于多分类的杂物检测。其中 500 幅用于模型训练(训练集),100 幅用做参数验证(验证集),测试集是在复烤厂现场部署完系统后,共测试了半个小时,投放包含八个类的 447 个杂物和 53 个不属于训练集类别的杂物,主要包括烟盒、碎布、皮带,附件中给出了测试过程的部分视频。

4.3. 评价指标

模型评估主要采用平均交并比、虚警率和漏检率作为性能评价指标,平均交并比用于模型性能评价,而虚警率和漏检率用于检测与识别性能评价。

在语义分割的问题中,交并比就是该类的真实标签和预测值的交和并的比值。其中: TP、FP、FN 和 TN 的含义与 2.3 节 L_{Dice} 中的定义一致。

$$IoU = \frac{TP}{TP + FN + FP} \quad (12)$$

$mIoU$ 就是该数据集中的每一个类的交并比的平均,计算公式如下,其中 P_{ij} 表示将 i 类别预测为 j 类别。

$$mIoU = \frac{1}{k+1} \frac{\sum_{i=0}^k p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (13)$$

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP_i}{TP_i + FN_i + FP_i} \quad (14)$$

此外，也使用虚警率(FA)和漏检率(MD)指标来衡量分割性能，他们的定义如下。

$$FA = \frac{\sum_{i=0}^k FP_i}{\sum_{i=0}^k (FN_i + TP)} \quad (15)$$

$$MD = \frac{\sum_{i=0}^k FN_i}{\sum_{i=0}^k (FN_i + TP)} \quad (16)$$

4.4. 实现细节

进一步使用水平/垂直翻转对训练数据进行数据增强。使用 SGD 优化器，学习率初始化为 $5e-2$ ，并以余弦退火方式进行衰减，且衰减值 weight-decay 值设置为 $1e-4$ 。模型的 batch-size 设置为 8，执行 400 轮(epochs)可完成充分训练。

4.5. 综合测试与比较

我们使用所设计的模型组成高光谱杂物检测系统进行现场检测。我们选用随机森林方法，基于 YOLOv5 实例分割模型[14] [15] [16] [17]与本文方法作比较，8 类非烟杂物共投放 447 个，每类的投测试个数及检测结果如表 1 所示，可以看到，本文方法 FCMne 的检测虚警率和漏检率分别为 1.12%和 0.45%，三者中最低，随机森林和 YOLOv5 的综合性能不如 FCMnet，其中，随机森林虽具有较强的单个像素光谱特征表示和学习能力，但像素之间在图像空间上的语义关系建模并不强，导致较高的虚警率(5.82%)和漏检率(4.47%)。相比之下，YOLOv5 不会出现较多虚警(1.79%)，但受其模型本身对小目标的检测能力制约，特别是只有几个像素的非烟杂物的检测，容易出现漏检(漏检率达到 3.80%)。FCMnet 在抑制虚警和漏检两方面较为平衡，两种错误发生的更少些。

Table 1. Detection results of eight non-tobacco materials

表 1. 八类非烟物质检测结果

比较方法	类别	电线	捆扎带	工布胶皮	PE	麻绳	PVC	胶片	螺丝	总计
	测试数	93	57	50	58	54	61	39	35	447
随机森林 [12]	虚警数	7	1	1	4	3	4	4	2	26
	虚警率	7.53%	1.75%	2.00%	6.78%	5.56%	6.45%	10.26%	5.71%	5.82%
	漏检数	2	1	2	2	5	2	5	1	20
	漏检率	2.15%	1.75%	4.00%	3.39%	9.26%	3.23%	12.82%	2.86%	4.47%
YOLOv5 [14]	虚警数	1	0	0	2	3	1	0	1	8
	虚警率	1.08%	0	0	3.40%	5.56%	1.61%	0	2.86%	1.79%
	漏检数	3	3	2	2	1	2	3	1	17
	漏检率	3.23%	5.26%	4.00%	3.40%	1.85%	3.23%	7.69%	2.86%	3.80%
MFCnet	虚警数	1	0	0	0	3	0	0	1	5
	虚警率	1.03%	0	0	0	5.56%	0%	0%	2.86%	1.12%
	漏检数	0	0	0	1	0	0	0	1	2
	漏检率	0	0	0	1.69%	0	0	0	2.86%	0.45%

