光电技术及应用

DOI：10.16818/j.issn1001－5868.20250222002

# 基于时域频域特征融合的轻量化圆棒类产品字符缺陷检测网络

吴庆华[1,2*]，张哲铭[1,2]，赵德华[3]

（1. 湖北工业大学 机械工程学院，武汉 430068；2. 现代制造质量工程湖北省重点实验室，武汉 430068；
3. 湖北中烟工业有限责任公司，武汉 430040）

**摘　要：** 针对字符重叠、错位及偏移等圆棒类产品表面字符缺陷检测难度大、检测精度低的问题，提出一种基于时域频域特征融合的轻量化圆棒类产品字符缺陷检测网络。将Scharr算子、时域特征和频域特征融合，获得清晰度高的字符边缘和细节信息。引入共享卷积设计轻量化检测头，在保证精度的前提下提高检测速度。实验结果表明，改进算法的平均精度均值达96.37%、召回率达91.13%、精确率达95.47%，较原始网络分别提高6.97%，4.62%，6.3%；检测速度达87.9 fps，较原始网络提高6.6 fps。

**关键词：** 圆棒类产品缺陷检测；字符缺陷；卷积神经网络；时域频域特征融合；轻量化

**中图分类号：** TP391.41　**文章编号：** 1001－5868(2026)03－0528－08

## Lightweight Cigarette Character Defect Detection Network Based on Time- and Frequency-Domain Feature Fusion

WU Qinghua[1,2], ZHANG Zheming[1,2], ZHAO Dehua[3]

(1. School of Mechanical Engineering, Hubei University of Technology, Wuhan 430068, CHN; 2. Hubei Key Laboratory of Modern Manufacturing Quality Engineering, Wuhan 430068, CHN; 3. Wuhan Cigarette Factory, China Tobacco Hubei Industrial LLC, Wuhan 430040, CHN)

**Abstract:** To overcome the challenges of low detection accuracy in identifying surface character defects on cylindrical products such as cigarettes, a lightweight defect-detection network based on time- and frequency-domain feature fusion is proposed. High-definition character edges and detailed information are obtained by integrating the Scharr operator with time- and frequency-domain features, thereby effectively capturing both local image variations and structural periodicity characteristics. A shared convolution mechanism is incorporated to develop a lightweight detection head, thus improving inference speed while maintaining detection accuracy. The optimized network achieves a mean average precision of 96.37%, a recall rate of 91.13%, and a precision of 95.47%, corresponding to enhancements of 6.97%, 4.62%, and 6.3% over those of the baseline model, respectively. Furthermore, the detection speed increases to 87.9 fps, which is 6.6 fps higher than that of the original implementation. This method provides an effective solution for industrial defect detection in cylindrical products and demonstrates strong applicability to cigarette quality-control systems.

**收稿日期：**2025－02－02.

**基金项目：**国家自然科学基金项目(51275158)；湖北中烟工业有限责任公司科研项目(2022JSGY4WH2B041).

* **通信作者：**吴庆华

**Key words**：　cigarette defect detection；character defects；convolutional neural network；time- and frequency-domain feature fusion；lightweight network

## 0　引言

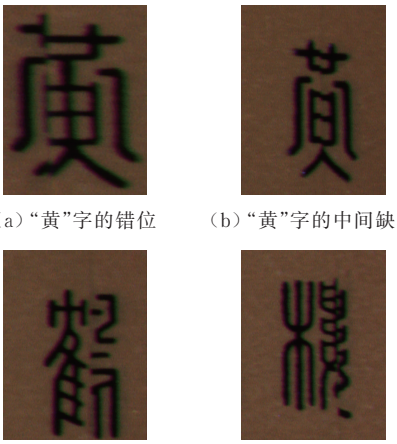
针对圆棒类产品缺陷的研究多集中于刺破、黄斑等表面缺陷[1]，忽略了卷接纸上的字符缺陷。字符承载生产批次、规格等关键信息，其完整性与清晰度直接影响产品质量[2]，但实际生产中字符重叠、错位、偏移等问题频发，严重降低合格率[3]。因此，高效检测字符缺陷已成为行业急需解决的关键问题。陈甦欣等[4]针对芯片表面的字符检测提出一种基于改进的区域生长算法分割和定位芯片表面字符区域方法，能够完整分割并定位芯片表面字符区域，其定位精度和效率较原有方法分别提升 5.3% 和 15.4%。苗相彬等[5]针对喷点字符边缘不清晰、模糊等问题，设计了一种基于秩正则几何变换的无监督字符缺陷算法，将检测精度从 0.935 提高到 0.983。

