

DOI:10.13232/j.cnki.jnju.2023.03.004

双端输入型嵌套融合多尺度信息的织物瑕疵检测

曲 皓¹,狄 岚^{1*},梁久祯²,刘 昊²

(1.江南大学人工智能与计算机学院,无锡,214000;2.常州大学计算机与人工智能学院,常州,213164)

摘 要:针对织物瑕疵检测中复杂纹理区域误判和边缘检测模糊问题,提出一种双端输入型网络架构 WNet,使用两个骨干分支分别提取多尺度局部和全局特征,依靠自注意力机制的全局建模能力,在卷积深层网络中补充全局信息,减少深层网络中纹理特征的冗余.为了减少深层网络中局部细节信息的丢失,提出一种轻量级双分支池化金字塔,将浅层多尺度细节特征引入深层模块.搭建多尺度嵌套双分支模块,将各级尺度下不同感受野的特征信息进行融合,该模块更加关注瑕疵特征,可以生成较精确的预测图.实验证明,提出的方法在 ZJU-Leaper 的四个数据集上的综合评价指标较优,尤其是 f_{measure} 、阳性预测率和阴性预测率较高.

关键词:织物瑕疵检测, WNet, 自注意力机制, 双分支池化金字塔, 多尺度嵌套双分支模块

中图分类号: TP391

文献标志码: A

Double-ended input type nested fusion of multi-scale information for fabric defect detection

Qu Hao¹, Di Lan^{1*}, Liang Jiuzhen², Liu Hao²

(1. College of Artificial Intelligence and Computer, Jiangnan University, Wuxi, 214000, China;

2. College of Computer and Artificial Intelligence, Changzhou University, Changzhou, 213164, China)

Abstract: To address the problems of complex texture region misclassification and blurred edge detection in fabric defect detection, this paper proposes a double-ended input network architecture WNet which uses two backbone branches to extract multi-scale local and global features, respectively. By relying on the global modeling capability of the self-attentive mechanism in the right backbone branch, the global information is added to the convolutional deep network to reduce the redundancy of texture features in deep network. To reduce the loss of local detail information in deep network, this paper proposes a lightweight dual-branch pooling pyramid to introduce shallow multi-scale detail features into the deep module. In the feature fusion stage, a multi-scale dual-branching nested module is built to screen high-gradient texture features and defective features. This module pays more attention to the feature information of different perceptual fields at different scales, which betterly screen out the defective features and generate more accurate prediction maps. Experimentally, the comprehensive evaluation indexes of the method on four datasets of ZJU-Leaper are proved to be better, especially f_{measure} , Positive Predictive Value and Negative Predictive Value.

Key words: fabric defect detection, WNet, self-attention, dual-branch pooling pyramid, multi-scale dual-branch nested module

织物瑕疵检测在纺织工业生产过程中有至关重要的作用,随着产业链升级,生产线产生的瑕疵

类型越来越多,缺陷和背景纹理的区分度也越来越低,需要泛化能力更强的算法,在复杂多变的检

基金项目:江苏省石油化工过程关键设备数字孪生技术工程研究中心开放课题(DT2020720)

收稿日期:2023-01-09

* 通讯联系人, E-mail: dilan@jiangnan.edu.cn

测环境下仍能保持稳定的精度.已有的织物瑕疵呈多态性,主要分污渍类瑕疵和破损类瑕疵.污渍类瑕疵与背景纹理在颜色空间上有明显区别,但过度关注颜色特征会导致高梯度纹理区域的误判.破损类瑕疵和背景纹理存在高度内在相似性,有时会相互穿插而导致边缘预测不够准确,这一直都是该类瑕疵的检测难点.

1 相关工作

传统织物瑕疵的检测方法主要分四种:基于模型的方法^[1]、低秩分解方法^[2-3]、局部最优方法^[4]和视觉显著性方法^[5],但这些方法执行复杂纹理类型的瑕疵检测任务时检测精度较低.传统检测方法需要对简单纹理织物进行单独建模,而基于深度学习的方法可以对复杂纹理织物进行统一建模,降低了成本,提高了检测效率.基于深度学习的检测方法在过去几年得到了很大的发展,基于深度学习的织物瑕疵检测方法不仅在周期性简单纹理织物中的检测效果十分理想,对复杂纹理织物的瑕疵也有较高的检测精度.

Liu et al^[6]提出一种多阶段GAN(Generative Adversarial Networks)框架来应对迭代快速的新型面料:第一阶段,GAN在新的无缺陷样品中生成合理的瑕疵来更新数据集,在此基础上改进语义分割网络完成第二阶段的织物瑕疵检测任务. Chen et al^[7]提出基于遗传算法的Gabor Faster R-CNN用于显著性织物瑕疵检测,能较好地减少纹理干扰,但对于复杂纹理中的低显著性瑕疵检测,效果不理想. Jing et al^[8]将Mobile-Unet用于织物瑕疵检测来解决实际工业中不同种类瑕疵因数量差异导致的数据不平衡问题,但容易对梯度较大的无瑕疵纹理区域产生误判.现有的深度学习方法针对复杂纹理织物中的不规则瑕疵依然存在边缘检测模糊和纹理区域错检问题.

现有的织物瑕疵检测算法通常借助预训练好的模型(如VGG-16^[9], ResNet-18(34, 50, 101)^[10], Dense-169^[11]等)作为骨干网络来提高模型的收敛速度和检测精度.近年的研究表明^[12-14],多尺度信息可以更好地提取不同尺寸瑕疵的特征,有助于提高瑕疵检测的精度.自注意力机制具有较强的全局建模能力,Transformer^[15]被广泛应用在计

算机视觉领域中. Dosovitskiy et al^[16]提出 vision Transformer,在各大目标检测数据集上取得了不错的效果. Liu et al^[17]提出基于转换窗口的Swin Transformer,在COCO数据集上的检测效果十分优异. Tian et al^[18]串联Swin Transformer和特征金字塔,搭建双向对象-上下文优先排序模型,将Swin Transformer作为检测任务的骨干网络.

许多研究者更关注特征融合方式的效能. Lin et al^[19]构造了一种自顶向下、带有横向连接的特征金字塔来构建多个尺度的深层语义特征. Hou et al^[20]提出DSSNet,对深层的输出先上采样,然后拼接浅层的输入,以短连接的方式集成多层次特征. Zhao et al^[21]在PSPNet中设计了Pyramid Pooling Module,聚合不同区域的上下文信息来提高网络获取全局信息的能力.上述方法采用单分支骨干网络,通过添加深层模块捕获全局特征.但在深层网络中,局部特征冗余容易引起瑕疵边缘的错误预测,全局特征稀疏容易引起背景区域的错检.在面对复杂纹理样本时,由于解码块不能很好地从特征图中筛选出瑕疵,很容易对高梯度纹理区域判断错误,使误检率升高.

许多研究受焦点损失^[22]的启发,更关注边界像素以及显著性特征图中的模糊像素判断^[23-24].通过分配更高的权重来增大这些模糊像素的占比,这在理论和技术上确实令人信服,但不是所有的模糊像素都在边界附近.最近的研究^[25]指出,深度网络预测存在两种困境,即高置信度但不准确的预测以及低置信度但准确的预测.

为了克服数据不平衡引起的边缘预测模糊的问题, Tu et al^[26]提出SEAL function,将交叉熵与分割错误的超像素联系起来,在计算损失值时对分割错误的像素值赋予更大权重. Jiang et al^[27]提出多标签损失函数来弱化数据不平衡的影响,将模型的输出经Softmax归一化后,再通过Sigmoid函数将输出映射到 $[0.5, 1]$,但这样会弱化偏离像素对损失值的贡献,大大影响模型的收敛速度.