此外,模型的混淆矩阵如图7所示,其中,0表示背景(正常烟叶和传送带)。从表中可以看出,除第8类(螺丝)易被识别成背景或其它类外,其余类的识别性能均达到了较高的水平,证明MFCnet具有较强的识别能力。

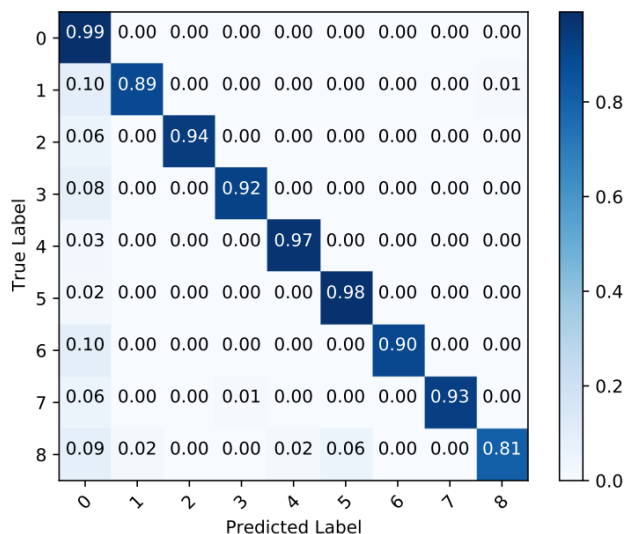


Figure 7. Consusion matrix
图7. 混淆矩阵

在实际测试中,我们还验证了将所有不同类非烟物质合并,统一看成一个类,即异常类,与正常烟叶进行区分,并且在前面实验采集的447个8种杂物样本基础上,还混入了53个不在8类中的非烟杂物,包括烟盒、碎布、皮带三类,于是共得到500个异常类样本。本测试希望以二分类方式来检验模型对“只需识别是否是杂物”这一任务下的检测能力以及对未见过的非烟杂物的检测能力。模型输出头改成二分类,训练样本将所有杂物归为一个类,具体测试方法与多类测试类似,其结果由表2给出。从已有的8类杂物图像上实验结果看(表2第三列),将8类杂物合为一个类别检测后,三种方法的检测虚警率和漏检率均增加了一些,总体不大,这主要是异常类的类内方差较大,给决策边界带来一些困难。MFCnet的虚警率和漏检率依然最低,分别维持在1.79%和1.57%。

对新引入的未见过的杂物进行检测时(表2第四列),三种方法的虚警率和漏检率都有不同程度的增大。YOLOv5的两种错误率增大的最明显,其中漏检率上升到了64.15%,主要原因是漏检的类似样本并没有出现在训练样本中,说明YOLOv5对训练样本的依赖较大;随机森林方法在检测新杂物时,受到的影响没有YOLOv5那么大,但其虚警率仍然较高,达到24.53%;本文方法MFCnet受到的影响最小,分别为13.21%和9.43%。在500幅图像上的综合来看(表2最后一列),MFCnet虚警率和漏检率最低,分别维持在3%和2.4%。

Table 2. Comparison of detection performance with all non-tobacco materials merged to an anomaly class and with instances from unseen class mixed

表2. 将所有杂物当成一个异常类并混入新类别杂物后的测试结果对比

比较方法	类别	8类杂物合成一个类	新增杂物(不属于8类)	总计
	测试个数	447	53	500
随机森林[12]	虚警数	31	13	44
	(虚检率)	6.94%	24.53%	8.80%
	漏检数	25	6	31
	(漏检率)	5.59%	11.32%	6.20%

续表

YOLOv5 [14]	虚警数	13	15	28
	(虚警率)	2.91%	28.30%	5.60%
	漏检数	26	34	60
	(漏检率)	5.82%	64.15%	12.00%
MFCnet	虚警数	8	7	15
	(虚警率)	1.79%	13.21%	3.00%
	漏检数	7	5	12
	(漏检率)	1.57%	9.43%	2.40%

