

文章编号 1004-924X(2023)10-1563-17

基于标签嵌入方法的纺织品瑕疵识别网络

刘 颖*, 姜 威, 李冠典, 陈 磊, 赵 爽

(长春理工大学 电子信息工程学院, 吉林 长春 130000)

摘要:卷积神经网络(Convolutional Neural Network, CNN)可用于工业生产环境下的纺织品疵点的鉴别与分类。针对实际场景下的纺织品瑕疵存在瑕疵类型视觉区分度小和实际数据样本采集时的瑕疵类别不平衡问题,本文提出了基于标签嵌入方法的纺织品瑕疵识别网络(Textile Defect Recognition Network Based on Label Embedding, TDRNet)。首先,算法调整了基础骨干网络的结构,从而提高模型的分类精度;接着算法还设计了标签嵌入模块(Label Embedded Module, LEM),并使用该模块来生成模型的类别权重偏移;然后,本文提出了分布感知损失函数(Distribution Perception Loss, DP Loss)调整算法的类别分布,从而减小同类瑕疵特征的类内距并增大异类瑕疵特征的类间距;最后,本文引入了Seesaw Loss损失函数,通过抑制少数类别的负样本梯度并提高对误分类时的样本损失来动态平衡模型训练过程中在不同样本下的更新梯度,以缓解少数类别的误分类率。在自制的“广东智能制造”布匹瑕疵分类数据集中,本文提出的框架在粗粒度分类和细粒度分类两个任务上的top1错误率可达16.35%和17.12%,而top5错误率在细粒度分类任务上低至5.20%。与其他分类模型相比,TDRNet在对比实验中取得了最优的结果。此外,TDRNet与近5年经典的细粒度分类模型进行了比较,并取得了SOTA结果,这充分表明了TDRNet的先进性。

关键词:卷积神经网络;纺织品瑕疵识别;标签嵌入;分布感知损失;Seesaw损失

中图分类号:TP391 **文献标识码:**A **doi:**10.37188/OPE.20233110.1563

Textile defect recognition network based on label embedding

LIU Ying*, JIANG Wei, LI Guandian, CHEN Lei, ZHAO Shuang

(College of Electronic Information Engineering, Changchun University of Science and Technology,
Changchun 130000, China)

* Corresponding author, E-mail: liuying02@cust.edu.cn

Abstract: A convolutional neural network (CNN) can be used in the industrial production environment to identify and classify textile defects. To overcome the problems in the visual discrimination of small defect types and imbalance of textile defect categories in actual scenes, a textile defect recognition network (TDRNet) based on label embedding method is proposed. First, the backbone structure is adjusted to improve the classification accuracy of the model. Then, a label embedded module (LEM) is constructed to generate the category weight offset of the model. Subsequently, a distribution perception loss function (DP loss) is proposed to adjust the class distribution of the algorithm; this reduces the distance of homogeneous defect features and increases the distance of heterogeneous features. Finally, the seesaw loss function is introduced to dynamically balance the gradient update for different samples during the model training process by suppressing the negative sample gradient of a few categories and increasing the sample loss during

收稿日期:2022-06-10;修订日期:2022-08-10.

基金项目:吉林省自然科学基金资助项目(No. 20210101182JC)

misclassification, thereby alleviating the misclassification rate of a few categories. In the self-made "Guangdong intelligent manufacturing" cloth defect classification dataset, the top1 error rate of our framework for rough-grained and fine-grained classifications reached 16.35% and 17.12%, respectively, whereas the top5 error rate of fine-grained classification was as low as 5.20%. Compared with other classification models, TDRNet achieved the best results. In addition, TDRNet was compared with the classical fine-grained classification model in recent five years and achieved state-of-the-art (SOTA) performance, fully demonstrating the enhancements provided.

Key words: convolutional neural network; textile defect recognition; label embedding; distribution perception loss; Seesaw Loss

1 引 言

在纺织品生产过程中,由于设备运行不稳定、人工操作失误等原因,纺织品不可避免地会出现黄渍、毛斑、弓纱、吊经、结洞等瑕疵,不利于后续纺织品的生产效率的提升和质量评级。因此,纺织品瑕疵识别是纺织品生产的必要环节。

随着计算机技术的迅猛发展,以计算机视觉技术为核心的图像缺陷检测技术在纺织工业中得到广泛应用。传统的纺织品瑕疵识别方法大致可分为两类,即基于纺织品纹理的识别方法和基于图像频谱分析的识别方法。基于纹理的瑕疵识别方法依据纺织品的纹理结构是否遭受破坏来检测纺织品瑕疵^[1]。Chen等采用阈值方法分割纺织品图像并从中提取出布匹纹理特征描述子,在该局部描述子基础上,使用其位置和长度直方图参数来定义可用于纺织品缺陷检测的统计测量方法^[2];Cohen等使用了马尔科夫随机场(Markov Random Field, MRF)的方法,通过比较分析常规图像和瑕疵图像尺度和方向上的MRF参数值来检测瑕疵,然而该方法在小样本下缺乏普适性^[3]。基于频谱分析的方法则利用Gabor滤波器(Gabor Filters, GF)等数学工具来获取纺织品图像的频域特征,并通过比较频域特征来检测瑕疵^[4]。Chiu等人将图像变换到频域,并使用极大似然估计分割纺织品瑕疵,但该方法在小目标任务上难以胜任^[5];Jing等人则使用Gabor滤波器组来计算瑕疵特征,并使用核主成分分析的方法减少数据冗余,但方法的计算量较大^[5];而Liang等人则结合了小波变换和模糊神经网络的优势,实现了纺织品纱线的质量

监测^[6]。

传统的图像识别算法实现复杂,最终的检测精度严重依赖于提取的特征集合。近年来,深度学习技术在强大硬件性能支持下崛起,并在图像处理的各个应用领域中表现出巨大的优势^[8],如缺陷检测^[9-10]、遥感监测^[11-13]、医学影像分析^[14-15]等。在图像分类任务中,基于CNN的模型将特征提取和样本分类统一为一个端到端的整体,且模型的泛化能力远高于传统分类模型^[16]。2012年,Krizhevsky等人提出 AlexNet,在当年图像分类领域取得了SOTA的结果^[17];2016年,He等人提出的 ResNet在CNN中引入了残差连接,极大地缓解了模型训练过程中的网络退化问题^[18];2018年,Hu等人将注意力机制引入到CNN架构中,提高了网络模型的有效容量,进而提高了模型的性能^[19]。在纺织品瑕疵识别领域,基于CNN的方法也大放异彩。Ouyang等在CNN的基础上嵌入了成对电位激活层(Pairwise-potential Activation Layer, PPAL),从而改善了复杂特征的疵点检测的准确性^[20];陆贵家基于特征金字塔网络,在 Cascade R-CNN中嵌入花色布匹特征融合模块,以解决复杂纹理下瑕疵识别的准确性^[21];Peng等人对瑕疵类别的先验锚框进行改进,提出了基于瑕疵真值的稀疏先验锚卷积网络,以更准确和有效地定位极端疵点^[22]。

虽然上述基于深度学习的方法在纺织品瑕疵识别任务中的表现比传统图像处理方法更为优秀,但对于纺织品样本存在的瑕疵类别难以区分和类别不平衡等挑战性难题,仍未给出有效的解决方案。有鉴于此,本文提出了基于标签嵌入方法的纺织品瑕疵识别网络(Textile Defect Rec-

ognition Network Based on Label Embedding, TDRNet)。首先,为保证模型具备工业实用价值,TDRNet对图像进行预处理,并在不降低检测精度的基础上调整了网络结构;接着,本文引入了标签嵌入学习的方法,在训练时引入标签信息来引导不同类别的特征在类别特征空间中的分布;为了提高难分样本的分类准确率,本文设计了分布感知损失 DP Loss 以约束类别特征的内距和类间距;此外,为有效地缓解类别不平衡带来的模型“走捷径”问题,传统的交叉熵损失被替换为 Seesaw Loss,以在网络训练过程中动态地降低负样本梯度信息并增加对样本错误分类时的损失。在基于“广东智能制造”布匹瑕疵数据集基础上创建的纺织品分类数据集的两个分类任务上进行的一系列实验结果表明,本文提出的方法在纺织品瑕疵识别任务是有效的,且比起主流的 CNN 模型,TDRNet 取得了更为优越的识别效果。

本文的贡献总结如下:

(1) 针对布匹瑕疵图像照幅大分辨率高的问题,本文改进了骨干网络的结构,提出了改进型 ResNet;