以上大部分改进损失函数的方法都采用外部加权的方式,目的都是增加错误预测像素损失值的贡献,引导模型沿正确方向迭代.值得注意的是,若过度地增大错误预测像素的损失贡献,容易

混淆瑕疵和背景交界区域,进而陷入局部最优,导致高梯度纹理区域的错检。

为了解决上述问题,本文提出融合双路多尺度信息的网络模型(WNet)用于复杂纹理织物的瑕疵检测。为了提高训练速度,采用预训练的ResNet-18(RBlock-0~4)作为骨干网络左分支。为了降低深层网络中背景特征对瑕疵特征的干扰,采用预训练的Swin Transformer(STBlock-1~4)作为骨干网络右分支,借助自注意力机制出色的全局建模能力在卷积深层网络中补充全局上下文信息。为了减少深层网络中瑕疵细节信息丢失,提出一种轻量级双分支池化金字塔(Dual-Branch Pooling Pyramid,DBPP)将浅层多尺度细节特征引入深层模块。特征融合阶段,在双分支池化金字塔的基础上设计新的特征融合模块,即多尺度嵌套双分支模块(Multiscale Dual-Branch Nested Module,DBN)。最后,为了降低高梯度纹理区域的误检率,对二元交叉熵损失函数进行改进来提高最终的检测精度。

本文的主要创新:

(1)提出一种双端输入型网络框架WNet,用两个预训练的骨干网络分别提取多尺度局部和全局特征。提出由DBPP,RBlock-0及其相应特征融合模块构成的瑕疵特征增强器,将低级多尺度细节特征导入最终的特征融合模块,降低瑕疵细节信息损失,提高瑕疵边缘预测的准确性。RBlock-1~4,STBlock-1~4及其对应的特征融合模块组成纹理特征过滤器,在卷积深层网络中补充全局上下文信息,为特征融合模块提供更丰富的纹理特征信息,有利于特征融合模块对高梯度纹理特征中的瑕疵特征进行甄别。最后,将全部尺度特征融合器输出的特征图融合,获得预测图像。

(2)为了减少深层网络中瑕疵细节信息的缺失,提出一种轻量化双分支池化金字塔,以较少的计算消耗在深层网络中引入多尺度低级细节特征。在特征融合阶段,为了进一步提炼、甄别、增强输入特征,筛选高梯度纹理特征及瑕疵特征,通过嵌套DBPP构建特征融合器DBN。每一尺度的特征融合模块都融合多个感受野范围内的信息,为最终的输出融合模块补充更全面的上下文信息。

(3)用改进二元交叉熵损失函数 $W_G BCE$

Loss进行深度监督,在四个瑕疵数据集上进行综合评估,并与其他最先进的方法进行比较。WNet在所有测试数据集上的综合评价指标均较优,尤其是 $f_{measure}$ 、阳性预测率和阴性预测率较高。

2 融合双路多尺度信息的织物瑕疵检测方法

以预训练的骨干网络作为编码器能极大地提高模型的收敛速度和检测精度,但低尺度特征提取模块信息来源单一,下采样过程中细节信息的丢失会影响深层全局上下文信息的提取,在网络深层模块中会出现局部特征冗余和全局特征稀疏的问题,大大影响瑕疵边缘的预测精准度以及高梯度纹理区域的判断。因此,本文提出一种融合双路多尺度信息的织物瑕疵检测方法WNet。

2.1 WNet模型架构 本文提出一种网络架构思想:以 n 个不同骨干网络作为支撑,使用一个特征融合器对多种不同空间内的特征信息进行融合、筛选。依赖多阶段的损失函数对不同阶段给予不同的损失权重和计算方式,引导各阶段的迭代趋向。理论上可以根据任务类型设置 n 个骨干分支来实现多端输入,但是具有多骨干分支的架构过于复杂,对特征融合器的过滤能力要求较高,还会带来高昂的计算开销。因此,针对织物瑕疵检测,本文采用双端输入的网络模型WNet。

WNet由瑕疵特征增强器和背景特征过滤器构成(图1)。DBPP,RBlock-0及其相应特征融合模块组成瑕疵特征增强器。卷积浅层网络中蕴含丰富的细节特征,将原始多尺度细节特征引入最后的特征融合模块来减少瑕疵细节信息的损失,提高瑕疵边缘预测精度。消融实验去除DBPP的 $f_{measure}$ 降低了0.95%,RBlock-1~4,STBlock-1~4及其对应的特征融合模块组成背景特征过滤器。在卷积深层网络中补充全局上下文信息可以为特征融合模块提供更丰富的纹理特征信息,有利于特征融合模块对高梯度纹理特征中的瑕疵特征进行筛选。消融实验中不使用ST骨干分支,检测精度降低了6.67%。在特征融合阶段,各级模块在损失函数引导下生成各个阶段的输出特征图。所有特征图首先上采样至原始分辨率,然后进行最

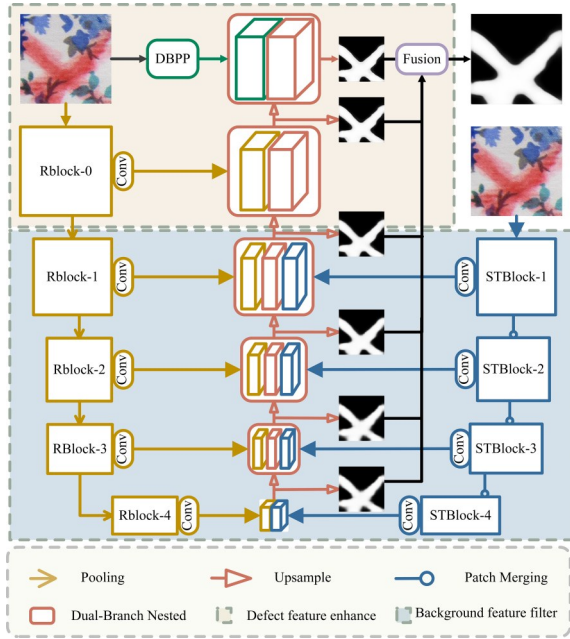


图1 WNet的网络结构

Fig. 1 The architecture of WNet

终融合,生成最终的预测图像。

为了以较低的计算成本捕获多尺度不同感受野的特征信息,本文设计了轻量化的DBPP,在此基础上搭建DBN。DBN作为特征融合模块能较好地筛选瑕疵特征,2.2将详细介绍模块构成。瑕疵特征增强器的两个输出特征图和背景特征过滤器的四个输出特征图首先被采样至输入图像分辨率,再经过融合模块生成预测图像。

为了提高预测置信度,将高斯概率分布权重 W_G 赋予二元交叉熵损失函(BCE Loss),提高瑕疵检测精度。在消融实验中,使用 W_G BCE Loss的 f_{measure} 比标准BCE Loss提高0.79%。假设 x 为原始输入图像,则WNet可以概括为式(1):

$$\begin{aligned} WNet(x) = & DBN(DBPP(x) \oplus Residual_w)_0 \oplus \\ & DBN(RBlock_0 \oplus Residual_w)_1 \oplus \\ & DBN(RBlock_1 \oplus STBlock_1 \oplus Residual_w)_2 \oplus \\ & DBN(RBlock_2 \oplus STBlock_2 \oplus Residual_w)_3 \oplus \\ & DBN(RBlock_3 \oplus STBlock_3 \oplus Residual_w)_4 \oplus \\ & \oplus Conv(RBlock_4 \oplus STBlock_4) \end{aligned} \quad (1)$$

其中, $DBN(\cdot)_0$ 为0级嵌套双分支模块, $Residual_w$ 为剩余 W 结构的输出特征图, \oplus 为拼接融合。

2.2 双分支嵌入特征融合模块

2.2.1 双分支池化金字塔

浅层高分辨率特征图保留了瑕疵的边缘、形状、大小等丰富的局部信息,而局部特征提取的充分性决定了预测图的细节呈现,进而决定检测精度的上限。若特征提取模块太关注局部信息,忽略全局信息,容易引发纹理像素点及区域的误判,可以通过大核卷积堆积或者扩张卷积的方式关注全局信息。如图2所示,假设原始特征图的高为 H ,宽为 W ,中心像素点 $(H/2, W/2)$ 需要与顶点 $(1, 1)$ $(1, W)$ $(H, 1)$ 和 (W, H) 的特征信息进行交互,一般采用卷积层堆叠或空洞卷积池化金字塔(Atrous Spatial Pyramid Pooling, ASPP^[14])的方式。图中紫色块代表中心像素点,红色块代表边缘目标像素点,蓝色块代表 3×3 卷积核的感受野,黄色块代表膨胀倍率为2的 3×3 卷积核的感受野,虚线表示当前颜色卷积核的已捕获信息的范围。

如图2a所示,普通的卷积层堆叠方式,需要不断地堆叠层数,渐进式地学习全局上下文信息,如式(2)和式(3)所示:

$$P_{xy}^1 = F_1(P_{xy}) = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} w_{ij}^1 P_{ij} \quad (2)$$

$$P_{xy}^{L_s} = F_{L_s-1}(P_{xy}^{L_s-1}) = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} w_{ij}^{L_s-1} P_{ij}^{L_s-1} \quad (3)$$