图 8 给出了一组展示各方法检测结果的可视化图，从图中可以清楚的看到，对比非烟杂物的掩模真值，随机森林方法虚警较多，YoloV5 方法漏检较多，且也存在好几个虚检，而本文提出的 MCFnet 虚警和漏检均较低。

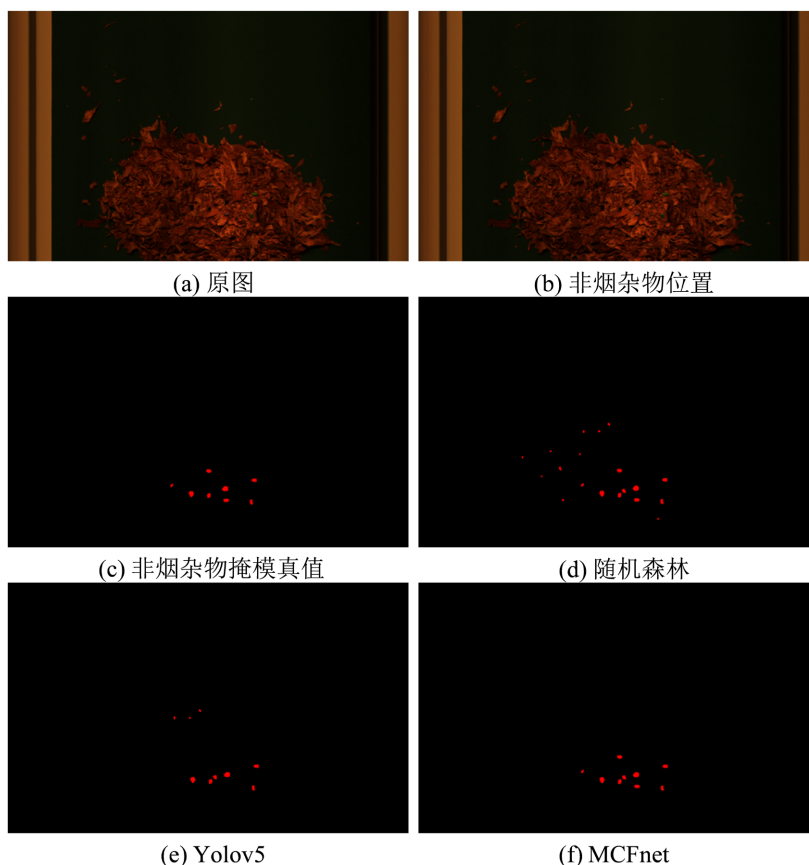


Figure 8. Visualization results for all methods

图 8. 所有方法非烟物质检测可视化结果图

4.6. 消融实验

(1) 不同模型骨架比较

我们分别使用 U-NET 及 DeeplabV3 模型架构中的不同骨架作为我们的模型骨架，并比较它们在我们的非烟杂物数据集上检测 mIoU 值，由表 3 可以发现，使用 U-NET 作为 MCFnet 的模型骨架，mIoU 性能最高。其中，BaseConv 就是图 1 给出的架构图，即降采样五层卷积，下采样四层卷积。

Table 3. Ablation study of the selection of model architecture and Backbone**表 3.** 模型架构及骨架选择消融实验

模型架构	骨架	mIoU
Deeplabv3	Resnet50	46.4
Deeplabv3	BaseConv	47.3
U-NET	Resnet50	69.2
U-NET	BaseConv	69.7

(2) 引入多类型卷积比较

在 U-NET + BaseConv 基础上, 我们研究逐步引入多类型卷积的有效性, 由表 4 可以看出, 当上采样阶段五层全部使用中心差分残差模块时, mIoU 从 69.7 达到 71.0。而当上采样第一层使用中心卷积层, 其它四层全部使用中心差分残差模块时, mIoU 又达到 71.2, 在此基础上, 进一步在中间层分别加入 1、2、3 个快速傅里叶卷积残差模块时, 模型 mIoU 均有了提升, 其中, 加入 2 个时指标最高。最后, 我们还测试了如果将 MCFnet 的下采样层全部改为中心差分残差模块, 模型 mIoU 反而下降了。