(2) 针对布匹瑕疵存在难分类别的问题,本文根据 NLP 领域中的标签嵌入思想,设计了标签嵌入模块,用该模块生成当前网络的类别权重偏移,并利用该偏移优化 softmax 分类器权重参数;

(3) 为了使得标签嵌入模块学习到正确的类别权重偏移,本文提出了分布感知损失 DP Loss,以使网络提取到更为强大的分类特征,并减少分类器参数的类内距和增大类间距;

(4) 由于数据集存在类别不平衡的问题,因此 TDRNet 引入了 Seesaw Loss 来缓解过多的负样本梯度信息和误分类损失对网络优化的影响,从而使模型往正确的方向更新梯度。

2 CNN 图像分类模型及其损失函数

2.1 残差 CNN 分类模型

残差单元 (Residual Unit)^[18] 是 CNN 模型必

不可少的组件。该组件可以防止参数学习过程中模型出现退化问题。残差单元与普通级联卷积块的区别在于末端 ReLU 输出前增加了从输入到输出的残差连接,如图 1 所示。

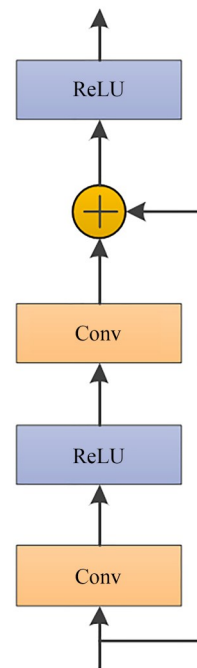


图 1 残差单元示意图

Fig. 1 Structure of residual unit

由于残差连接提高了更新梯度信息的传播效率,且残差单元参数空间更为平滑,这使得带残差单元的 CNN 模型更容易学习到有效特征,因此与早期的 CNN 架构相比,包含残差单元的 CNN 模型具备更强的分类性能。

2.2 残差 CNN 的分类与训练

CNN 可以通过添加 SVM (Support Vector Machine), sigmoid 或 softmax 等分类器实现图像分类。在训练图像分类 CNN 模型时,一般使用小批量随机梯度下降算法 (Mini-batch Stochastic Gradient Descent, Mini-batch SGD) 或对应改进方法来更新模型参数。在训练集 $D = \{(\mathbf{x}^{(n)}, y^{(n)})\}_{n=1}^N$ 上,若使用损失函数 $L(y, \hat{y})$ 对残差 CNN 进行优化,则结构化风险函数为:

$$R(\mathbf{W}, \mathbf{b}) = \frac{1}{N} \sum_{n=1}^N L(y^{(n)}, \hat{y}^{(n)}) + \frac{1}{2} \lambda \|\mathbf{W}\|_F^2, \quad (1)$$

其中: $\mathbf{W}$ 和 $\mathbf{b}$ 为网络权重和偏置, $L(y^{(n)}, \hat{y}^{(n)})$ 为损

失函数,若使用 softmax 分类器作为网络的分类头,则 $L(y^{(n)}, \hat{y}^{(n)})$ 选用交叉熵损失函数。 $\|W\|_F^2$ 为权重矩阵 W 的 Frobenius 范数,用来惩罚网络的权重参数以防止网络训练过程中出现过拟合。 $\lambda > 0$ 为超参数,其越大则 W 越接近于 0。

3 基于标签嵌入的纺织品瑕疵识别算法

3.1 网络架构

本文以 CNN 架构为基础,针对纺织品瑕疵识别任务存在的挑战性设计 TDRNet。其训练架构如图 2 所示。

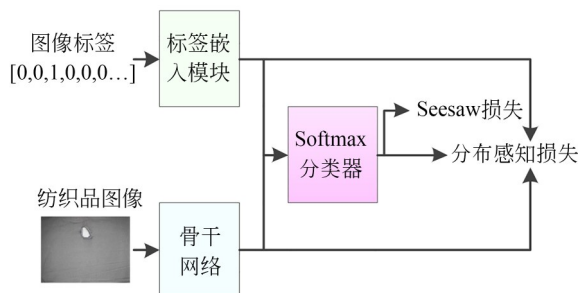


图 2 TDRNet 网络架构示意图

Fig. 2 Overall of TDRNet architecture when training

TDRNet 框架由骨干网络、标签嵌入模块和 softmax 分类器组成。其中,纺织品图像输入骨干网络,以提取纺织品图像的语义特征;图像标签则输入标签嵌入模块,以学习图像标签与图像语义信息的距离,进而减小同类瑕疵样本间的类内距和增加异类瑕疵的类间距,来增加网络对瑕疵的区分度。在该框架中,骨干网络的输出作为 softmax 分类器的输入特征,标签嵌入模块的输出作为分类器的类别权重偏移,softmax 分类器则通过度量分类器权重与输出语义特征的距离,从而实现瑕疵的分类。

在训练阶段,TDRNet 使用两个损失函数,分布感知损失 DP Loss 和 Seesaw Loss 进行优化,其中,DP Loss 用以学习当前类别分布相对理想类别分布的偏移,Seesaw Loss 则用来学习瑕疵的类别特征。需要指出的是,在测试阶段,TDRNet 将移除标签嵌入模块,直接使用骨干网络进行推理,如图 3 所示。

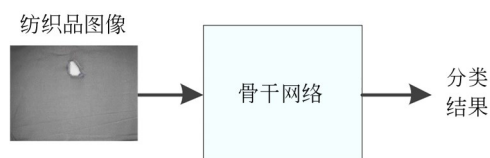


图 3 TDRNet推理架构示意图

Fig. 3 Overall of TDRNet Architecture When Testing

3.2 网络模型

3.2.1 骨干网络

残差单元的假设函数可表示为:

$$f_L(\theta, x) = h_L(\theta, x) - x, \quad (2)$$

其中: $f_L(\theta, x)$ 为残差目标函数, $h_L(\theta, x)$ 为目标函数, x 为残差单元的输入特征。由公式(2)可知,残差单元可以用 $(f_L(\theta, x) + x)$ 去逼近 $h_L(\theta, x)$,根据第 2 章内容可知,由残差单元堆叠而成的 ResNet 参数空间的流形更为光滑,这使得整个网络更为容易学习。因此,本文提出的 TDRNet 使用 ResNet 作为骨干网络架构。

但 ResNet 并不适合直接作为 TDRNet 的骨干网络。原始的 ResNet 的图像输入尺寸为 224×224 ,而工业环境中采集到的纺织品样本为 2560×1920 的 4K 图像。直接使用大尺度图像作为网络输入会使得网络显存占用大幅增加,训练速度降低,进而增加网络训练难度。而且受限网络的结构,直接使用大尺度图像训练模型还会带来模型感受野不足,无法充分提取图像全局特征的问题。为此,有必要对 ResNet 进行改进,使其适配纺织品瑕疵检测任务。TDRNet 的骨干网络结构如表 1 所示。

TDRNet 对原始的 ResNet-50 和 ResNet-101 的结构做了如下改进:首先, 2560×1920 的原始数据集被裁剪成 720×720 数据集,并将该尺寸作为 TDRNet 骨干网络的输入尺寸,在保证分类准确度的情况下降低了模型的计算开销;接着,将输入 7×7 卷积核换为 3 个 3×3 的卷积核,以获取更强的语义特征;然后,骨干网络在保持每一阶段的卷积步长(Stride)为 2 的前提下,在 Conv4 阶段后将卷积核的空洞率(Dilated Rate)调整为 2,以增强网络深层输出的特征图的全局语义;最后,TDRNet 的骨干网络在 Conv4 阶段之后添加了 Conv5 阶段,以进一步聚合图像的高级语义特征。

表 1 TDRNet骨干网络结构
Tab. 1 Backbone of TDRNet

Layer name	Output size	Stride	Dilated rate	50-layer	101-laer
Input Layer	720×720	— —	— —	— —	— —
Conv0	360×360	2	1	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \\ 3 \times 3, 128 \end{bmatrix}$ $3 \times 3, \text{max pool}$	
Conv1_x	180×180	2	1	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv2_x	90×90	2	2	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Conv3_x	45×45	2	2	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1\,024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1\,024 \end{bmatrix} \times 23$
Conv4_x	23×23	2	2	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2\,048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2\,048 \end{bmatrix} \times 3$
Conv5	12×12	2	2	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2\,048 \end{bmatrix} \times 1$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2\,048 \end{bmatrix} \times 1$
— —	1×1	— —	— —	Global Average Pooling, FC, softmax	