P_{xy}^1 表示原始特征图 P 经历一层卷积后 (x, y) 的像素值,中心特征与边缘特征进行交互需要堆叠 $L_s(2^{L_s-1} = H/2)$ 个卷积层。但卷积层增多使参数量大幅提升,过多卷积层也会引发梯度消失。

使用ASPP中的多倍率池化卷积方式,如图2b所示,以最大倍率为 $H/2$ 的空洞卷积即可学习全局信息,如式(4)所示:

$$P_{xy}^D = \sum_{i=x-D}^{x+D} \sum_{j=y-D}^{y+D} w_{ij}^D P_{ij} \quad (4)$$

其中, D 表示卷积空洞倍率。为了加强不同空洞倍率感受野内特征信息的交互性,往往需要增加多个不同倍率的空洞卷积层。但由于高倍率空洞卷积要保证输出特征图的尺寸与原始特征图相同,边界处的卷积操作添加了过多填充,这不仅引入了过多的噪声,而且提升了该模块的参数数量。

为解决上述问题,本文提出双分支池化金字塔DBPP。如图2c所示,通过池化融合邻域内四

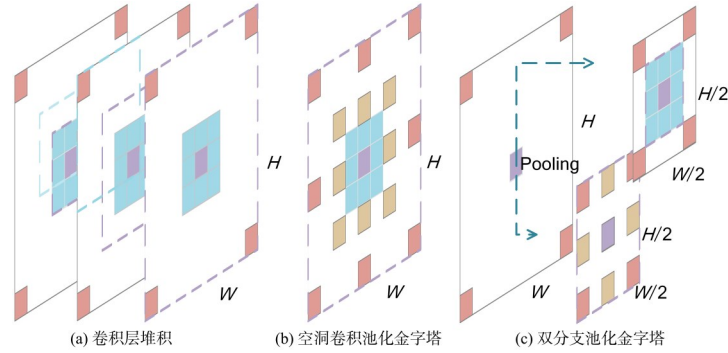


图 2 DBPP 模块与 ASPP 模块以及卷积堆叠的对比

Fig. 2 Comparison of DBPP module with ASPP module and convolutional stacking

个像素的特征信息,同时降低特征图大小. 这种操作可以极大地降低模型运算量,双分支分别采用 3×3 卷积和 $3 \times 3_{\text{dilation}=2}$ 卷积,保证感受野范围均匀扩大. 对每一幅原始特征图 $P_{(H,W)}$ 先进行一次池化,再进行两个分支的卷积,如式(5)所示:

$$\begin{cases} P_L^1\left(\frac{x}{2}, \frac{y}{2}\right) = \sum_{i=\frac{x}{2}-1}^{\frac{x}{2}+1} \sum_{j=\frac{y}{2}-1}^{\frac{y}{2}+1} w_{ij}^L P_{ij}^P \\ P_R^1\left(\frac{x}{2}, \frac{y}{2}\right) = \sum_{i=\frac{x}{2}-D}^{\frac{x}{2}+D} \sum_{j=\frac{y}{2}-D}^{\frac{y}{2}+D} w_{ij}^R P_{ij}^P \end{cases} \quad (5)$$

中心特征与边缘特征建立联系仅需要搭建 $L_n(2^{L_n+1}=H/2)$ 个卷积层.

2.2.2 嵌套双分支模块 通过嵌套 DBPP 实现的多尺度嵌套双分支模块(DBN)如图 3 所示,主要用于解码阶段的特征融合,也用于浅层高分辨率特征图的细节特征提取.

DBN 的形状类似二叉树,是编解码结构的特征提取模块. 两个特征提取分支关注不同感受野

中的局部信息,可对原始特征图的局部信息进行交叉验证. 通过与原始分支的融合,类似给原始特征图赋予空间注意力机制,起瑕疵增强的作用. 最后一个融合输出加在原第一层卷积层上,求出当前模块输出. 假设 x 为原始输入图像,则经过 DBN 处理后的输出概括为:

$$\begin{cases} DB(x) = \\ f(F_L^1(f_1(X)) \oplus F_R^1(f_1(X)) \oplus f_1(X)) + f_0(x) \\ X = f_0(x) \end{cases} \quad (6)$$

其中, f 代表 Conv, F_L^1 代表第一级左分支处理后的特征图, F_R^1 表示第一级右分支处理后的特征图.

多尺度嵌套双分支模块的编码层深度 L_n 的上限由输入图像的尺寸决定,该模块的参数量会随着中间层数 M 的增大而增大. 图 4 展示了嵌套双分支模块与其他先进的特征提取模块 FLOPs (Floating-Point Operations per Second) 的比较. 将 $256 \times 256 \times 3$ 的输入特征图转换为 $256 \times 256 \times 64$ 的输出特征图,计算中间通道数 M 从 32 到

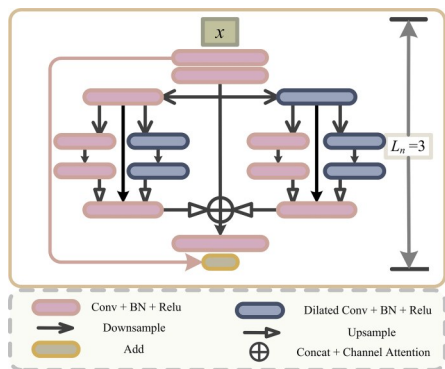


图 3 多尺度嵌套双分支模块结构图

Fig. 3 Multiscale dual-branch nested module

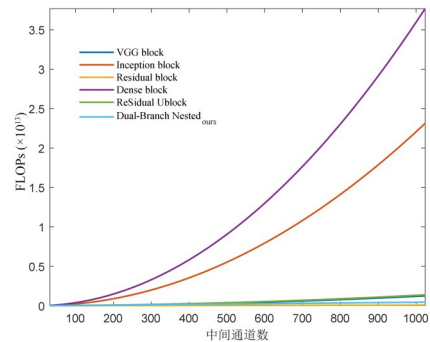


图 4 不同模块随中间通道数变化的 FLOPs

Fig. 4 The FLOPs of different modules changing with the number of intermediate channels

1024 的每个特征提取模块的成本. 由图可见, DBN 的 FLOPs 随着中间通道数的增大, 增长速度慢于其他的特征提取模块. 本文将双分支嵌套模块作为解码阶段的特征融合器, 每个尺度采用不同深度的 DBN, 在增加网络深度的同时不会显著提高计算消耗.

2.3 改进二元交叉熵损失函数 置信度偏低的预测在织物瑕疵检测任务中十分常见, 特别是在瑕疵与高梯度纹理的交界处. 损失函数的目的是缩小预测值和实际值之间的差异, 即提高预测置信度. 损失函数可以总结为:

$$\begin{cases} L = \operatorname{argmin} \left\{ \sum_{i=1}^W \sum_{j=1}^H |P_{ij} - G_{ij}| \right\} \\ L = \operatorname{argmin} \left\{ \sum_{m=1}^M w_m |P_{G=0} - 0| + \sum_{n=1}^N w_n |P_{G=1} - 1| \right\} \\ c = \frac{|1 - P_{G=0}|}{100\%} = \frac{|P_{G=1}|}{100\%} \quad M + N \leq W \times N \end{cases} \quad (7)$$

$P_{G=0} \neq 0, P_{G=1} \neq 1$

其中, L 为理想预测约束 (损失函数值), P 为预测图, G 为 Ground Truth, c 为预测置信度, i, j 分别表示图像的横纵坐标, M 为真实值为 0 但预测值不为零的像素点数量, N 为真实值为 1 但预测值不等于 1 的像素点数量. 理想状态下 $L = 0$, 实际上高置信度 ($0.5 < c < 1$) 和低置信度 ($0 < c < 0.5$) 会达到数值均衡, 这可以理解为随着预测接近 Ground Truth 时焦点损失使敏感度太低, 而当预测稍微偏离 Ground Truth 时敏感度增长过快又会陷入局部最优.