Table 4. Ablation study of involving multiple types of convolutions**表 4.** 引入多类型卷积消融实验

模型架构	mIoU
U-NET + BaseConv	69.7
U-NET + BaseConv + 上采样五层全部使用中心差分残差模块	71.0
U-NET + BaseConv + 上采样第一层中心卷积 + 上采样其它四层全部使用中心差分残差模块	71.2
U-NET + BaseConv + 1 个快速傅里叶卷积残差模块	72.3
U-NET + BaseConv + 2 个快速傅里叶卷积残差模块(MCFnet)	72.5
U-NET + BaseConv + 3 个快速傅里叶卷积残差模块	72.2
MCFnet 的下采样层全部改为中心差分残差模块	71.4

(3) 不同背景权重比较

由于非烟物质远远少于背景(烟叶与传送带), 导致每幅图像中两者的空间占比严重不均, 因此, 我们在交叉熵损失中给背景赋予较低一些权重, 让模型更更多的考虑非烟物质的分割。表 5 给出了交叉熵损失中使用不同背景权重时的模型性能。

Table 5. Ablation study of weight selection for cross-entropy loss**表 5.** 交叉熵损失权重选择消融实验

背景权重	mIoU
0.30	62.8
0.25	65.2
0.20	70.1
0.15	72.4
0.10	72.9
0.05	72.2
0.01	67.3

(4) 不同损失函数比较

表 6 给出了使用不同损失函数训练 MCFnet 后得到的模型性能比较, 可以看出, 加权交叉熵损失与 Dice 损失的组合损失得到最好的性能结果。

Table 6. Ablation study of using different loss function
表 6. 不同损失函数消融实验

损失函数	mIoU
Cross-entropy Loss	71.6
加权 Cross-entropy Loss	72.9
Focal Loss	72.8
Dice Loss	73.2
加权 Cross-entropy Loss + Focal Loss	73.7
加权 Cross-entropy Loss + Dice Loss	74.1

5. 结论

本文提出了一种新的基于高光谱的非烟叶杂物检测方法。该方法以 U-NET 语义分割模型为骨架，适配高光谱三维张量数据输入。采用中心差分卷积、快速傅里叶卷积和标准卷积三类不同卷积，分别嵌入 U-NET 模型的降采样段，中间段和上采样段，较好地提取了高光谱的空谱联合特征；进一步给出了与该模型相适应的组合损失模型，以保证模型的训练充分稳定；设计与实现的非烟物质检测系统，并在实际复烤烟厂进行了测试。测试结果表明所提模型及方法对杂物检测的具有较低的虚警率和漏检率。下一步将对模型进行压缩和参数优化，实现模型与除杂系统的联动测试。

基金项目

本文由国家自然科学基金支持，基金号：61703209，同时由湖南中烟工业有限责任公司科技项目资助，项目编号：KY2023ZB0010，项目名称：基于高光谱技术的复烤打包工序非烟物质在线识别监测研究与应用。