圆棒类产品表面，字符笔画与接缝混合形成的阴影会导致难以判断字符是否重叠；同时，表面弧度引起的字符扭曲也会使上述检测方法难以正常提取字符边缘信息，最终导致检测精度较低。乔帆等[6]针对多维时间序列存在维度高、维度间关系复杂、数据形态多变，难以选取区分力高的特征的问题，提出一种基于异构特征融合的多维时间序列分类算法，其融合时域、频域和区间统计值这三种特征，并对特征进行聚类以找到最有代表性的特征。

基于此，将时域频域特征相融合能够提升对复杂信号的检测精度，图像缺陷检测本质上也属于信号处理，时域分析方法能够直接提取字符的空间位置和边缘信息，频域分析方法则可以通过频率特性分离字符与背景的干扰。因此，本文提出一种融合时域和频域特征的卷积神经检测网络。设计了频域时域特征融合（FTFF）模块，通过多种视角捕获字符缺陷细节的结构特征。同时，设计了一种轻量化的检测头，结合 GroupNorm、共享卷积和 Scale 层的特点，优化检测头的结构和性能，在保持检测精度的前提下，有效提高了检测速度。

## 1　检测网络设计

圆棒类产品字符外观缺陷数据集来自实验团队研发的圆棒类产品缺陷检测仪器，如图 1 是典型的圆棒类产品字符缺陷图例。



（a）“黄”字的错位　（b）“黄”字的中间缺笔　（c）“鹤”字右侧左部缺笔　（d）“楼”字的右侧错位

图 1　圆棒类产品字符缺陷图像

### 1.1　总体结构设计

对圆棒类产品字符缺陷样本进行统计分析，设计如图 2 所示缺陷检测网络结构。将 FTFF 模块添加在骨干网络中，Scharr 算子与时域卷积和频域卷积结合，获取含有更多细节的字符缺陷特征。同时设计了一种轻量化检测头，融合共享卷积，在保持检测精度的同时提高检测速度。

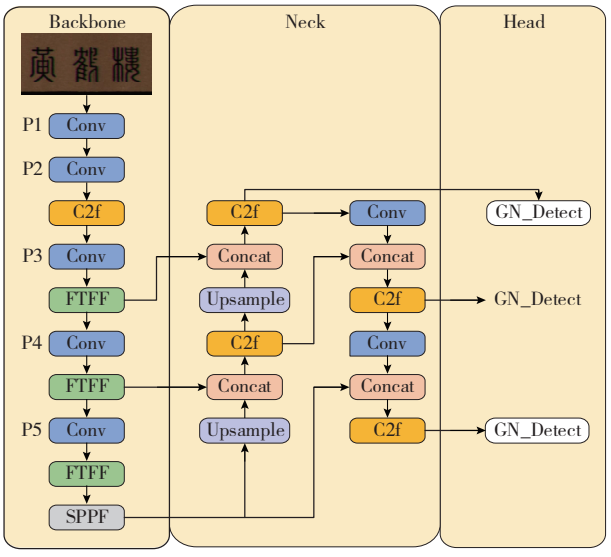



图 2　改进的网络结构

Backbone 部分用于从圆棒类产品字符图像中提取特征信息，这些特征信息是后续网络层进行目标检测的基础。Neck 部分的作用是进行圆棒类产品字符特征信息的融合和增强。Head 部分是检测

网络的决策部分，负责输出最终的圆棒类产品字符识别结果。

### 1.2 FTFF模块

在检测圆棒类产品字符的过程中，边缘检测能够提取字符的轮廓信息，增强字符与背景之间的差异，获得字符边缘的不连续或异常部分，这使得边缘检测成为缺陷定位的关键线索[7]。常用的边缘检测算子包括Sobel，Prewitt，Laplacian和Scharr算子等。其中，Sobel算子通过计算图像梯度来突出边缘，其抗噪性能较好，但对细节的敏感性稍弱。Prewitt算子与Sobel类似，但在处理高频噪声时表现稍显逊色。Laplacian算子通过二阶导数捕捉边缘，能够提取更多细节，但对噪声非常敏感，容易产生误检。相比之下，Scharr算子是一种优化的梯度算子，具有更高的旋转对称性和精度，在检测边缘的同时能保持较好的抗噪性能和细节保留能力[8]。在实际测试中，Scharr算子在处理复杂纹理或噪声较多的图像时，边缘清晰度和细节表现均优于其他算子。因此，Scharr算子的优越性能为字符缺陷检测中的特征提取提供了新的思路。结合其在边缘检测中的优势，设计了一种基于Scharr算子的特征提取模块，可有效提升对字符边缘的精确提取。Scharr算子结构如图3所示。