3. 2. 2 标签嵌入模块

由于纺织品的某些瑕疵类别相似度高,视觉上难以区分其类别,直接使用CNN模型进行瑕疵分类会降低其分类准确率。图4为不同类别的瑕疵图像。图4(a)展示的松经样例中经纱在交织点部分有非正常拱起;图4(b)断经是由于经纱断裂造成的经纱距离出现异状变大;而图4(c)吊经则展示了单根或数根经纱吊紧而纬纱弯曲的现象,图4(d)纬缩则因织造过程中多余的纬纱卷入造成。对比4个瑕疵样例可知,这些瑕疵外观难以分辨,若非熟练工人,且不借助触摸、显微放大等手段,很难将其区分开。在训练瑕疵识别网络时,类似的难分瑕疵种类会干扰网络模型的参数学习,使得网络收敛速度变慢,甚至学习失败。因此,从某种程度上讲,纺织品瑕疵识别是细粒度分类问题。

针对上述问题,本文构建了标签嵌入模块(Label Embeded Module, LEM),在假设标签语义互不相似,且互不相似程度等同的基础上,

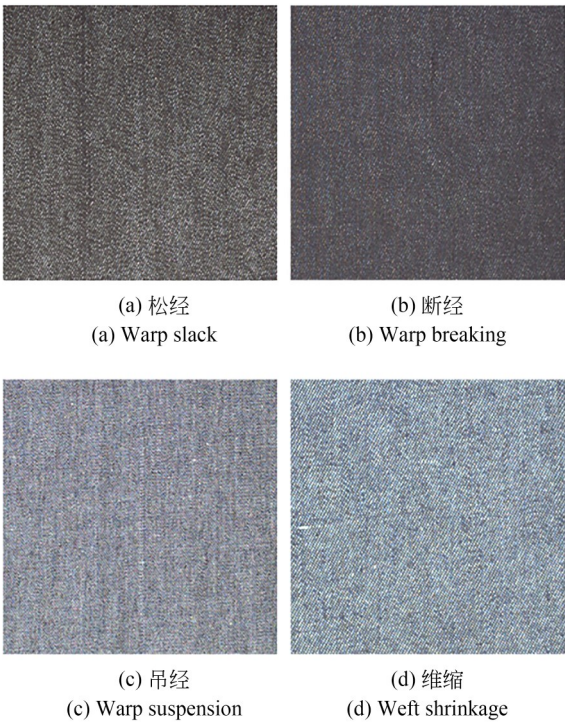


图 4 纺织品瑕疵类别样例
Fig. 4 Example of textile defect category

LEM将各个类别的标签映射为独热码并作为标签嵌入模块的输入,且该输入可在网络的训练中隐式地调整其分布;并逐渐学习到真实的标签语义输出。在每一次训练的前向传播阶段中,该模块将逐步生成类别权重偏移,接着该偏移量用以调整骨干网络和 softmax 分类器的参数,从而扩大难分类样本特征的余弦距离。标签嵌入模块

LEM的结构如图 5(a)所示。

标签嵌入模块主体为一个 4 层的小型全连接前馈神经网络。前三层的全连接层的结构如图 5(b)所示,由线性层、BN 层和 ReLU 激活层组成, C 为类别数,表示全连接层的宽度;第四层结构如图 5(c)所示,由线性层和 Tanh 激活层组成。

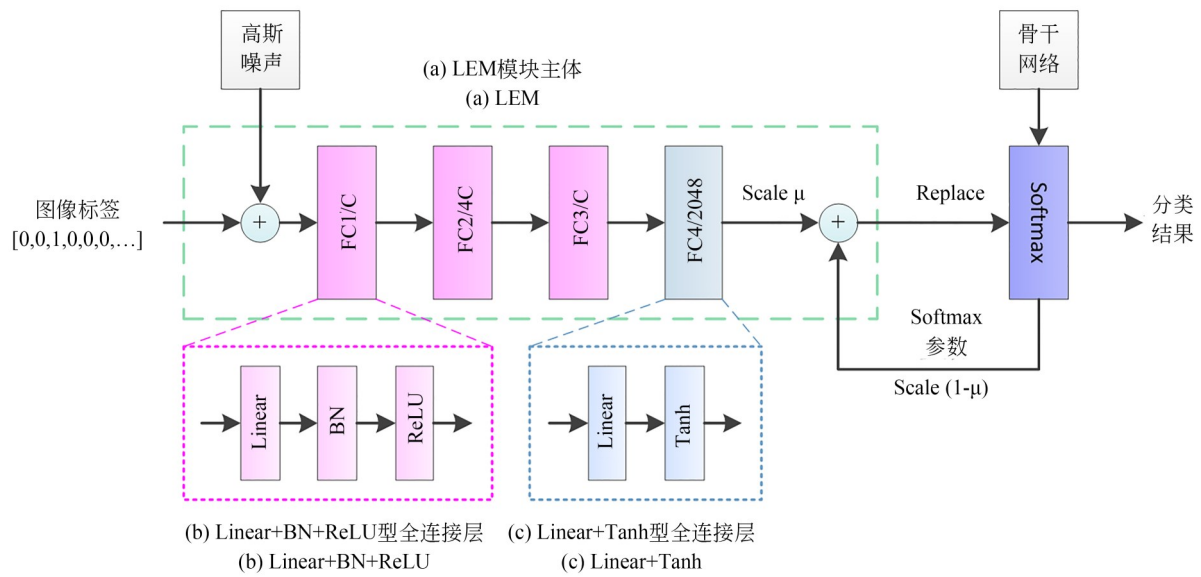


图5 LEM模块结构示意图

Fig. 5 Structure of Label Embedded Module

该网络将图像类别标签转换为独热标签向量后将其作为模块的输入,由于直接将独热向量输入全连接层会导致输入神经元失活,相当于添加了带类别偏好的 dropout,这为 LEM 引入不必要的类别先验,从而导致模块在训练初期便过拟合,因此 LEM 模块的输入额外引入了 $X \sim N(0, 0.05)$ 的高斯噪声;接着,带高斯噪声的标签独热向量经过 4 层全连接层后,经 Tanh 激活函数输出对称的类别权重偏移,其维度大小为 $\text{batch_size} \times 2048 \times 1$;然后,偏移量与 softmax 分类器的参数使用指数移动平均法进行融合,且该融合过程发生在训练过程中每一批次的前向传播阶段。需要注意的是,该结果只与对应的输入标签相匹配的类别权重相加。最后,输出的结果作为新的 softmax 分类器参数。TDRNet 的分类器权重将得到更新过程可表示为:

$$W_{\text{softmax}}^{\text{iter}} = (1 - \mu) \cdot W_{\text{softmax}}^{\text{iter}-1} + \mu \cdot X_{\text{LEM}}^{\text{iter}}, \quad (3)$$

其中: W_{softmax} 表示 softmax 分类器权重参数, iter 表示网络的第 iter 次训练迭代, X_{LEM} 为 LEM 模块输出, μ 为缩放系数。即经过 μ 缩放后的 X_{LEM} 输出将与 softmax 分类器中对应经过 $(1 - \mu)$ 缩放的权重参数 W_{softmax} 相加,然后再对原参数进行迭代。在 LEM 模块中引入缩放系数可以在迭代过程中平滑模型权重,以防止过大的偏移量导致模型参数更新不稳定甚至学习失败,同时,缩放系数的大小还控制着网络模型遗忘历史分类权重的速度。

在 TDRNet 中,骨干网络将学习纺织品图像的语义特征向量,而 LEM 模块将学习训练样本标签与理想特征语义之间潜在的对应关系,这让 softmax 分类器得以显式地遵循“减小类内距——增大类间距”的范式调整网络在训练时不同瑕疵类别在类别概率空间中的分布,进而利用调整后的类别分布来约束网络末端的梯度更新方

向,从而得到类内紧凑、类间分明的图像语义特征。需要指出的是,LEM模块作为TDRNet训练阶段的即插即用模块,在推理阶段,该模块可以移除,这不影响最终的分类结果,且可以提高推断效率。

3.3 损失函数

3.3.1 分布感知损失

若仅用交叉熵损失约束分类器,则LEM模块将蜕化为分类器权重参数动态梯度信息预测器,这是不可接受的。嵌入TDRNet的LEM模块需要学习如何生成当前语义向量相对理想语义向量的偏移。因此,本文提出了分布感知损失(Distribution Perception Loss, DP Loss)以确保LEM模块学习到正确的语义偏移范式。本文以余弦距离作为特征距离度量,并通过“减小类内距——增大类间距”的方法提高难分类别的区分度,则DP Loss被定义为:

$$L_{DP} = L_{intra_dis} + L_{inter_dis}, \quad (4)$$

其中: L_{DP} 由两部分组成。 L_{intra_dis} 和 L_{inter_dis} 分别表示类内损失和类间损失。DP Loss将骨干网络的倒数第二层输出、LEM模块输出和分类器权重作为优化项,对于 L_{intra_dis} ,有:

$$L_{intra_dis} = \sum_{i=1}^C y_i (1 - \cos \langle X_{backbone}, X_{LEM} \rangle), \quad (5)$$