参考文献

- [1] 刘配文, 温圣贤. 打叶复烤环节中非烟杂物的控制措施[J]. 作物研究, 2013(27): 51-54.
- [2] 张长华, 赵红枫, 胡伟, 等. 烟草原料中主要非烟物质的成因分析[J]. 2013, 34(1): 90-93.
- [3] 烤烟中非烟物质控制技术规程, 中华人民共和国烟草行业标准, YC/T 370-2010 [S]. 北京: 国家烟草专卖局, 2010: 4-7.
- [4] Li, Z.K., Fan, Y.J., Zou, Y.S., Wu, M.Y. and Liu, G.Y. (2014) Study and Application of Impurity Removal Methods in Tobacco Production. *Advanced Materials Research*, **1049-1050**, 1131-1134. <https://doi.org/10.4028/www.scientific.net/AMR.1049-1050.1131>
- [5] Kai, C., Qian, X., Bo, X., Chao, M. and Wei, Z.Z. (2019) A Machine Vision Algorithm for Foreign Bodies Detection in Tobacco Conveyor. 2019 *International Conference on Sensing and Instrumentation in IoT Era (ISSI)*, Lisbon, Portugal, 29-30 August 2019, 1-6. <https://doi.org/10.1109/ISSI47111.2019.9043657>
- [6] Zhang, F. and Zhang, X. (2011) Classification and Quality Evaluation of Tobacco Leaves Based on Image Processing and Fuzzy Comprehensive Evaluation. *Sensors*, **11**, 2369-2384. <https://doi.org/10.3390/s110302369>
- [7] Sun, J., Zhou, X., Wu, X., et al. (2016) Identification of Moisture Content in Tobacco Plant Leaves Using Outlier Sample Eliminating Algorithms and Hyperspectral Data. *Biochemical and Biophysical Research Communications*, **471**, 226-232. <https://doi.org/10.1016/j.bbrc.2016.01.125>
- [8] Li, Y. and Shen, Y. (2023) Design and Application of Tobacco Impurity Removal Model Based on Convolutional Neural Network. 2023 *IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI)*, Changchun, 26-28 May 2023, 1600-1605. <https://doi.org/10.1109/ICETCI57876.2023.10176382>
- [9] Ang, K.L.-M. and Seng, J.K.P. (2021) Big Data and Machine Learning with Hyperspectral Information in Agriculture. *IEEE Access*, **9**, 36699-36718. <https://doi.org/10.1109/ACCESS.2021.3051196>
- [10] Kumar, A., et al. (2020) UAV Based Remote Sensing for Tassel Detection and Growth Stage Estimation of Maize Crop Using Multispectral Images. *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, Waikoloa, HI, USA, 26 September-2 October 2020, 1588-1591.

-
- <https://doi.org/10.1109/IGARSS39084.2020.9323266>
- [11] Zhu, H., Chu, B., Zhang, C., *et al.* (2017) Hyperspectral Imaging or Presymptomatic Detection of Tobacco Disease with Successive Projections Algorithm and Machine-Learning Classifiers. *Scientific Reports*, **7**, Article Number: 4125. <https://doi.org/10.1038/s41598-017-04501-2>
- [12] Zhang, L., Ma, X., Li, Z. and Liu, Y. (2019) Application of Hyperspectral Imaging Technology in Classification of Tobacco Leaves and Impurities. 2019 *2nd International Conference on Safety Produce Informatization (IICSPI)* Chongqing, 28-30 November 2019, 157-160. <https://ieeexplore.ieee.org/abstract/document/9095975>
- [13] Tang, J., Zhou, H., Wang, T., Jin, Z., Wang, Y. and Wang, X. (2022) Cascaded Foreign Object Detection in Manufacturing Processes Using Convolutional Neural Networks and Synthetic Data Generation Methodology. *Journal of Intelligent Manufacturing*, **34**, 2925-2941. <https://doi.org/10.1007/s10845-022-01976-3>
- [14] Amri, M.B., Yedjour, D., El Amin Larabi, M. and Bakhti, K. (2022) Stadium Detection from Alsat-2 and Google-Earth Multispectral Images Using YOLOv5 and Mask R-CNN. 2022 *4th International Conference on Pattern Analysis and Intelligent Systems (PAIS)*, Oum El Bouaghi, 12-13 October 2022, 1-4. <https://doi.org/10.1109/PAIS56586.2022.9946887>
- [15] Wang, Y., Feng, W., Jiang, K., Li, Q., Lv, R. and Tu, J. (2023) Real-Time Damaged Building Region Detection Based on Improved YOLOv5s and Embedded System from UAV Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **16**, 4205-4217. <https://doi.org/10.1109/JSTARS.2023.3268312>
- [16] Zhu, X., Lyu, S., Wang, X. and Zhao, Q. (2021) TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-Captured Scenarios. 2021 *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, BC, 11-17 October 2021, 2778-2788. <https://doi.org/10.1109/ICCVW54120.2021.00312>
- [17] Wang, Y., Ouyang, Z., Han, R., Yin, Z. and Yang, Z. (2022) YOLOMask: Real-Time Instance Segmentation with Integrating YOLOv5 and Oriented Mask. 2022 *IEEE 22nd International Conference on Communication Technology (ICCT)*, Nanjing, 11-14 November 2022, 1646-1650. <https://doi.org/10.1109/ICCT56141.2022.10073387>