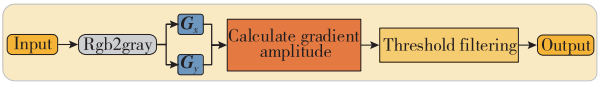



图3 Scharr结构图

Scharr算子是一种基于一阶导数的边缘检测算子，其原理是在图像梯度计算中对权值进行优化，使边缘检测结果具有更高的精度[9]。相较于传统的Sobel算子，Scharr算子通过引入更大的权重分布，增强了对边缘方向变化的敏感性。其在水平方向的卷积核$G_x$和垂直方向的卷积核$G_y$分别为：

$$G_x=\begin{bmatrix}-3 & 0 & 3\\ -10 & 0 & 10\\ -3 & 0 & 3\end{bmatrix} \tag{1}$$

$$G_y=\begin{bmatrix}-3 & -10 & 3\\ 0 & 0 & 0\\ 3 & 10 & 3\end{bmatrix} \tag{2}$$

在卷积操作中，Scharr算子计算图像灰度梯度的两个分量$G_x$和$G_y$，然后通过梯度幅值公式合成边缘强度：

$$G=\sqrt{G_x^2+G_y^2} \tag{3}$$

尽管Scharr算子在字符边缘检测中展现了出色的性能，能够高效捕捉边缘细节并保持良好的抗噪能力，但单独依靠边缘信息来识别字符缺陷仍存在一定的局限性。圆棒类产品表面字符缺陷往往不仅只局限于边缘的变化，还可能在图像的局部区域或全局结构中产生不同的影响。因此，单一的边缘检测方法不足以全面应对这些复杂的缺陷类型。

为解决该问题，可将Scharr算子提取的边缘特征与时域和频域信息相结合，可更加全面、准确地识别字符缺陷，克服单一特征在复杂背景或细微缺陷下的不足，从而提升字符缺陷检测的精度和鲁棒性。设计的特征提取模块结构如图4所示。

该模块分别对图片的时域特征和频域特征进行提取。时域特征的提取首先对原始图像进行Scharr算子处理提取边缘信息。接着进行时域卷积处理，对图像进行进一步的特征强化，强调图像中的局部变化如纹理和形状。处理后的时域特征再