$y_i \in \{0, 1\}$ 是独热码标签,当当前类与当前标签相匹配时, $y_i = 1$,否则 $y_i = 0$ 。式(5)中 $X_{backbone}$ 和 X_{LEM} 分别表示骨干网络的倒数第二层的语义特征输出与LEM模块输出的类别权重偏移, C 为类别数。而对于 L_{inter_dis} ,本文将定义为:

$$L_{inter_dis} = \sum_{i=1}^C y_i \left(1 + \cos \left\langle \frac{(1+t)W_{mean} - tW^{(i)}}{\mu X_{LEM} + (1-\mu)W^{(i)}}, \right\rangle \right), \quad (6)$$

其中: W_{mean} 表示softmax分类器的各个类别的分类参数向量和, $W^{(i)}$ 表示softmax分类器在类 i 下的分类参数, $t = 1/(C-1)$,为不包含背景类别的瑕疵类别数的倒数, μ 为上文所提缩放系数。

分布感知损失的作用机制如图6所示。根据式(5),类内损失 L_{intra_dis} 的优化目标为 $X_{backbone}$ 和 X_{LEM} 的余弦相似度。因为在理想状态下, $\cos \langle X_{backbone}, X_{LEM} \rangle \in [-1, 1]$,所以:

$$\lim_{iter \rightarrow +\infty} (1 - \cos \langle X_{backbone}, X_{LEM} \rangle) = 0. \quad (7)$$

根据公式(7),当 $X_{backbone}$ 与 X_{LEM} 相似度越高,则类内损失越小,因为 X_{LEM} 由LEM模块生成,而LEM在输入带噪的独热标签的情形下,期望输出固定的分布偏移量。因此, L_{intra_dis} 将使得骨干网络的倒数第二层的输出语义 $X_{backbone}$ 向 X_{LEM} 收敛,如图6(a)所示,随着网络的不断训练,原本同

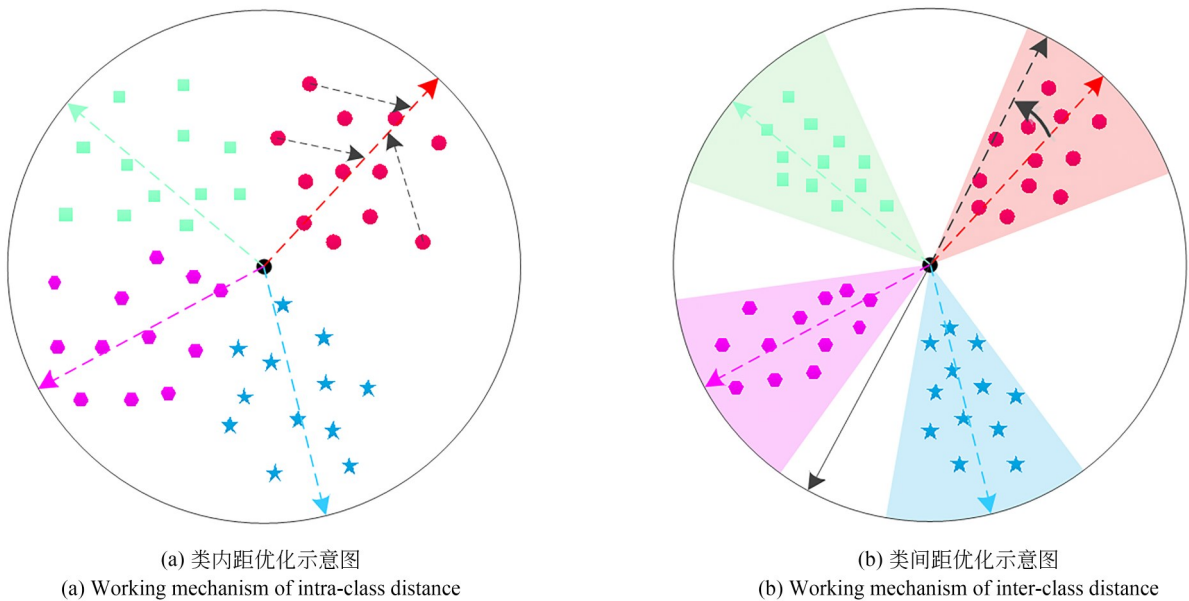


图6 DP Loss原理示意图

Fig. 6 principle of DP loss

类瑕疵样本在类别特征空间中松散的特征语义将向 X_{LEM} 偏移,从而显著减小了类内距。类似地,对类间损失,优化目标应为增加当前类与其

他所有类的距离,设第 i 类类心为 c_i ,并用类心距表征类间距,则当前类 c_i 与其他所有类的平均余弦距离可表示为:

$$D_{\text{inter}} = \sum_{i=1}^c y_i \frac{1}{C-1} \sum_{j=1, i \neq j}^C \cos \langle c_j, c_i \rangle = \sum_{i=1}^c y_i \frac{1}{C-1} \sum_{j=1, i \neq j}^C c_j c_i^T = \sum_{i=1}^c y_i \left(\frac{1}{C-1} \sum_{j=1, i \neq j}^C c_j \right) c_i^T = \sum_{i=1}^c y_i \cos \left\langle \frac{1}{C-1} \sum_{j=1, i \neq j}^C c_j, c_i \right\rangle, \quad (8)$$

由公式(8)可知,求当前类与其他类的平均余弦距离,等价于求当前类类心与其他类类心均

值的余弦距离,在 TDRNet 中, $c_i = \mu X_{LEM} + (1 - \mu) W^{(i)}$, 而 $c_j = W^{(j)}$, $i \neq j$, 因此:

$$\begin{aligned} D_{\text{inter}} &= \sum_{i=1}^c y_i \cos \left\langle \frac{1}{C-1} \sum_{j=1, i \neq j}^C W^{(j)}, \mu X_{LEM} + (1 - \mu) W^{(i)} \right\rangle = \\ &= \sum_{i=1}^c y_i \cos \left\langle \frac{W_{\text{sum}}^{(j)} - W^{(i)}}{C-1}, \mu X_{LEM} + (1 - \mu) W^{(i)} \right\rangle = \\ &= \sum_{i=1}^c y_i \cos \left\langle \frac{C}{C-1} W_{\text{mean}} - \frac{W^{(i)}}{C-1}, \mu X_{LEM} + (1 - \mu) W^{(i)} \right\rangle = \\ &= \sum_{i=1}^c y_i \cos \langle (1 + t) W_{\text{mean}} - t W^{(i)}, \mu X_{LEM} + (1 - \mu) W^{(i)} \rangle, \end{aligned} \quad (9)$$

其中, $t = \frac{1}{C-1}$ 。由于 $D_{\text{inter}} \in [-1, 1]$, 对于基于类间平均距离的类间损失,希望:

$$\lim_{\text{iter} \rightarrow \infty} D_{\text{inter}} = - \sum_{i=1}^c y_i. \quad (10)$$

根据公式(10),可构建类间损失为:

$$L_{\text{inter_dis}} = \sum_{i=1}^c y_i (1 + D_{\text{inter}}). \quad (11)$$

即公式(6)。 $L_{\text{inter_dis}}$ 的优化项为嵌入 LEM 模块的后的 softmax 分类器权重,随着训练次数的增加,该类间损失将逐渐降低,当前类类心与其

他类类心的距离也逐渐变大,如图 6(b)所示。

使用 DP Loss 后,当网络收敛时, TDRNet 将使骨干网络的各个瑕疵类别的输出语义特征朝 X_{LEM} 趋近,而分类器的权重参数将更替为参数权重与 X_{LEM} 历史数据的指数移动平滑结果,且该结果使得类心距最大化。