针对上述问题, 本文在标准二元交叉熵损失函数的基础上进行改进, 通过高斯概率分布权重 W_G 赋予偏离 Ground Truth 的预测更大的损失贡献. 为了减少假阳性和假阴性预测, 利用 F 误差平滑损失函数曲线, 使显著性预测图更偏向真阳性和真阴性预测. 损失函数的定义如下:

$$\begin{cases} L = -\frac{1}{K} \sum_{k=1}^K \{ W_G [y_k \lg(x_k) + (1 - y_k) \lg(1 - x_k)] + \lambda F \} \\ F = x_k(1 - x_k) \\ W_G = 1 + \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x_k - \mu)^2}{2\sigma^2}} \bigg/ \max P(x) \end{cases} \quad (8)$$

其中, K 表示像素个数, $\mu = 0.5, \sigma = 0.2, \lambda = 0.1$, σ 和 λ 的取值由参数消融实验确定. 本文沿用多阶段损失计算方式^[28-29], WNet 包含六个尺度的输出和一个最终融合输出, 将每个阶段的 W_G -BCE Loss 相加作为总的损失函数:

$$L_{\text{Total}} = \sum_{s=1}^6 w_{\text{stage}}^s L_{\text{stage}}^s + w_{\text{fuse}} L_{\text{fuse}} \quad (9)$$

其中, s 代表多阶段输出. 图 5 展示了 W_G BCE Loss 与其他改进的 BCE Loss 在相同预测值下的损失函数值的比较. 递减曲线的 Ground Truth 为 1, 递增曲线的 Ground Truth 为 0.

由图可见, 预测值趋近标签值的范围在 $[0.15, 0.6]$, W_G BCE Loss 的敏感性显著优于 BCE Loss 和 Closs, 在 $[0.6, 0.8]$, W_G BCE Loss 的梯度最平滑, 可以规避局部最优的产生. 改进后的损失函数可以提高预测置信度, 提高检测精度.

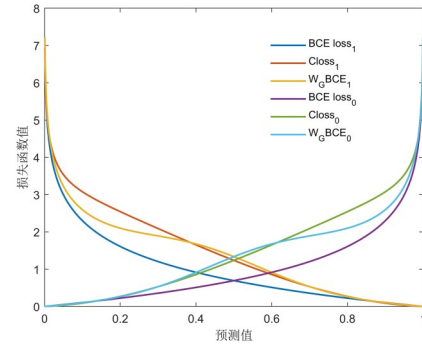


图 5 W_G BCE Loss 与其他损失函数的可视化比较

Fig. 5 Visual comparison of W_G BCE loss with other loss functions

3 实验与结果分析

实验配置: PyTorch 1.9.1, cuda 11.0, Python 3.7, 基于 PyTorch 深度学习框架构建网络模型; 处理器为 CPU 5600x @ 4.6 GHz, GPU NVIDIA 3060 ti, 显存 8 G, 内存 16 G. 五个 DBPP 的中间通道数为 $[8, 16, 32, 64, 128]$, DBPP 左右分支卷积核大小都为 3, 右分支空洞卷积倍率为 2. 学习率设置为 10^{-4} , 使用 Adam 算法训练模型.

3.1 数据集 使用浙江大学构建的织物瑕疵数据集 ZJU-Leaper^[30], 它由 19 种不同纹理类型的 98777 张织物图组成, 带有详细的注释. 选用前 15 种复杂纹理织物的 18855 张瑕疵样本, 图 6 展示了

15种织物样本的图片. 织物样本按纹理复杂度被分为四个组, 命名为 Group 1~4. 第一组包含 White Plain, Thick Stripe, Thin Stripe; 第二组包含 Dot Pattern, Houndstooth, Gingham, Knot Pattern; 第三组包含 Twill Plaid, Blue Plaid, Brown Plaid, Gray Plaid, Red Plaid; 第四组包含 Floral Print1, Floral Print2, Floral Print3. 为了提高训练速度, 将图像尺寸压缩至 256×256 构成新的数据集, 按 9:1 的比例将数据样本分为训练集和测试集. 实验中, 输入图像首先缩放至 320×320 , 然后随机裁切为 224×224 , 再随机经过旋转、镜像、翻转操作进行图像增强.

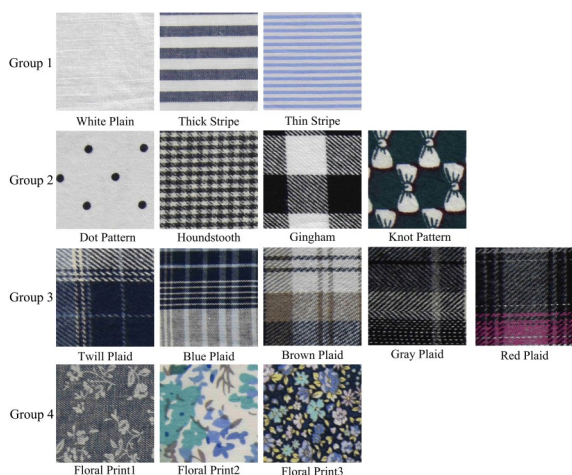


图 6 ZJU-Leaper 数据集中四组织物 15 种背景纹理的示例图像

Fig. 6 Example images for 15 patterns of four groups of ZJU-Leaper dataset

3.2 评价指标 使用 $f_{\text{measure}}^{[31]}$ 进行评价:

$$f_{\text{measure}} = \frac{(\gamma^2 + 1) \times \text{Precision} \times \text{Recall}}{\text{Precision} + \gamma^2 \times \text{Recall}} \quad (10)$$

具体的几个指标: *Precision* 为查准率, 即阳性预测率 (Positive Predictive Value, *PPV*); *Recall* 为查全率, 即真阳性率 (True Positive Rate, *TPR*); 假阳性率 (False Positive Rate, *FPR*), 即误检率; 阴性预测率 (Negative Predictive Value, *NPV*). 另外, 依据参考文献[3], 将 γ 设置为 1.

显然, f_{measure} 综合考量了 *TPR* 和 *PPV*. f_{measure} , *TPR*, *PPV*, *NPV* 越高, *FPR* 越低, 说明该方法的瑕疵检测性能越好. 为了综合评估本文的模型并排除低置信预测的影响, 添加另外三个指标, 分别

为 *MAE*, *PR* 曲线和 *ROC* 曲线.

MAE: 预测值和 Ground Truth 像素差值的绝对值之差的平均值, 定义为:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|$$

MAE 越小, 说明预测值与真值偏差越小.

PR 曲线: 给定一个预测的概率图, 通过 0~255 之间的不同阈值进行二值掩码, 并与 Ground Truth 进行比较, 计算 *Precision* 和 *Recall*, 可以绘制 *PR* 曲线.

ROC 曲线: 纵轴是 *TPR*, 横轴是 *FPR*, 当数据集中的阳性和阴性样本分布不均时, *ROC* 曲线比 *PR* 曲线更能保持稳定.

3.3 定量比较 将 WNet 与其他先进方法进行比较, 分别为 PicaNet^[32], F3Net^[24], U²Net^[29], GCPANet^[33], SCWS^[34], DDNet^[35], RCSB^[25]. 表 1~4 列出了四个数据集上七个评价指标的实验结果, 表中黑体字表示结果最优. 所有算法中, 分组训练模型的检测精度普遍高于整体训练模型, 这个结果也符合经验判断.

许多方法对瑕疵边缘产生了大面积的误判, 使 *TP* (预测值和真实值同为正例) 和 *FP* (预测值为正例, 真实值为负例) 显著升高, 一定程度上提高了 *TPR*. 高梯度纹理区域的误检也会提高 *FP*, 使 *NPV* 提升缓慢甚至下降. 例如, SCWS 和 F3Net 的 *TPR* 较高, *NPV* 和 f_{measure} 较低, 尤其是 *MAE* 显著高于其他方法, 说明这些方法在织物瑕疵区域和纹理区域产生了大量的错误判断. 后续的定性比较实验也证实部分方法产生了大量的漏检和误检.

由表 1~4 可见, 本文方法 WNet 没有显著降低运算速度, 在四个测试数据集上的七个评价指标均表现了良好的竞争性能. 与其他先进方法相比, 在四个测试数据集上, WNet 的 f_{measure} , *PPV* 和 *NPV* 都比较理想. WNet 在 Group 1 和 Group 2 中的性能最出色, f_{measure} , *PPV* 和 *NPV* 均显著高于其他方法. 在纹理较复杂的 Group 3 和 Group 4 上, 各种方法的性能均有一定程度的下降, 但 WNet 的综合性能依旧比较理想. 尽管 WNet 的运算速度稍慢于 GCPANet, 但其泛化能力有利于其应用于较复杂的检测环境.