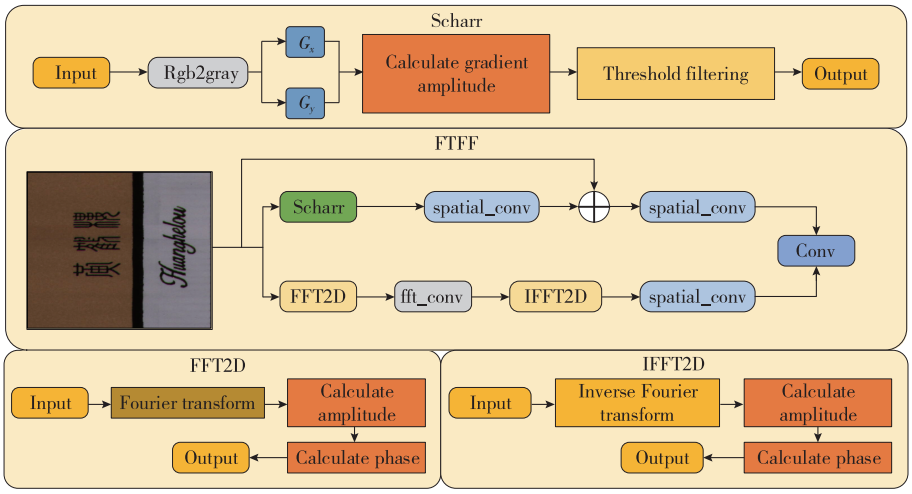



图4 FTFF模块结构图

与原始图像特征进行融合，确保在提取局部信息的同时，保留图像的整体结构。流程如图 5 所示。

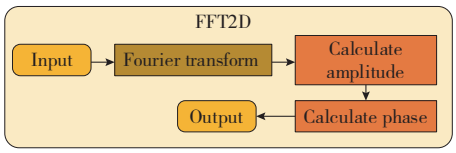



图 5　傅里叶变换流程

频域特征的提取则首先通过傅里叶变换将原始图像从时域转换到频域，该操作能够从全局角度揭示字符的频率分布，捕捉到图像的周期性变化和潜在的缺陷特征[10]。在频域中，通过频域卷积处理对图像的高频和低频信息进行筛选和强化，抑制背景噪声，同时突出能够有效反映字符缺陷特征的频率成分。再通过逆傅里叶变换将频域特征信息转换回时域特征信息，以便与时域特征进行融合。流程如图 6 所示。

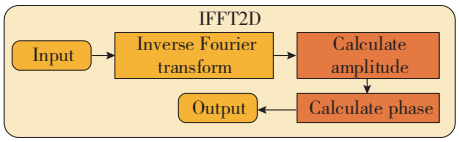



图 6　逆傅里叶变换流程

最后，将提取的时域特征和频域特征分别进行卷积处理，进一步增强各自特征的表现力。在完成时域和频域特征提取后，需要将这两类特征进行融合，以充分利用时域特征的局部边缘信息和频域特征的全局结构信息。采用自适应加权的融合方式，首先对时域特征 $F_t$ 和频域特征 $F_f$ 进行归一化，以保证不同特征尺度的一致性。然后采用加权融合策略，通过学习可训练参数 $\alpha$ 和 $\beta$ 来控制时域与频域特征的贡献比例，融合公式如下：

$$F_{融合}=\alpha\times F_t+\beta\times F_f \tag{4}$$

其中，$\alpha$ 和 $\beta$ 通过自适应学习的方式动态调整，以适应不同缺陷类型的特征分布。在融合后，进一步通过 1×1 组卷积进行通道压缩，减少冗余信息。相比直接拼接或简单加权求和，该方法能够有效保留字符缺陷的重要特征，同时减少计算量，提高检测效率。通过这种融合方法，时域特征能够提供精确的局部缺陷定位，频域特征则提供全局的结构信息，二者的互补性显著提高了圆棒类产品字符缺陷检测的准确性与鲁棒性。

### 1.3　轻量化检测头

由于在字符缺陷检测中融合了时域特征与频域特征，网络需要同时处理局部细节与全局频率信息，这不可避免地增加了计算量，尤其是在特征提取和融合阶段。该策略虽显著提升了检测精度与鲁棒性，却也造成网络检测速度下降，难以满足实际应用对实时性和高效性的要求。为解决该问题，设计了一种轻量化的检测头，通过优化结构和减少参数，有效降低了计算复杂度。其结构如图 7 所示。

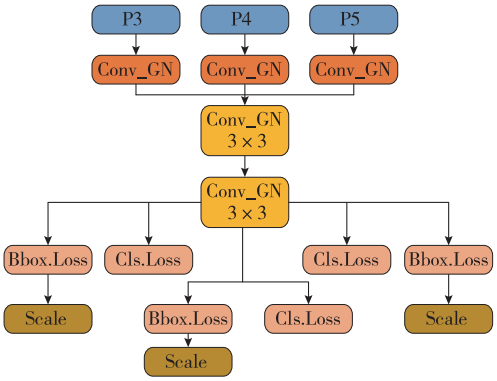



图 7　轻量化检测头结构

在检测头的结构中使用了共享卷积，即在分类和定位任务中共享卷积层参数。通过这一措施，有效减少了检测头的参数数量，使得网络更加轻量化。共享卷积在分类和定位两个任务之间共享底层特征，可在减少冗余参数的同时提高特征利用效率[11]。由于分类和定位任务需处理的特征存在一定相关性，适当使用共享卷积可以避免性能损失。

同时，在共享卷积的基础上引入 Scale 层，按检测头针对的目标尺寸对特征进行动态缩放，确保在处理不同尺度目标时具有更好的适应性。Scale 层能够对特征进行线性缩放，以调整不同目标尺度的特征响应，弥补共享卷积带来的检测目标尺度不一致的不足，使检测头在适应目标尺度变化时更加灵活[12]。Scale 层的流程如图 8 所示。