3.3.2 Seesaw Loss

工业环境下采集到的纺织品瑕疵数据集存在长尾分布问题。本文构建的 GDIM-CD 数据集同样如是。图 7 展示了 GDIM-CD 数据集在粗、

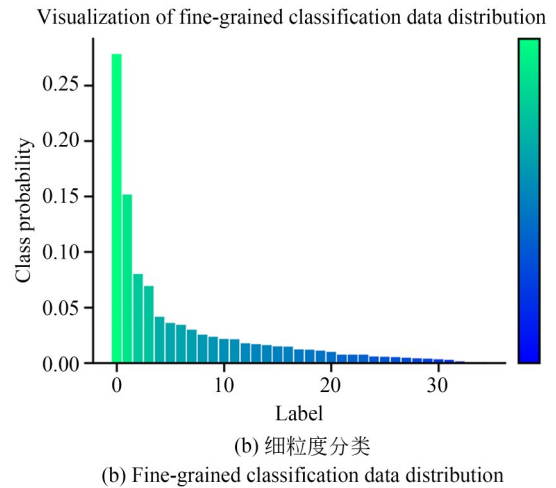
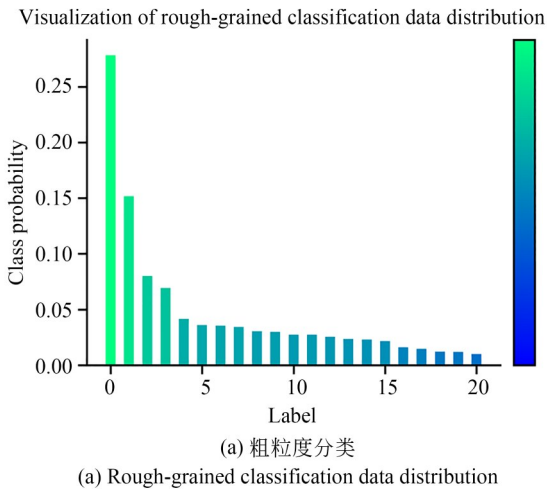


图 7 GDIM-CD 数据集数据分布

Fig. 7 the dataset label distribution of Guangdong intelligent manufacturing category dataset

细两个粒度两个分类任务上的样本分布情况。根据图 7(a),在粗粒度分类任务的数据中,稀有样本类别的数据量远比大量样本类别要少,而图 7(b)表明这一点更为明显。

数据集的类偏斜特性会为网络训练增加难度。这是因为来自多数类别的样本会在训练过程中对少数类别的样本施加过量负样本梯度,进而抑制了少数类别的 loss 响应,从而导致网络不能均衡地按类别更新梯度信息。这意味着使用常规的交叉熵损失在类偏斜数据集上不会取得好的效果,因为越是常见的类别,则其分类的准确率就越高;反之,越是稀有的类别,则其分类的准确率就越低。这可解释为大量的常见样本的梯度信息掩蔽了稀有类别的梯度信息。因此,GDIM-CD 的类偏斜若不得到有效处理,将会使模型在训练后倾向于将样本分类为常见瑕疵,这将导致模型缺乏对稀有类别的辨别能力。针对上述问题,TDRNet 将常规的交叉熵替换为 Seesaw Loss,通过借助其动态调节正负样本梯度比例的特性以缓解类偏斜带来的模型精度损失问题。

Seesaw Loss 在常规的交叉熵损失的基础上,使用调节因子对正负样本梯度进行平衡。在网络训练过程中,Seesaw Loss 统计施加在当前类别上正负样本的梯度值和样本数量,并利用样本数量作为权重来调整正负样本的梯度值,从而保证正负样本的动态平衡。Seesaw Loss 的数学表达式可写为:

$$\begin{cases} L_{\text{seesaw}}(z) = -\sum_{i=1}^C y_i \log(\hat{\sigma}_i) \\ \hat{\sigma}_i = \frac{e^{z_i}}{\sum_{j \neq 1}^C S_{ij} e^{z_j} + e^{z_i}} \end{cases}, \quad (12)$$

其中: $y_i \in \{0, 1\}$ 是独热码标签, $z = [z_1, z_2, \dots, z_C]$ 是每一个类别预测的 logit, S_{ij} 是平衡系数,此时,对于一个第 i 类的样本,它施加在第 j 类上的负样本梯度为:

$$\frac{\partial L_{\text{seesaw}}(z)}{\partial z_j} = S_{ij} \frac{e^{z_j}}{e^{z_i}} \hat{\sigma}_i. \quad (13)$$

由式(6)可知,通过调节 S_{ij} , Seesaw Loss 可以达到放大或者缩小第 i 类样本施加在第 j 类上的负样本梯度的效果,因此, Seesaw Loss 可以通过调节平衡系数 S_{ij} 来达到平衡正负样本梯度的

目的。为了避免因引入平衡系数而破坏交叉熵损失函数的原本优良性质,平衡系数的设计除了考虑减少过量的尾部类别的负样本梯度外,还考虑误分类时要避免盲目减少负样本梯度。因此, S_{ij} 可表示为:

$$\begin{cases} S_{ij} = M_{ij} \cdot C_{ij} \\ M_{ij} = \min \left\{ 1, \left(\frac{N_j}{N_i} \right)^p \right\} \\ C_{ij} = \max \left\{ 1, \left(\frac{\sigma_j}{\sigma_i} \right)^q \right\} \end{cases}, \quad (14)$$

其中: M_{ij} 为用来缓解尾部类别上过量的负样本梯度,而 C_{ij} 用来对误分类样本减少的梯度进行补充。 N 为训练过程中的在线累积动态数量, σ 为样本的分类置信度, p 和 q 分别为可调节的超参数用以在训练过程中控制梯度更新的大小。

4 实验与结果分析

4.1 数据集与评价指标

为了验证 TDRNet 的性能,本文以“广东智能制造纺织品瑕疵检测数据集”^[23]为基础,制作了广东智能制造分类数据集(Guangdong Intelligent Manufacturing Category Dataset, GDIM-CD),该数据集源图像采集自从佛山南海纺织车间现场。从 2 560×1 920 源数据集生成本文所用数据集的步骤如下:

- (1) 对无瑕疵图像,将图像划分为 2×3 图像块,得到 6 幅 854×960 的图像,在该基础上,在不超出图像边界的前提下对每个图像块进行随机位置裁剪,生成 720×720 的无瑕疵纺织品图像;
- (2) 对有瑕疵图像,根据数据集的标签确认瑕疵的位置和大小,在不超过图像边界、范围内没有其他瑕疵、瑕疵的宽或高小于 720 像素尺寸的前提下,尽量生成 720×720 的图像,同时进行类似步骤 1 中的随机生成;

(3) 对有瑕疵图像,在不超过图像边界、范围内没有其他瑕疵的前提下,若最后生成的宽高尺寸达不到或超过 720,则在训练的时候进行在线增强,采用填充和宽高等比例缩放的方式将图像尺寸变换到 720×720;

(4) 如果在不超过图像边界时,最小的瑕疵图像尺寸下仍包含了其他瑕疵,则计算其他瑕疵

在该瑕疵图像中的 IOU 值,若值大于 0.2,则舍弃该瑕疵图像,若小于 0.2,则包含该瑕疵图像。

制作完的数据集包含 20 506 张分辨率大小为 720×720 的纺织品图像。该数据集分别拥有 14 342 个和 3 057 个训练和验证样本,此外还有 3 107 张包含无瑕疵样本的测试图样。

GDIM-CD 的每个训练样本拥有粗粒度和细粒度两个标注标签。其中,粗粒度标注将训练集划分为 20 个瑕疵类别和 1 个无疵点类别;细粒度标注将训练集划分为 34 个瑕疵类别和 1 个无疵点类别。GDIM-CD 足以支撑深度学习模型在粗、细粒度纺织品瑕疵分类任务上进行训练与测试。GDIM-CD 训练样本的标签和对应类别如下表 2 所示。

为了评估 TDRNet 的分类精度,本文使用 top-1 错误率和 top-5 错误率来评价网络性能。其中,top-1 错误率指模型对瑕疵进行分类时,概率最大的类别不为瑕疵真实类别的样本在占总体的比例;top-5 错误率为模型对瑕疵进行分类时,概率最高的五种类别中不包含真实类别的样本占总体的比例。top-1 和 top-5 错误率两个评价指标可较为公正客观地刻画模型的分类能力。

4.2 网络训练

4.2.1 实验环境

为了训练和测试 TDRNet,实验环境所需的软硬件配置如表 3 所示。

服务器上的 CPU 和 GPU 数量为 1,内存条数量为 2。该实验环境可以满足 TDRNet 模型训练与测试的需要。

4.2.2 训练策略

TDRNet 使用改进型的 ResNet-50 和 ResNet-101 作为骨干网络。实验中使用小批量

表 2 GDIM-CD 数据集瑕疵类别

Tab. 2 Defect Category of GDIM-CD

瑕疵名称	粗分类标签	细分类标签	瑕疵名称	粗分类标签	细分类标签
无疵点	0	0	星跳	16	18
破洞	1	1	跳花	16	19
水渍	2	2	断氨纶	17	20
油渍	2	3	稀密档	18	21
污渍	2	4	浪纹档	18	22
三丝	3	5	色差档	18	23
结头	4	6	磨痕	19	24
花板跳	5	7	轧痕	19	25
百脚	6	8	修痕	19	26
毛粒	7	9	烧毛痕	19	27
粗经	8	10	死皱	20	28
松经	9	11	云织	20	29
断经	10	12	双纬	20	30
吊经	11	13	双经	20	31
粗维	12	14	跳纱	20	32
纬缩	13	15	筘路	20	33
浆斑	14	16	纬纱	20	34
整经结	15	17	不良		