表1 本文算法和对比方法在ZJU-Group 1上的定量实验结果的比较

Table 1 Quantitative experimental results of our model and other models on ZJU-Group 1

方法	f	TPR	FPR	PPV	NPV	MAE	Time (ms)
PicaNet (Group)	75.62%	77.89%	2.28%	73.48%	98.12%	0.020	99
PicaNet (Total)	65.56%	67.85%	2.66%	63.43%	97.30%	0.004	99
F3Net (Group)	76.92%	82.92%	2.38%	71.72%	98.17%	0.032	23
F3Net (Total)	74.93%	79.32%	2.11%	71.00%	97.83%	0.098	23
U ² Net (Group)	76.01%	78.77%	2.21%	73.43%	98.13%	0.016	27
U ² Net (Total)	73.78%	77.18%	2.43%	70.67%	97.69%	0.010	27
GCPANet (Group)	76.80%	82.75%	2.61%	71.64%	97.93%	0.150	20
GCPANet (Total)	76.30%	81.51%	2.24%	71.72%	97.73%	0.125	20
SCWS (Group)	72.90%	79.05%	8.08%	67.64%	93.10%	0.080	22
SCWS (Total)	74.12%	82.15%	7.01%	67.51%	96.45%	0.069	22
DDNet (Group)	76.87%	81.08%	2.55%	73.07%	98.09%	0.014	20
DDNet (Total)	74.18%	77.89%	3.14%	70.81%	97.88%	0.012	20
RCSB (Group)	74.85%	76.43%	1.82%	73.34%	98.07%	0.015	76
RCSB (Total)	72.28%	77.41%	8.74%	67.78%	96.38%	0.065	76
WNet (Group) (Ours)	78.08%	80.66%	2.19%	75.67%	98.61%	0.016	25
WNet (Total)	75.77%	80.65%	2.61%	71.43%	98.33%	0.014	25

表2 本文算法和对比方法在ZJU-Group 2上的定量实验结果的比较

Table 2 Quantitative experimental results of our model and other models on ZJU-Group 2

方法	f	TPR	FPR	PPV	NPV	MAE	Time (ms)
PicaNet (Group)	67.48%	73.46%	4.16%	62.40%	98.02%	0.012	95
PicaNet (Total)	65.48%	68.42%	2.88%	62.77%	96.93%	0.007	95
F3Net (Group)	76.36%	82.15%	2.42%	71.34%	98.11%	0.128	19
F3Net (Total)	73.76%	78.74%	2.55%	69.37%	98.00%	0.146	19
U ² Net (Group)	73.34%	77.69%	2.80%	69.46%	98.11%	0.014	25
U ² Net (Total)	70.34%	74.70%	3.01%	66.46%	97.73%	0.013	25
GCPANet (Group)	74.38%	80.32%	3.11%	69.26%	97.13%	0.196	20
GCPANet (Total)	75.19%	80.69%	2.62%	70.40%	97.35%	0.171	20
SCWS (Group)	69.83%	77.29%	9.32%	63.68%	93.51%	0.090	20
SCWS (Total)	69.02%	82.64%	11.50%	59.25%	92.21%	0.113	20
DDNet (Group)	74.54%	78.72%	2.96%	70.78%	98.17%	0.015	18
DDNet (Total)	74.29%	78.74%	4.31%	70.33%	98.05%	0.017	18
RCSB (Group)	72.03%	79.55%	5.41%	65.80%	98.81%	0.012	73
RCSB (Total)	71.76%	76.24%	5.30%	67.78%	97.73%	0.035	73
WNet (Group) (Ours)	76.66%	79.34%	2.29%	74.16%	98.45%	0.013	24
WNet (Total)	74.64%	78.30%	2.47%	71.30%	98.30%	0.014	24

图7是本文方法和对比方法的PR曲线和ROC曲线。由图可见,WNet的曲线十分平滑,在大部分数据集中都涵盖了其他模型的曲线,说明WNet的结果更接近二元预测。值得注意的是,由

于SCWS的输出特征加强模块过分增强了预测的置信度,所以其ROC曲线出现了过短的情况。

以 $\bar{f}=\frac{1}{n}\sum_{i=1}^nf_i(n=8)$ 评判模型的泛化能力。

表 3 本文算法和对比方法在 ZJU-Group 3 上的定量实验结果的比较

Table 3 Quantitative experimental results of our model and other models on ZJU-Group 3

方法	f	TPR	FPR	PPV	NPV	MAE	Time (ms)
PicaNet (Group)	63.72%	67.91%	2.63%	60.02%	98.23%	0.021	95
PicaNet (Total)	55.66%	61.07%	3.42%	51.14%	97.56%	0.006	95
F3Net (Group)	72.51%	77.95%	2.08%	67.78%	98.38%	0.035	20
F3Net (Total)	70.30%	74.47%	1.83%	66.58%	98.35%	0.127	20
U ² Net (Group)	71.89%	75.63%	2.07%	68.49%	98.57%	0.012	25
U ² Net (Total)	63.92%	71.34%	2.91%	57.91%	97.92%	0.011	25
GCPANet (Group)	70.97%	77.29%	5.55%	65.60%	97.68%	0.230	18
GCPANet (Total)	72.48%	79.07%	2.39%	66.90%	97.84%	0.191	18
SCWS (Group)	72.96%	81.49%	10.26%	66.05%	91.35%	0.102	20
SCWS (Total)	70.54%	80.60%	10.10%	62.71%	92.22%	0.101	20
DDNet (Group)	70.11%	76.03%	3.96%	65.04%	98.52%	0.013	19
DDNet (Total)	69.94%	75.75%	6.04%	64.96%	98.34%	0.022	19
RCSB (Group)	71.74%	80.75%	2.79%	64.55%	98.99%	0.008	74
RCSB (Total)	64.54%	67.86%	6.52%	61.52%	98.25%	0.048	74
WNet (Group) (Ours)	74.16%	79.60%	2.05%	69.42%	99.07%	0.013	23
WNet (Total)	70.62%	76.62%	2.16%	65.49%	98.80%	0.010	23

表 4 本文算法和对比方法在 ZJU-Group 4 上的定量实验结果的比较

Table 4 Quantitative experimental results of our model and other models on ZJU-Group 4

方法	f	TPR	FPR	PPV	NPV	MAE	Time (ms)
PicaNet (Group)	63.12%	68.15%	3.26%	58.79%	97.16%	0.010	95
PicaNet (Total)	63.52%	68.14%	3.11%	59.49%	97.03%	0.006	95
F3Net (Group)	73.95%	79.06%	2.36%	69.46%	98.01%	0.044	21
F3Net (Total)	71.71%	77.83%	2.77%	66.49%	98.14%	0.177	21
U ² Net (Group)	73.66%	78.88%	2.76%	69.09%	98.28%	0.012	26
U ² Net (Total)	70.14%	75.42%	3.05%	65.55%	98.02%	0.014	26
GCPANet (Group)	72.04%	78.39%	5.39%	66.64%	96.99%	0.350	19
GCPANet (Total)	74.74%	80.81%	2.68%	69.51%	97.63%	0.160	19
SCWS (Group)	71.17%	86.32%	9.76%	60.55%	93.42%	0.033	21
SCWS (Total)	70.11%	82.93%	10.06%	60.72%	93.19%	0.100	21
DDNet (Group)	72.76%	78.81%	4.88%	67.57%	98.28%	0.024	20
DDNet (Total)	72.69%	78.76%	5.62%	67.49%	98.26%	0.029	20
RCSB (Group)	71.47%	74.57%	3.85%	68.62%	98.53%	0.018	78
RCSB (Total)	71.01%	74.14%	4.05%	68.13%	98.50%	0.026	78
WNet (Group) (Ours)	74.86%	78.16%	2.22%	71.84%	98.45%	0.014	25
WNet (Total)	73.50%	76.10%	2.20%	71.08%	98.26%	0.013	25

表 5 对比了各个方法的 Model Size、GFLOPs、参数量 Params 和综合泛化能力 \bar{f} , 表中黑体字表示结果最优。

由表可见, 在保证模型泛化能力的基础上,

WNet 没有显著增加计算消耗, 模型运算量和参数量都较低。与其他方法相比, WNet 相对轻量化, 模型保存也仅需 113 M。对比其他方法, WNet 的 \bar{f} 最高, 为 74.79%, 进一步证明 WNet 对多种织