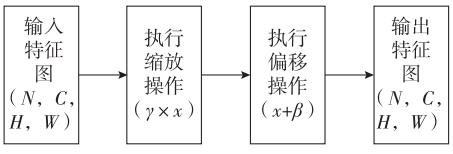



图 8　Scale 层流程图

由图可知，输入是一个形状为 $(N,C,H,W)$ 的四维张量，$N$ 代表图像个数，$C$ 代表图像的通道数，$H$ 和 $W$ 分别代表图像的高度和宽度。其核心公式为：

$$\mathrm{Output}(i,j,k,l)=\gamma_k\times\mathrm{Input}(i,j,k,l)+\beta_k \tag{5}$$

式中，$\gamma_k$是缩放参数，用于调整每个通道的幅度；$\beta_k$是偏移参数，用于调整每个通道的偏移；$i,j,k,l$分别表示图片索引、空间位置和通道索引。通过缩放和偏移，可以重新分布经过归一化的特征，使网络更好适应数据[13]。

但是在轻量化的过程中，不可避免地会损失一些检测精度。因此，本文使用GroupNorm替代BatchNorm。GroupNorm通过每一组通道的归一化，不依赖于批量大小，按通道划分小组，计算每组内部的均值和方差进行归一化。这种归一化方式提高了特征表达能力，有效提升网络在定位和分类上的表现。其流程如图9所示。

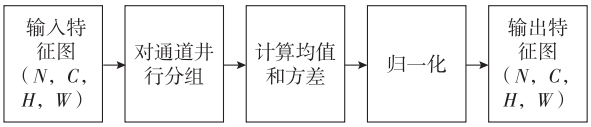



图9 GroupNorm流程图

在读取图像后，将图像的每个通道分成$G$个组，每组包含$C/G$个通道，计算每个组所有像素点的均值和方差[14]：

$$\mu_g=\frac{1}{N\times H\times W\times \frac{C}{G}}\sum_{i\in g}x_i \tag{6}$$

$$\sigma_g^2=\frac{1}{N\times H\times W\times \frac{C}{G}}\sum_{i\in g}(x_i-\mu_g)^2 \tag{7}$$

然后对每个像素点进行归一化：

$$\hat{x}_c=\frac{x_c-\mu_g}{\sqrt{\sigma_g^2}} \tag{8}$$

通过对每组通道进行归一化，GroupNorm能有效减少特征分布的漂移，稳定轻量化检测头的特征表达，有效补正轻量化检测头导致的精度下降问题，显著提升检测性能和鲁棒性。

## 2 实验结果与分析

### 2.1 实验准备

全部实验基于Windows操作系统，实验环境为Python3.9、CUDA11.0及Pytorch1.7。计算机CPU为AMD Ryzen 7 5800H with Radeon Graphics，8 GB的NVIDIA GeForce RTX3060，运行内存为16 GB。设置相关参数：训练轮数为200、批量大小为8、输入图片尺寸为1 280×640，随机种子为0且固定随机数、优化器SGD，初始学习率为0.001，权重衰减系数为0.000 5，动量因子为0.937。

实验使用的圆棒类产品样本均为实际生产所得，利用图10所示视觉系统采集图片。在获得大量图片数据后，筛选出成图效果优秀的2 372张图片，利用LabelImg软件手动对目标缺陷所在区域进行矩形框的标注。标注完成后，将数据集按7∶2∶1的比例分为训练集、测试集和验证集。

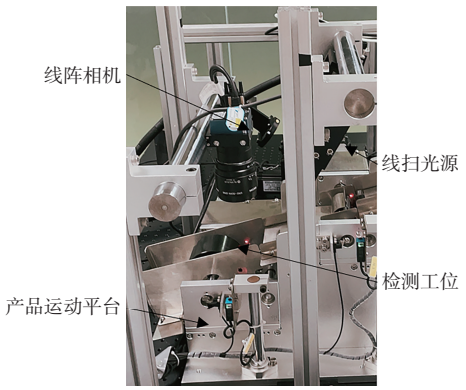



图10 圆棒类产品表面缺陷检测视觉实验平台

### 2.2 评估算法性能

采用目标检测中常用的准确率（P）、召回率（R）、平均精度均值（mAP）、平均检测速度（FPS）作为网络的评价指标[15]。在保持相同数据集划分的前提下，选取如Faster R-CNN[16]，SSD[17]，YOLOv8[18]，ResNeSt[19]，Mask Transfiner[20]，CMF-Net[12]和EfficientAD[21]等不同目标缺陷检测算法和改进算法进行性能对比，结果如表1所示。