梯度随机下降算法(Mini-batch Stochastic Gradient Descent)作为模型的优化器进行网络学习。优化器初始学习率设置为 0.01,动量设置为 0.9,衰减系数设置为 1E-4,TDRNet 一共训练 300 个 epoch,且模型使用 StepLR 训练策略,在 epoch 为 (60,140,200,240,270)的时候对学习率进行调整,调整率 γ 设置为 0.2,另外,LEM 模块的缩放系数 μ 与学习率进行同步调整,分别设置为(0,1E-5,3E-3,3E-4,1E-5,1E-6)。

表 3 TDRNet 实验环境

Tab. 3 Experimental environment of TDRNet

软硬件名称	硬件型号/软件版本
CPU	Intel(R) Core(R) i7-10700KF@3.8GHz
GPU	NVIDIA RTX3090
内存	Fury HX432C16FB3K2/32
操作系统	Ubuntu 20.04.3 LTS (GNU/Linux 5.11.0-37-generic x86_64)
CUDA 版本	11.1.1
Python 版本	3.9.6
Pytorch 版本	1.9.0

在模型训练的过程中,训练集得到了在线增强,增强数据集的策略有上下翻转、左右翻转、随即旋转、色彩抖动,以使模型具备更强的鲁棒性和更好的泛化能力。

4.2.3 训练细节

初始学习率和学习率的调整对模型最终性能有至关重要的影响。本文使用文献[24]提出的学习率搜索方法搜寻初始学习率 α 。该方法使 α 按指数级增长,并在 $[1E-4, 1E-1]$ 区间内搜索最佳学习率。在以改进型ResNet-50作为骨干网络的TDRNet上训练了400个iteration后,学习率搜索的实验结果如图8所示。

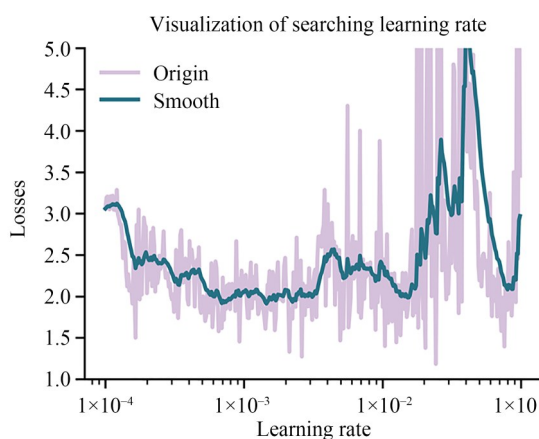


图8 基于改进型ResNet50的TDRNet初始学习率搜索可视化

Fig.8 visualization of searching initial learning rate of TDRNet based on improved Resnet-50

由于网络权重进行了随机初始化,训练初期网络损失的波动比较剧烈,从原始损失值中很难观察到最佳初始学习率。因此,本文使用指数滤波器对原始损失曲线进行平滑滤波处理,从而得到平滑后的损失曲线。在图8中,淡紫色曲线是原始学习率搜索曲线,墨绿色曲线为平滑后学习率搜索曲线。由于过小的学习率会导致网络收敛速度变慢甚至无法跳出局部最优,而过大的学习率又会导致损失不减反增。在1个epoch内,当学习率小于最佳学习率时,损失曲线整体呈下降趋势,而大于最佳学习率时,损失曲线整体呈上升趋势。由此在第1个epoch中,损失曲线上损失值最低点对应的学习率即为最佳初始学习率。根据图8,最佳学习率在落在 $[9E-3, 2E-$

2]区间,为了便于调参和避免使用过大的学习率导致网络无法收敛到近似全局最优,本文保守地将初始学习率设置为0.01。在选定初始学习率后,在整个网络的训练中,学习率调整策略如图9所示。

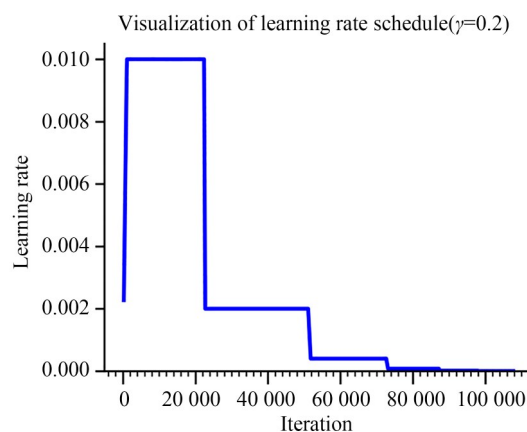


图9 基于改进型ResNet50的TDRNet学习率策略可视化($\gamma=0.2$)

Fig.9 visualization of learning rate schedule ($\gamma=0.2$) of TDRNet based on improved Resnet-50

需要说明的是,TDRNet的训练还使用了学习率预热策略。由于在训练初期,在随机初始化的网络参数下,模型不具备数据集中的先验知识,较高的学习率会导致网络朝错误的方向学习。因此,本文在训练初期采用了学习率线性增长的方法,在训练的前3个epoch内,从0开始对学习率进行预热,以保证网络训练初期参数更新的稳定性并加快网络的收敛速度。

最终,基于改进型ResNet-50的训练曲线如图10所示。图10(a)为Seesaw Loss训练损失曲线;图10(b)为DP Loss训练损失曲线。

由图10(a)可知,Seesaw Loss曲线波动较为剧烈,过大的损失波动导致模型难以快速收敛到近似的全局最优。这是由于在GDIM-CD中,训练集存在大量难分样本导致的。为缓解Seesaw Loss损失波动剧烈的问题,在实验设施受限的条件下,本文采用了批次扩增策略近似地扩增批次大小,从而降低损失方差。具体来说,本文以每4个批次作为一个周期进行梯度累加和梯度更新操作。在每个周期的前3个批次内,网络只累积梯度而不更新参数,在第4个训练批次,在累

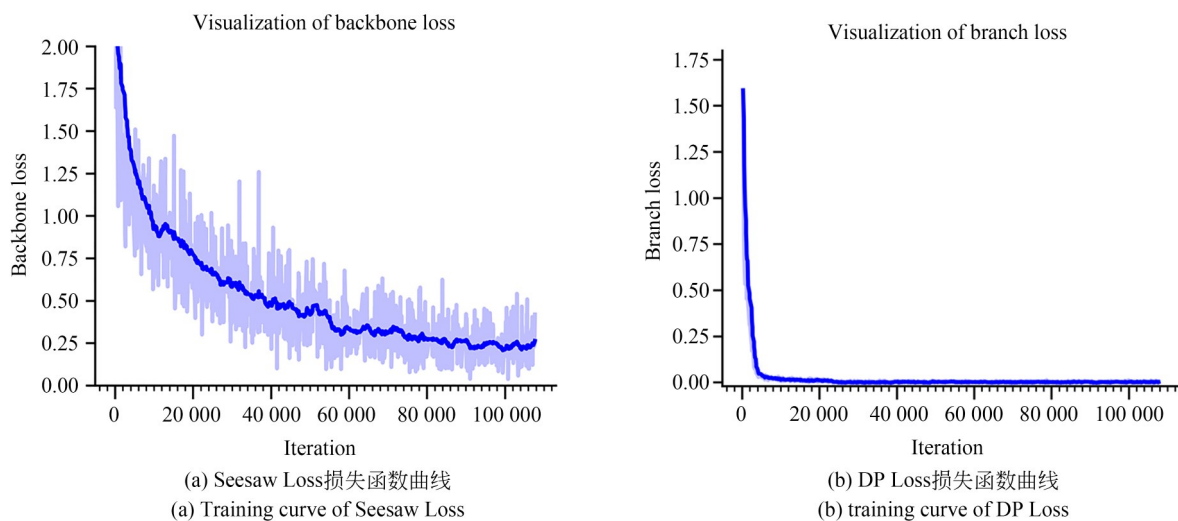


图 10 基于改进型 ResNet-50 的 TDRNet 网络损失曲线
Fig. 10 loss curve of TDRNet based on improved Resnet-50