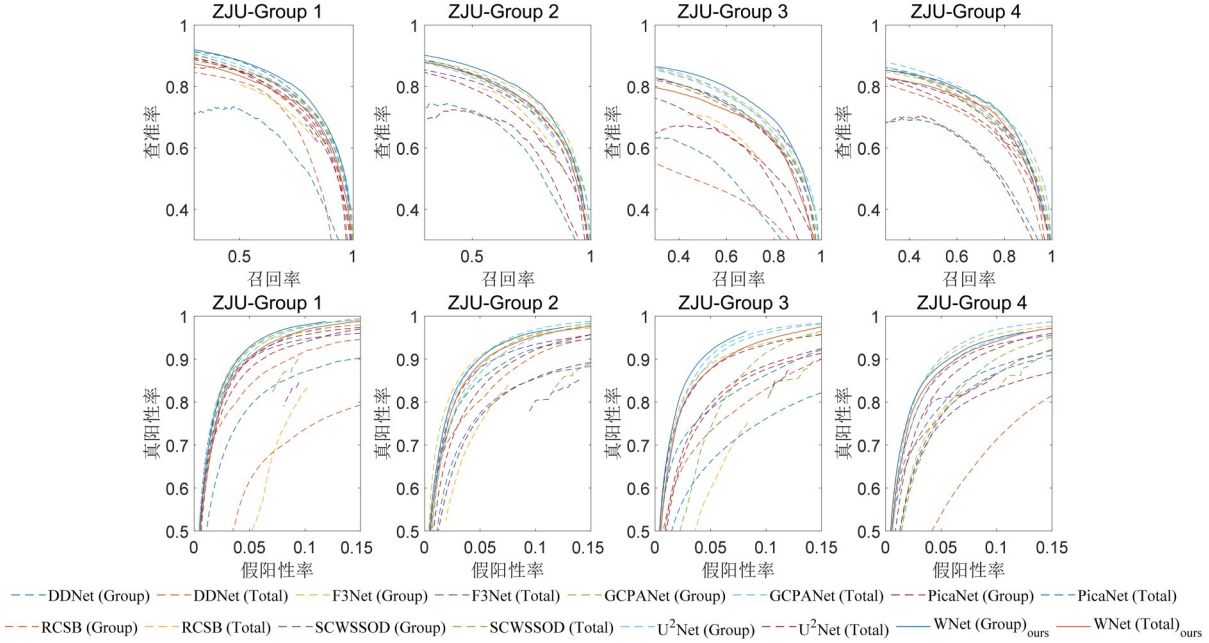


图7 各个方法的PR曲线(上一行)和ROC曲线(下一行)对比

Fig. 7 PR curves (up) and ROC curves (down) of each method

表5 各个方法的大小、运算量、参数量和泛化能力的对比

Table 5 Model size, GFLOPs, parameters and generalization ability of each model

方法	\bar{f}	GFLOPs	Params (M)	Size (M)
PicaNet	65.02%	80.6	49.7	189
F3Net	73.80%	6.7	25.5	97.8
U ² Net	71.64%	28.8	44	168
GCPANet	74.11%	16.6	67.1	256
SCWS	71.33%	21.8	63.5	242
DDNet	73.17%	4.6	33.3	12.8
RCSB	71.31%	218	33.5	137
WNet	74.79%	7.3	27.7	113

物纹理瑕疵的泛用性. 超轻量模型DDNet也预示WNet有很大的参数优化空间,这也是下一步的研究方向.

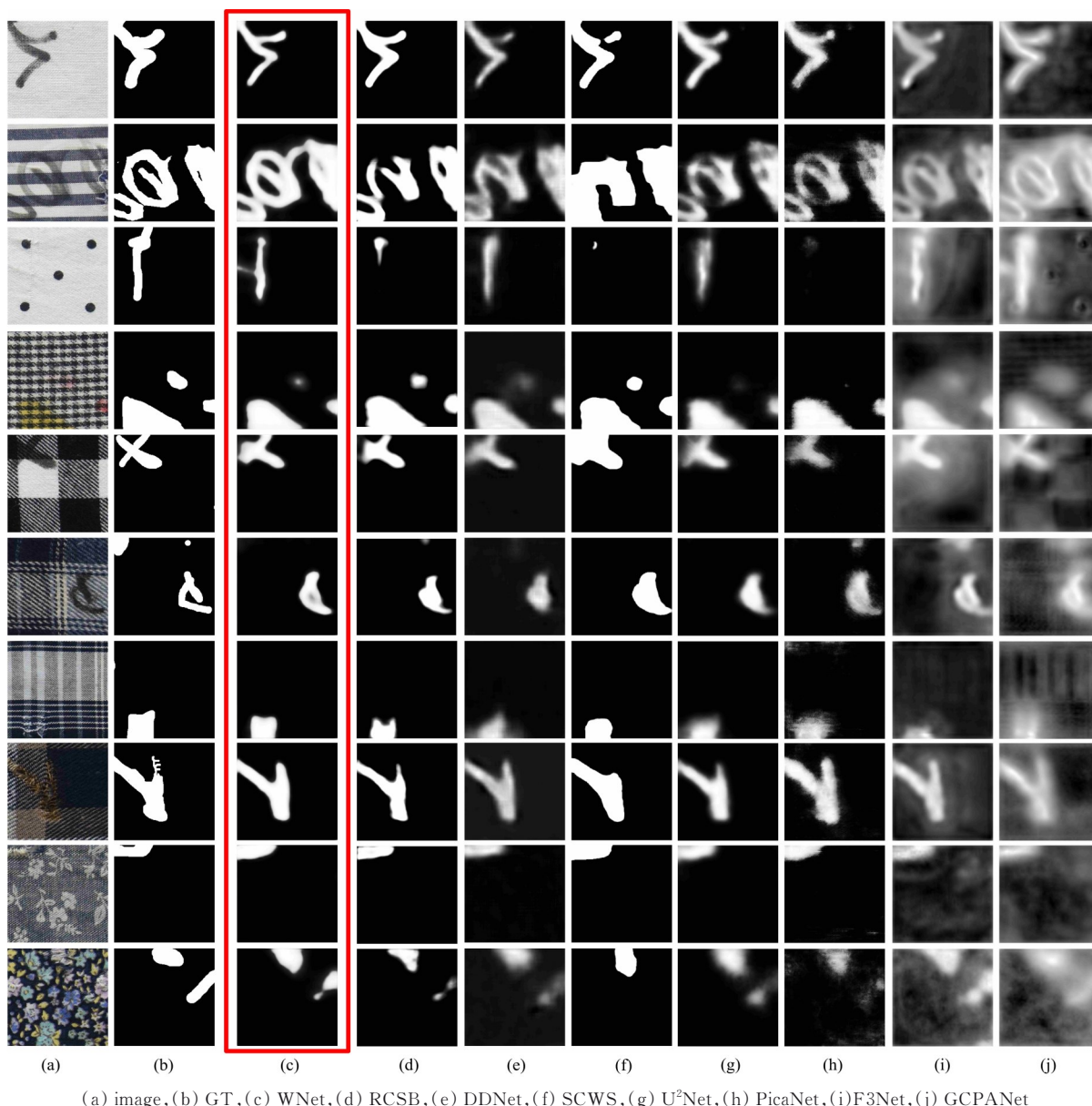
3.4 定性比较 可视化对比如图8所示. 由图可见, F3Net和GCPANet尽管其 TPR 和 PPV 相对高于其他方法,但这两种方法出现了大规模的边缘模糊,也存在大规模的高梯度纹理区域的误判,造成其 MAE 显著高于其他方法.

SCWS会强迫融合关联特征,它的预测图置信度过高. 如图8中第二行和第六行所示, SCWS将大面积的瑕疵边缘纹理误判为瑕疵,一定程度

上提高了 TP 进而影响了评价指标 TPR . 第三行中,低显著性褶皱类瑕疵与纹理的关联度很高, SCWS特征融合模块会过滤瑕疵特征,因此出现了大面积的漏检.