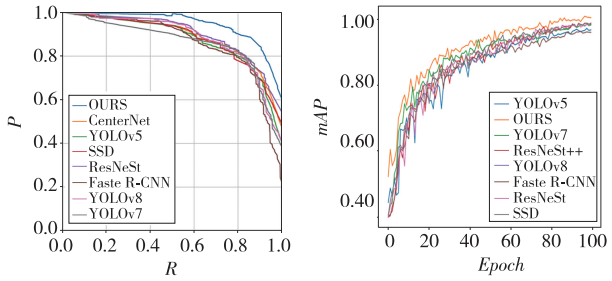
**表1 不同检测算法性能对比**

| 算法 | $P$/% | $R$/% | $mAP$@0.5/% | $FPS$/s$^{-1}$ |
|---|---|---|---|---|
| Faster R-CNN | 75.22 | 78.08 | 79.66 | 18.1 |
| SSD | 75.64 | 76.23 | 87.65 | 60.1 |
| YOLOv8 | 89.17 | 86.51 | 89.40 | 81.3 |
| ResNeSt | 89.44 | 86.02 | 87.68 | 38.1 |
| Mask Transfiner | 94.68 | 92.16 | 94.28 | 47.5 |
| CMF-Net | 91.29 | 90.31 | 92.94 | 37.1 |
| EfficientAD | 92.36 | 92.43 | 93.81 | 142.7 |
| 本文 | 95.47 | 91.13 | 96.37 | 87.9 |

结果表明，所提改进卷积神经网络在精确率（95.47%）、$mAP$@0.5（96.37%）和召回率（91.13%）等核心指标上均优于现有方法。同时，在不损失精度的前提下，检测速度达到87.9 fps，相比YOLOv8提升6.6 fps，且明显优于Mask Transfiner，CMF-Net和ResNeSt算法。虽然EfficientAD在速度上具有优势（142.7 fps），但其检

测精度低于本文方法，在字符缺陷检测任务上未能达到更优效果。

如图 11 所示的 *P*-*R* 曲线和 *mAP* 曲线说明了网络之间的性能差异。改进后的网络表现出最佳性能，其 *P*-*R* 曲线始终位于顶部，表明在整个召回范围内具有较高精度。*mAP*@0.5 值在整个训练过程中始终保持在 0.9 以上，证明了其上级的检测精度和网络稳定性。结果表明，整合 FTFF 和轻量化检测头的技术有效提高了基线网络中对圆棒类产品字符缺陷检测的准确性和鲁棒性，改进策略显著提高了网络处理圆棒类产品字符缺陷的能力。




（a）*P*-*R* 曲线　　（b）*mAP* 曲线

图 11　*P*-*R* 曲线和 *mAP* 曲线

### 2.3　消融实验

为验证所提出的改进的有效性，设计一组消融实验。将各改进模块单独添加到算法中，从而验证各改进模块的有效性，并进行组合添加，验证各改进模块是否能够共同提升算法性能。结果见表 2。

**表 2　消融实验结果**

| 组别 | FTFF | 轻量化检测头 | *mAP*@0.5/% | *FPS*/s$^{-1}$ |
|---|---|---|---|---|
| 第一组 | − | − | 89.40 | 81.3 |
| 第二组 | + | − | 97.23 | 54.4 |
| 第三组 | − | + | 88.64 | 101.2 |
| 第四组 | + | + | 96.37 | 87.9 |

其中，FTFF、轻量化检测头表示为本文提出的改进方法。引入 FTFF 方法后，算法的 *mAP*@0.5 提升了 7.8%，验证了该方法对圆棒类产品字符缺陷的检测有提升。加入轻量化检测头后，与原始网络相比，检测速度提升了 20 fps，证明该方法能够提升网络对圆棒类产品字符缺陷的检测速度。在采用 FTFF 方法的基础上，进一步融合检测头，其 *mAP*@0.5 保持在 96.37%，*FPS* 提升了 6.6 fps，验证了本文算法的可行性。