积梯度后才更新网络参数。在图 10(b)中, DP Loss 损失曲线的收敛接近于理想状态,这是由于尽管使用高斯噪声扩增了样本,但送入 LEM 的数据样本规模仍远远小于 LEM 模块参数,因此网络经过 30 个 epoch 的训练后便达到了收敛状态。在训练 LEM 模块时,为了避免给 softmax 分类器的权重引入不必要的噪声,加大网络的学习成本并延长网络的训练周期,在开始训练的前 60 个 epoch, LEM 模块的缩放系数需设置为 0,即在训练的前期, LEM 模块以跟踪骨干网络的语义信息为主,以确保 TDRNet 准确快速地收敛。

随着网络的不断训练,网络性能也不断提高。以改进型 ResNet-50 作为骨干网络的 TDRNet 的精度曲线如图 11 所示。从图中可以看出,在 TDRNet 训练初期,网络的精度会随着 loss 的快速降低而快速提高,在训练中后期,网络逐渐达到收敛状态,精度不再上升。

4.3 实验分析

4.3.1 消融研究

为了验证框架的有效性,本文以 ResNet-50 作为所有方法的基准,在 GDIM-CD 数据集上的粗粒度分类任务中对 TDRNet 进行了一系列消融研究。消融研究结果如表 4 所示。

在消融研究中, TDRNet 不使用预训练模型,骨干网络和 LEM 模块都对权重进行随机初始化,且它们都使用同样的数据增强方法对训练

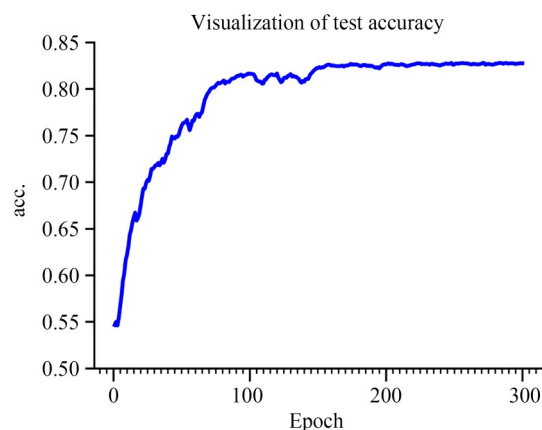


图 11 基于改进型 ResNet-50 的 TDRNet 精度曲线
Fig. 11 accuracy curve of TDRNet based on improved Resnet-50

集进行数据增强,在消融研究中,在不使用 Seesaw Loss 的时候,网络使用交叉熵损失作为分类损失训练模型。在作为基准的 ResNet-50 模型中,其在 GDIM-CD 测试集上粗粒度分类任务的 top1 错误率为 20.52%。在仅添加了 LEM 模块的前提下,其 top1 错误率为 20.59%,与没添加模块想必,网络的性能并没有得到提升,这与预期是符合的,即单纯使用 LEM 模块并不能提升网络的性能。在嵌入 LEM 模块后使用 DP Loss,网络性能即得到提升,错误率下降了 2.42%;在单独使用 Seesaw Loss 的时候,网络的错误率降低为 18.97%,下降了 1.59%。更进一步地,在基准

表 4 TDRNet粗粒度分类任务上的消融研究结果
Tab. 4 Ablation study of TDRNet on rough-grained task

Backbone	LEM Model	DP Loss	Seesaw Loss	Top1 err. (%)
ResNet-50 (baseline)	×	×	×	20.52
ResNet-50	✓	×	×	20.59
ResNet-50	✓	✓	×	18.10
ResNet-50	×	×	✓	18.97
ResNet-50	✓	✓	✓	17.32
Improved ResNet-50	×	×	×	19.94
Improved ResNet-50 (TDRNet-50)	✓	✓	✓	16.80
ResNet-101	×	×	×	19.58
ResNet-101	✓	✓	✓	16.61
Improved ResNet-101	×	×	×	19.23
Improved ResNet-101 (TDRNet-101)	✓	✓	✓	16.35

方法上同时使用 LEM 模块、DP Loss 和 Seesaw Loss,结果表明网络 top1 错误率降低至 17.32%,与基准方法相比,性能提升幅度高达 3.20%。

紧接着,本文研究了骨干网络对网络分类性能的影响。与基准方法相比,只使用改进型 ResNet-50 同样可以使得网络的分类性能得到提升,减少了 0.58%,在 TDRNet-50 中,网络的 top1 分类错误率降低至 16.80%,对比基准方法与不使用改进型 ResNet-50,TDRNet -50 的 top1 错误率分别降低了 3.72% 和 0.52%。在使用 ResNet-101 和改进型 ResNet-101 之后,网络的 top1 错误率性能达到了 16.61% 和 16.35% 的水平。

在基准方法上的一系列消融研究表明,本文提出的一系列方法是有效的,LEM 和 DP Loss 的组合可以使得模型具有更为强大的特征提取能力;且 Seesaw Loss 在一定程度上缓解了负样本梯度和误分类损失对网络模型的不利影响。

4.3.2 对比实验

为了验证 TDRNet 的先进性,在 GDIM-CD 数据集上,本文将 TDRNet 与其他主流图像分类网络进行了性能对比。为确保比较的公平性,所有参与比较的网络模型均在相同的实验配置下进行训练。最终,在 GDIM-CD 测试集上,相关网络模型在细粒度分类任务上的分类性能结果如表 5 所示。

根据表 5 可知,TDRNet 具有中等规模的参数量、计算量和运行速度,但在 GDIM-CD 数据集的细粒度分类任务上,其取得了最佳的测试结果。其中,TDRNet-50 的 top1 错误率为 17.45%,top5 错误率为 5.20%;而 TDRNet-101 的 top1 错误率则为 17.12% 和 5.27%。与经典的 VGG16 相比,TDRNet 具备压倒性的优势,其 top1 错误率和 top5 错误率低 6.20% 和 5.53%;与先进的 CNN 架构 EfficientNet 相比,TDRNet 测试结果同样具备明显的优势,比起同等参数规模的 EfficientNet_B6,TDRNet-50 在 top1 错误率上的指标要低 0.90%,而 top5 错误率的指标则低 1.88%。与当前最先进的 Transformer 架构相比,本文所提方法同样有明显优势,但需要指出的是,在中小型数据集上,未经过预训练的 Transformer 模型缺乏先验信息,训练困难,其性能表现差强人意。但综合考虑相同的训练条件,TDRNet 框架仍是最先进的模型。

为了更进一步验证 TDRNet 的优势,在细粒度分类任务上,本文将 TDRNet 与近 5 年细粒度视觉分类领域中的主流模型进行了性能比较,对比结果如表 6 所示。从表中可知,TDRNet 在 GDIM-CD 数据集上取得了 SOTA 结果。与 MA-CNN 和 RA-CNN 相比,TDRNet 的 Top1 错误率低 2.69% 和 2.86%,Top5 错误率低 3.34% 和 3.09%;与同样使用 ResNet 作为基础

表 5 GDIM-CD 细粒度分类任务上的在主流分类模型对比实验结果

Tab. 5 Comparison of the typical classification model experimental results for fine-grained task on GDIM-CD

Model	Top1 err. / %	Top5 err. / %	Params. / M	FLOPs/G	FPS
EfficientNet_B0 ^[25]	25.11	12.64	4.03	4.40	33
DenseNet-169 ^[26]	23.11	12.44	12.52	35.55	13
EfficientNet_B4 ^[25]	20.20	8.15	17.59	16.70	16
DenseNet-201 ^[26]	19.49	8.46	18.13	45.44	11
ResNext-50 ^[27]	20.13	7.34	23.02	44.74	26
ResNet-50 ^[18]	21.27	7.60	23.55	43.13	36
TDRNet-50	17.45	5.20	32.34	58.91	29
EfficientNet_B6 ^[25]	18.35	7.08	40.78	36.91	11
ResNet-101 ^[18]	20.04	7.79	42.54	81.75	21
TDRNet-101	17.12	5.27	51.33	97.53	19
AlexNet ^[17]	27.34	11.18	57.09	7.37	277
WRN50 ^[28]	20.68	8.99	66.88	119.42	19
ViT_B_16 ^[29]	20.06	9.76	87.22	249.12	9
ViT_B_32 ^[29]	20.30	10.08	87.84	51.49	44
WRN101 ^[28]	19.13	8.63	124.88	237.08	11
VGG16 ^[30]	23.85	10.73	134.35	158.68	21
VGG19 ^[30]	24.01	10.54	139.66	201.67	18
ViT_L_16 ^[29]	21.55	11.09	305.20	815.69	3
ViT_L_32 ^[29]	21.39	11.38	306.02	175.66	17