WNet比U²Net, DDNet, F3Net的边缘过渡更清晰,且预测置信度更高. 与其他方法相比, WNet能充分过滤背景纹理特征,提取更多的织物细节特征. 在第二行和第六行, WNet没有对无暇区域产生大面积误判. 综上可见, WNet可以产生细节丰富的特征预测图.

3.5 消融实验 为了验证本文提出的双端输入、各个模块以及改进的损失函数的有效性,在ZJU-Group 4数据集上进行消融实验,表6展示了各个方法的定量比较,表中黑体字表示结果最优. 其中, Bce和WgBce分别表示采用Bce Loss和Wg-Bce Loss进行监督训练, R&S_Pret表示双路骨干网络载入预训练参数, NoPret表示骨干网络不载入预训练参数. 在编码阶段, Res&ST表示使用双路骨干, +ASPP表示使用ASPP作为细节补充模块, +DB表示使用嵌套双分支作为细节补充模块. 在解码阶段, ASPP和DB分别表示将ASPP和嵌套双分支模块作为解码模块, SBP代表使用单分支普通池化模块, 验证嵌套双分支模



(a) image, (b) GT, (c) WNet, (d) RCSB, (e) DDNet, (f) SCWS, (g) U²Net, (h) PicaNet, (i) F3Net, (j) GCPANet

图 8 WNet 与其他方法的定性比较

Fig. 8 Qualitative comparison of WNet with other methods

块的分支有效性. 由表可见, 使用单个残差分支的 f_{measure} 为 68.19%, 使用单个 Swin Transformer 分支的 f_{measure} 为 68.23%, 对比双路分支的 f_{measure} , 分别降低 6.67% 和 6.63%, 证明了本文双骨干分支的有效性. 将 WNet 中的 DBN 模块改用 ASPP 替代初级特征提取模块和解码模块, f_{measure} 分别降低 0.93% 和 1.93%. 另外, U²Net 使用 $W_{\text{G}}\text{BCE Loss}$ 的 f_{measure} 提升 0.36%, WNet 使用 $W_{\text{G}}\text{BCE Loss}$ 的 f_{measure} 提升 0.79%, 说明本文提出的 $W_{\text{G}}\text{BCE Loss}$

损失函数是有效的.

因为 ZJU-Group 2 中既有简单纹理类型的 Dot Pattern 纹理, 也有类周期低对比度的 Gingham 纹理和 Knot Pattern 纹理, 即 ZJU-Group 2 综合了 Group 1 和 Group 3 的纹理特征, 所以在 Group 2 上进行了相同的模块消融实验, 进一步验证算法模块的泛化有效性. 实验结果如表 7 所示, 表中黑体字表示结果最优. 由表可见, Group 2 上的实验结果与 Group 4 上的实验结果呈规律近似

表6 损失函数、DBN模块以及连接方式在ZJU-Group 4上的消融研究

Table 6 Ablation study for Loss function, each module and hybrid connection on ZJU-Group 4

损失函数及预训练	编码阶段	解码阶段	\bar{f}	TPR	FPR	PPV	NPV	FPS (ms)
Bce	U ² Net		73.66%	78.88%	2.76%	69.09%	98.28%	26
WgBce	U ² Net		74.02%	78.47%	2.40%	70.05%	98.35%	26
WgBce R&S_Pret	Res&ST+ASPP	DB	73.93%	78.36%	2.52%	69.97%	98.54%	27
WgBce R&S_Pret	Res&ST+ASPP	ASPP	73.56%	77.30%	2.37%	70.18%	98.45%	24
WgBce R&S_Pret	Res&ST+DB	ASPP	72.93%	73.53%	2.00%	72.34%	98.22%	24
WgBce S_Pret	ST+DB	DB	68.23%	74.40%	3.67%	63.00%	98.35%	22
WgBce R_Pret	Res+DB	DB	68.19%	73.02%	3.17%	63.97%	98.26%	18
WgBce R&S_Pret	Res&ST	DB	73.91%	77.97%	2.36%	70.25%	98.48%	23
WgBce No Pret	Res&ST+DB	DB	69.90%	75.11%	3.41%	65.36%	98.37%	24
Bce R&S_Pret	Res&ST+DB	DB	74.07%	76.70%	2.20%	71.62%	98.35%	24
WgBce R&S_Pret	Res&ST+SBP	SBP	73.35%	78.66%	2.57%	68.71%	98.61%	22
WgBce R&S_Pret	Res&ST+DB	DB	74.86%	78.16%	2.22%	71.84%	98.45%	24

表7 损失函数、DBN模块以及连接方式在ZJU-Group 2上的消融研究

Table 7 Ablation study for Loss function, each module and hybrid connection on ZJU-Group 2

损失函数及预训练	编码阶段	解码阶段	\bar{f}	TPR	FPR	PPV	NPV	FPS (ms)
WgBce R&S_Pret	Res&ST+ASPP	DB	75.78%	80.54%	2.57%	71.55%	98.59%	27
WgBce R&S_Pret	Res&ST+ASPP	ASPP	75.14%	79.22%	2.55%	71.46%	98.49%	24
WgBce R&S_Pret	Res&ST+DB	ASPP	75.09%	79.15%	2.52%	71.43%	98.40%	25
WgBce S_Pret	ST+DB	DB	71.82%	74.92%	2.66%	68.96%	98.05%	22
WgBce R_Pret	Res+DB	DB	72.81%	78.52%	2.96%	67.88%	98.42%	18
WgBce R&S_Pret	Res&ST	DB	75.21%	77.66%	2.09%	72.91%	98.12%	23
WgBce No Pret	Res&ST+DB	DB	72.89%	77.54%	2.80%	68.77%	98.39%	24
Bce R&S_Pret	Res&ST+DB	DB	76.14%	80.14%	2.67%	72.52%	98.50%	24
WgBce R&S_Pret	Res&ST+SBP	SBP	75.29%	79.81%	2.69%	71.25%	98.53%	22
WgBce R&S_Pret	Res&ST+DB	DB	76.66%	79.34%	2.29%	74.16%	98.45%	24

一致性,进一步说明本文的算法模块具备泛化能力。

最终实验使用的损失函数是 $W_{\text{G}}\text{BCE Loss}$, 它包含两个超参数 σ 和 λ 。图9展示了两个参数的可视化对比。由图可见,固定参数 σ 和 λ 越小,中值附近的损失贡献越大,中值附近的函数曲线越凸出;固定参数 λ 和 σ 越大,端点值损失贡献的斜率越大; σ 越趋近于0,中值附近产生拐点; σ 取值0.2时,中值附近的函数曲线更平滑。

综合考量,选用的参数区间为 $\sigma \in [0.1, 0.2, 0.3]$, $\lambda \in [0.05, 0.1, 0.15, 0.2]$ 。表8展示了两个参数的消融实验,表中黑体字表示结果最优。由表可见,固定参数 σ 和 λ 取值0.1时, f_{measure} 取得最

大值,处于波峰;固定参数 λ, σ 取值0.2时, f_{measure} 取得最大值,处于波峰。所以,最终的 σ 和 λ 分别取值为0.2和0.1。

4 结论

针对复杂纹理织物瑕疵检测,本文提出一种双端输入型嵌套融合多尺度信息的网络架构 WNet。为了降低深层网络中背景特征对瑕疵特征的干扰,借助右骨干分支的全局建模能力,在卷积深层网络中补充全局上下文信息。为了减少深层网络中瑕疵细节信息丢失,提出一种轻量级双分支池化金字塔,将浅层多尺度细节特征引入深层模块。在特征融合阶段,为了进一步细化、筛