同时，为更直观验证所提算法的有效性，将两种算法的字符缺陷检测可视化结果进行对比，如图 12 所示。可以看出，原始算法对三种字符进行检测时存在一定量的漏检情况，而本文改进的算法能够有效检出。




（a）原算法对“黄”　（b）改进后的算法对“黄”
（c）原算法对“鹤”　（d）改进后的算法对“鹤”
（e）原算法对“楼”　（f）改进后的算法对“楼”

图 12　网络改进前后的检测效果对比图

为使人工检测和数字化检测这两种不同检测方法的检测结果维持一致，需要对圆棒类产品字符缺陷进行一致性实验。使用目前主流的 Kappa 系数法对检测算法进行一致性评价，通过实验的 Kappa 系数指标，验证所设计的圆棒类产品字符缺陷检测算法相对于标准的一致性程度。实验内容包括由算法检测出的缺陷检测结果与标准值（即专家组人工判定结果）进行比较，并运用 Minitab 软件对测量结果进行一致性分析[22]。检测数据如表 3 所示。

**表 3　一致性实验结果**

| 检验员 | Kappa 值 | *P*/% | 漏判率/% | 误判率/% |
|---|---|---|---|---|
| 1 | 0.96 | <0.01 | 4.26 | 0.32 |
| 2 | 0.96 | <0.01 | 3.24 | 0.48 |
| 字符检测算法 | 0.96 | <0.01 | 1.64 | 0.00 |
| 所有检验员与标准 | 0.96 | <0.01 | | |

## 3　结论

本文提出一种基于时域与频域特征融合的轻量化检测网络，有效提升了字符缺陷检测的精度和鲁棒性。实验结果表明，该方法在精确率（95.47%）、召回率（91.13%）和 *mAP*（96.37%）方

面优于现有算法,同时在保证高精度的情况下,将检测速度提升至87.9 fps。相比传统方法,所提方法结合Scharr算子优化边缘特征,同时融合时域与频域信息,使检测更加精准。轻量化检测头的引入减少了计算开销,提升了检测速度。消融实验验证了FTFF模块与轻量化检测头的有效性,确保了模型在复杂字符缺陷检测任务中的稳定性。所提方法在精度、实时性和计算效率之间取得了良好平衡,为圆棒类产品字符缺陷检测提供了一种高效可靠的解决方案。

## 参考文献:


[1] Ding Y, Zhou H, Wu H, et al. SCS-YOLO: a defect detection model for cigarette appearance [J]. Electronics, 2024, 13(18): 3761.

[2] 张月华. 机器视觉与烟支缺陷自动检测技术的融合模式研究[J]. 无线互联科技, 2021, 18(10): 122-123.
Zhang Yuehua. Research on the fusion pattern of machine vision and automatic detection of tobacco defects[J]. Wireless Internet Science and Technology, 2021, 18(10): 122-123.

[3] 吴玉生, 李安虎, 万亚明, 等. 基于机器视觉的烟草在线检测技术研究进展[J]. 激光与光电子学进展, 2024, 61(8): 39-52.
Wu Yusheng, Li Anhu, Wan Yaming, et al. Progress in research on tobacco online inspection technology based on machine vision[J]. Laser & Optoelectronics Progress, 2024, 61(8): 39-52.

[4] 陈甦欣, 赵安宁, 罗乐文. 基于机器视觉的芯片字符区域分割和定位算法[J]. 组合机床与自动化加工技术, 2024(4): 10-13, 18.
Chen Suxin, Zhao Anning, Luo Le|Yue)Wen). Chip character area segmentation and localization algorithm based on machine vision[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2024(4): 10-13, 18.

[5] 苗相彬. 喷点字符识别模型轻量化与缺陷检测问题研究[D]. 成都: 西南交通大学, 2022.
Miao Xiangbin. Research on the problems of model slimming for dot character recognition and defect detection [D]. Chengdu: Southwest Jiaotong University, 2022.

[6] 乔 帆, 王 鹏, 汪 卫. 基于异构特征融合的多维时间序列分类算法[J]. 计算机科学, 2024, 51(2): 36-46.
Qiao Fan, Wang Peng, Wang Wei. Multivariate time series classification algorithm based on heterogeneous feature fusion [J]. Computer Science, 2024, 51(2): 36-46.