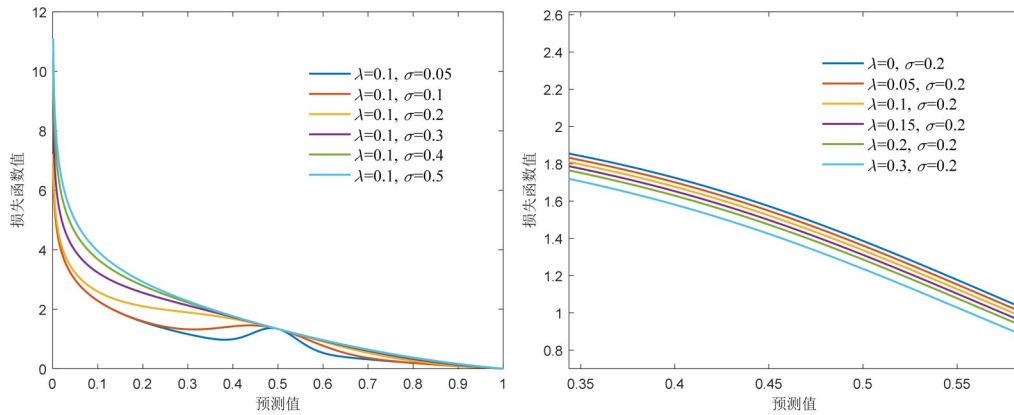


图 9 损失函数参数的可视化对比

Fig.9 Visual comparison of loss function parameters

表 8 ZJU-Group 4 上不同参数损失函数的消融实验

Table 8 Ablation study for different Loss Functions parameter on ZJU-Group 4

σ	λ	\bar{f}	TPR	FPR	PPV	NPV	MAE
0.1	0.05	71.13%	76.10%	2.69%	66.77%	98.45%	0.012
	0.1	74.50%	77.29%	2.16%	71.90%	98.39%	0.010
	0.15	73.86%	77.58%	2.28%	70.48%	98.42%	0.15
	0.2	73.80%	77.92%	2.42%	70.09%	98.46%	0.013
0.2	0.05	74.26%	78.83%	2.48%	70.18%	98.55%	0.015
	0.1	74.86%	78.16%	2.22%	71.84%	98.45%	0.014
	0.15	74.19%	77.46%	2.27%	71.18%	98.37%	0.012
	0.2	73.92%	79.28%	2.59%	69.24%	98.56%	0.013
0.3	0.05	73.10%	74.63%	2.03%	71.63%	98.32%	0.015
	0.1	73.73%	78.60%	2.62%	69.42%	98.60%	0.012
	0.15	73.73%	76.08%	2.11%	71.53%	98.30%	0.011
	0.2	73.74%	74.54%	2.01%	72.96%	98.25%	0.016

选、增强输入特征,甄别高梯度纹理特征和瑕疵特征,利用嵌套 DBPP 实现多尺度嵌套双分支模块,该模块更关注多尺度不同感受野范围内的特征信息,可以生成更加精确的多阶段瑕疵预测图像。最后,为了降低高梯度纹理区域的误检率,对二元交叉熵损失函数进行了改进,以提高预测置信度。提出的方法在四个瑕疵测试集上的综合评价指标较优,尤其是 f_{measure} 、阳性预测率和阴性预测率较高。但是,本文方法 WNet 的平均查准率依旧有较大的提升空间,同时,针对细小的织物瑕疵,检测效果依旧不够理想。

未来的研究将考虑改进特征融合方式,以更轻量高效的方式提高复杂纹理中瑕疵的检测精度。

参考文献

- [1] 马明寅,狄岚,梁久祯. 基于图像校正和模板分割的纺织品瑕疵检测. 南京大学学报(自然科学), 2021, 57(1): 29—41. (Ma M Y, Di L, Liang J Z. Fabric defect detection based on image correction and template segmentation. Journal of Nanjing University (Natural Science), 2021, 57(1): 29—41.)
- [2] 纪旋,梁久祯,侯振杰,等. 基于模板校正与低秩分解的纺织品瑕疵检测方法. 模式识别与人工智能, 2019, 32(3): 268—277. (Ji X, Liang J Z, Hou Z J, et al. Fabric defect detection based on template correction and low-rank decomposition. Pattern Recognition and Artificial Intelligence, 2019, 32(3): 268—277.)
- [3] Shi B S, Liang J Z, Di L, et al. Fabric defect

- detection via low-rank decomposition with gradient information and structured graph algorithm. *Information Sciences*, 2021(546):608—626.
- [4] 刘威,常兴治,梁久祯,等. 基于局部最优分析的纺织品瑕疵检测方法. *模式识别与人工智能*, 2018, 31(2): 182—189. (Liu W, Chang X Z, Liang J Z, et al. Fabric defect detection based on local optimum analysis. *Pattern Recognition and Artificial Intelligence*, 2018, 31(2):182—189.)
- [5] 龙涵彬,狄岚,梁久祯. 基于畸变校正与视觉显著特征的纺织品瑕疵检测. *模式识别与人工智能*, 2020, 33(12): 1122—1134. (Long H B, Di L, Liang J Z. Fabric defect detection based on distortion correction and visual salient features. *Pattern Recognition and Artificial Intelligence*, 2020, 33(12):1122—1134.)
- [6] Liu J H, Wang C Y, Su H, et al. Multistage GAN for fabric defect detection. *IEEE Transactions on Image Processing*, 2020(29):3388—3400.
- [7] Chen M Q, Yu L J, Zhi C, et al. Improved faster R-CNN for fabric defect detection based on Gabor filter with Genetic Algorithm optimization. *Computers in Industry*, 2022(134):103551.
- [8] Jing J F, Wang Z, Rätsch M, et al. Mobile-unet: An efficient convolutional neural network for fabric defect detection. *Textile Research Journal*, 2022, 92(1—2):30—42.
- [9] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 2014, 15(1):1929—1958.
- [10] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA: IEEE, 2016: 770—778.
- [11] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017:2261—2269.
- [12] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA: IEEE, 2015:3431—3440.
- [13] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834—848.
- [14] Chen L C, Zhu Y K, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation//*European Conference on Computer Vision*. Springer Berlin Heidelberg, 2018:833—851.
- [15] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach, CA, USA: ACM, 2017: 6000—6010.
- [16] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: Transformers for image recognition at scale//*The 9th International Conference on Learning Representations*. Vienna, Austria: ICLR, 2021:435—457.
- [17] Liu Z, Lin Y T, Cao Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows//2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE, 2021:9992—10002.
- [18] Tian X, Xu K, Yang X, et al. Bi-directional object-context prioritization learning for saliency ranking//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, LA, USA: IEEE, 2022:5872—5881.
- [19] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017: 936—944.
- [20] Hou Q B, Cheng M M, Hu X W, et al. Deeply supervised salient object detection with short connections//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA: IEEE, 2017:5300—5309.
- [21] Zhao H S, Shi J P, Qi X J, et al. Pyramid scene parsing network//*Proceedings of the IEEE Conference on Computer Vision and Pattern*

- Recognition. Honolulu, HI, USA: IEEE, 2017: 6230—6239.
- [22] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 2999—3007.
- [23] Chen Z X, Zhou H J, Lai J H, et al. Contour-aware loss: Boundary-aware learning for salient object segmentation. IEEE Transactions on Image Processing, 2021(30): 431—443.
- [24] Wei J, Wang S H, Huang Q M. F³Net: Fusion, feedback and focus for salient object detection//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2020: 12321—12328.
- [25] Ke Y Y, Takahiro T. Recursive contour-saliency blending network for accurate salient object detection//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, HI, USA: IEEE, 2022: 1360—1370.
- [26] Tu W C, Liu M Y, Jampani V, et al. Learning superpixels with segmentation-aware affinity loss//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 568—576.
- [27] Jiang W Q, Wu Y N, Qiao F S, et al. Model level ensemble for facial action unit recognition at the 3rd ABAW challenge//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New Orleans, LA, USA: IEEE, 2022: 2336—2343.
- [28] Xie S N, Tu Z W. Holistically-nested edge detection//2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 1359—1403.
- [29] Qin X B, Zhang Z C, Huang C Y, et al. U²-Net: Going deeper with nested U-structure for salient object detection. Pattern Recognition, 2020(106): 107404.
- [30] Zhang C K, Feng S Z, Wang X L Q, et al. ZJU-Leaper: A benchmark dataset for fabric defect detection and a comparative study. IEEE Transactions on Artificial Intelligence, 2020, 1(3): 219—232.
- [31] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA: IEEE, 2009: 1597—1604.
- [32] Liu N, Han J W, Yang M S. PiCANet: Learning pixel-wise contextual attention for saliency detection//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 3089—3098.
- [33] Chen Z Y, Xu Q Q, Cong R M, et al. Global context-aware progressive aggregation network for salient object detection//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2020: 10599—10606.
- [34] Yu S Y, Zhang B F, Xiao J M, et al. Structure-consistent weakly supervised salient object detection with local saliency coherence//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2021: 3234—3242.
- [35] Hussain T, Anwar S, Ullah A, et al. Densely deformable efficient salient object detection network. 2021, arXiv:2102.06407.

(责任编辑 杨可盛)