[7] 孙 洋, 孙 林, 夹尚丰, 等. 一种基于边缘形状特征的字符缺陷检测算法[J]. 计算机应用与软件, 2023, 40(9): 177-183.
Sun Yang, Sun Lin, Jia Shangfeng, et al. Character defect detection algorithm based on edge shape features[J]. Computer Applications and Software, 2023, 40(9): 177-183.

[8] 罗 山, 侯俊涛, 郑 彬. 刹车盘表面缺陷边缘检测算法研究[J]. 机械设计与制造, 2025(8): 192-195, 200.
Luo Shan, Hou Juntao, Zheng Bin. Research on edge detection algorithm of brake disc surface defect[J]. Machinery Design & Manufacture, 2025(8): 192-195, 200.

[9] 姚 昱, 郑 洋, 程子燚, 等. 基于Scharr算子的计算鬼成像边缘提取技术[J]. 激光与光电子学进展, 2024, 61(10): 129-136.
Yao Yu, Zheng Yang, Cheng Ziyi, et al. Edge detection technology based on computational ghost imaging using scharr operator [J]. Laser & Optoelectronics Progress, 2024, 61(10): 129-136.

[10] 周承乐, 石 茜, 李 军, 等. 光谱—频域属性模式融合的高光谱遥感图像变化检测[J]. 遥感学报, 2024, 28(1): 105-120.
Zhou Chengle, Shi Qian, Li Jun, et al. Spectral-frequency domain attribute pattern fusion for hyperspectral image change detection[J]. National Remote Sensing Bulletin, 2024, 28(1): 105-120.

[11] Tian Z, Shen C, Chen H, et al. FCOS: a simple and strong anchor-free object detector[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(4): 1922-1933.

[12] Ding Y, Zhou T, Xiang L, et al. Cross-contrast mutual fusion network for joint MRI reconstruction and super-resolution[J]. Pattern Recognition, 2024, 154: 110599.

[13] 闫建红, 冉同霄. 基于YOLOv8的轻量化无人机图像目标检测算法[J]. 图学学报, 2024, 45(6): 1328-1337.
Yan Jianhong, Ran Tongxiao. Lightweight UAV image target detection algorithm based on YOLOv8[J]. Journal of Graphics, 2024, 45(6): 1328-1337.

[14] 蒋世杰, 夏秀山, 翟 伟, 等. 基于ODE扩散模型的多类异常检测和定位[J]. 智能系统学报, 2025, 20(2): 376-388.
Jiang Shijie, Xia Xiushan, Zhai Wei, et al. ODE diffusion model for multiclass anomaly detection and localization[J]. CAAI Transactions on Intelligent Systems, 2025, 20(2): 376-388.

[15] 张重生, 陈 杰, 纵瑞星, 等. 基于Transformer的低质场景字符检测算法[J]. 北京邮电大学学报, 2022, 45(2): 124-130.
Zhang Chong|Zhong)Sheng), Chen Jie, Zong Ruixing, et al. Transformer based scene character detection over low quality images [J]. Journal of Beijing University of Posts and Telecommunications, 2022, 45(2): 124-130.

[16] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[17] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector [C]//Computer Vision – ECCV 2016, 2016: 21-37.

[18] Varghese R, M S. YOLOv8: a novel object detection algorithm with enhanced performance and robustness [C]// 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), 2024: 1-6.

[19] Zhang H, Wu C, Zhang Z, et al. ResNeSt: split-attention networks [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2022: 2735-2745.

[20] Ke L, Danelljan M, Li X, et al. Mask transfiner for high-quality instance segmentation [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 4402-4411.

[21] Batzner K, Heckler L, König R. EfficientAD: accurate visual anomaly detection at millisecond-level latencies [C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024: 127-137.

[22] 陈星月, 贾子彦, 李　青, 等. 基于改进YOLOv5s的免疫组化阳性细胞计数方法[J]. 中国医学物理学杂志, 2025, 42(2): 167-174.
Chen Xingyue, Jia Ziyan, Li Qing, et al. Improved YOLOv5s based method for immunohistochemically positive cell counting[J]. Chinese Journal of Medical Physics, 2025, 42(2): 167-174.

**作者简介：**

**吴庆华**(1978—)，男，湖北省天门市人，博士，副教授，研究方向为三维视觉检测与自动控制。