骨干网络的 TASN、DCL 比较, Top1 错误率和 Top5 错误率两个分类指标分别低 1.25%, 1.12% 和 1.52%, 1.84%; 与 WS-DAN 作对比, TDRNet 在 Top1 错误率和 Top5 错误率指标上仍有 0.68% 和 0.98% 的微弱优势。与使用 ViT 作骨干网络的 TransFG 比较, TRNet 分类性能同样更佳, 同样需要指出的是, TransFG 和上文视觉 Transformer 模型 ViT 一样, 是在没有预训练

的条件下得到的结果, 在实验条件受限的情形下可能无法发挥 Transformer 架构的优秀性能。但这仍说明了在同等的实验条件下, 本文的方法明显更加先进。

表 5 和表 6 的实验结果证明了本文方法在细粒度瑕疵分类上具备优良的性能, 从而证实了 TDRNet 的先进性。

表 6 GDIM-CD 细粒度分类任务上与细粒度分类模型的对比实验结果

Tab. 6 Comparison of the experimental results with fine-grained classification model for fine-grained task on GDIM-CD

Model	TDRNet	MA-CNN ^[31]	RA-CNN ^[32]	WS-DAN ^[33]	TASN ^[34]	DCL ^[35]	TransFG ^[36]
Backbone	ResNet-50	VGG-19	VGG-19	ResNet-50	ResNet-50	ResNet-50	ViT_B_16
Top1 err. / %	17.45	20.14	20.31	18.13	18.70	18.57	20.43
Top5 err. / %	5.20	8.54	8.29	6.18	6.72	7.04	8.97

5 结 论

本文根据工业场景下的纺织品瑕疵分类存在的实际问题, 提出了一种基于标签嵌入的纺织

品瑕疵分类网络。该网络调整了骨干网络的结构和输入图像的分辨率, 以保证检测精度和速度满足实际工业生产需求。为了增加算法对纺织品瑕疵的区分能力, 网络中增加了标签嵌入模

块,该模块利用标签先验信息来调整不同瑕疵语义在特征空间中的分布,为了使得标签嵌入模块具备学习特征语义偏移的能力,本文提出了分布感知损失来辅助网络训练;此外,网络使用 See-saw Loss 进行分类特征学习,以保证在数据集类别偏斜的情况下,网络能正确地更新网络的参数梯度。在以公开的“广东智能制造布匹瑕疵检测数据集”构建的 GDIM-CD 数据集上, TDRNet

框架在粗粒度分类任务上的 top1 错误率可低至 16.35%,在细粒度分类任务上的 top1 错误率和 top5 错误率则为 17.12% 和 5.20%。该实验结果表明,本文方法对于提升纺织品瑕疵的识别分类是有效的,且优于目前的主流方法。如何进一步降低瑕疵分类的错误率,以及检出纺织品瑕疵的位置和大小并提高检出精准度,是接下来本文需进一步研究的问题。

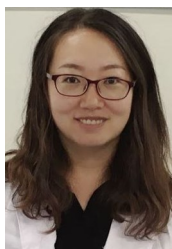
参考文献:

- [1] 庄集超. 基于深度卷积神经网络的布匹瑕疵点检测算法研究[D]. 秦皇岛:燕山大学, 2020.
ZHUANG J C. *Research on Detection Algorithm Based on Deep Convolution Neural Network for Fabric Defects*[D]. Qinhuangdao: Yanshan University, 2020. (in Chinese)
- [2] CHEN J H, JAIN A K. A structural approach to identify defects in textured images[C]. *Proceedings of the 1988 IEEE International Conference on Systems, Man, and Cybernetics*, 8-12, 1988, Beijing, China. IEEE, 2002: 29-32.
- [3] COHEN F S, FAN Z, ATTALI S. Automated inspection of textile fabrics using textural models[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, 13(8): 803-808.
- [4] 黄佩璠. 基于深度学习的布匹瑕疵检测识别的研究和应用[D]. 南昌:南昌大学, 2021.
HUANG P F. *Research and Application of Fabric Defect Detection and Recognition Based on Deep Learning* [D]. Nanchang: Nanchang University, 2021. (in Chinese)
- [5] CHIU S H, CHOU S, LIAW J J, *et al.* Textural defect segmentation using a Fourier-domain maximum likelihood estimation method[J]. *Textile Research Journal*, 2002, 72(3): 253-258.
- [6] JING J F, FAN X T, LI P F. Automated fabric defect detection based on multiple Gabor filters and KPCA [J]. *International Journal of Multimedia and Ubiquitous Engineering*, 2016, 11(6): 93-106.
- [7] LIANG Z, XU B, CHI Z, *et al.* Intelligent characterization and evaluation of yarn surface appearance using saliency map analysis, wavelet transform and fuzzy ARTMAP neural network[J]. *Expert Systems With Applications*, 2012, 39(4): 4201-4212.
- [8] 陈彦彤, 陈伟楠, 张献中, 等. 基于深度卷积神经网络的蝇类面部识别[J]. *光学精密工程*, 2020, 28(7): 1558-1567.
CHEN Y T, CHEN W N, ZHANG X Z, *et al.* Fly facial recognition based on deep convolutional neural network[J]. *Opt. Precision Eng.*, 2020, 28(7): 1558-1567. (in Chinese)
- [9] 王宸, 张秀峰, 刘超, 等. 改进 YOLOv3 的轮毂焊缝缺陷检测[J]. *光学精密工程*, 2021, 29(8): 1942-1954.
WANG C, ZHANG X F, LIU C, *et al.* Detection method of wheel hub weld defects based on the improved YOLOv3[J]. *Opt. Precision Eng.*, 2021, 29(8): 1942-1954. (in Chinese)
- [10] JI L Y, JIANG X Y, GAO Y B, *et al.* ADR-Net: context extraction network based on M-Net for medical image segmentation[J]. *Medical Physics*, 2020, 47(9): 4254-4264.
- [11] WANG C Y, LI L F. Multi-scale residual deep network for semantic segmentation of buildings with regularizer of shape representation [J]. *Remote Sensing*, 2020, 12(18): 2932.
- [12] LIU W J, ZHANG Y J, FAN H S, *et al.* A new multi-channel deep convolutional neural network for semantic segmentation of remote sensing image [J]. *IEEE Access*, 2020, 8: 131814-131825.
- [13] 陈欣, 万敏杰, 马超, 等. 采用多尺度特征融合 SSD 的遥感图像小目标检测[J]. *光学精密工程*, 2021, 29(11): 2672-2682.
CHEN X, WAN M J, MA C, *et al.* Recognition of small targets in remote sensing image using multi-scale feature fusion-based shot multi-box detector[J]. *Opt. Precision Eng.*, 2021, 29(11): 2672-2682. (in Chinese)
- [14] 王阳, 陈薇伊, 马军山. 基于卷积神经网络的乳腺癌良恶性诊断[J]. *软件工程*, 2022, 25(1): 6-9.
WANG Y, CHEN W Y, MA J S. Diagnosis of

- benign and malignant breast cancer based on convolutional neural network [J]. *Software Engineer*, 2022, 25(1): 6-9. (in Chinese)
- [15] UZEN H, TURKOGLU M, HANBAY D. Texture defect classification with multiple pooling and filter ensemble based on deep neural network [J]. *Expert Systems With Applications*, 2021, 175: 114838.
- [16] 余永维, 韩鑫, 杜柳青. 基于 Inception-SSD 算法的零件识别 [J]. *光学精密工程*, 2020, 28(8): 1799-1809.
- YU Y W, HAN X, DU L Q. Target part recognition based Inception-SSD algorithm [J]. *Opt. Precision Eng.*, 2020, 28(8): 1799-1809. (in Chinese)
- [17] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [18] HE K M, ZHANG X Y, REN S Q, *et al.* Deep Residual Learning for Image Recognition [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [19] HU J, SHEN L, SUN G. Squeeze-and-Excitation Networks [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [20] OUYANG W B, XU B G, HOU J, *et al.* Fabric defect detection using activation layer embedded convolutional neural network [J]. *IEEE Access*, 2019, 7: 70130-70140.
- [21] 陆贵家. 基于 Cascade R-CNN 改进的花色布匹瑕疵智能识别方法 [J]. *现代信息科技*, 2020, 4(23): 20-24.
- LU G J. Improved intelligent recognition method of pattern and color fabric defects based on cascade R-CNN [J]. *Modern Information Technology*, 2020, 4(23): 20-24. (in Chinese)
- [22] PENG P R, WANG Y, HAO C, *et al.* Automatic fabric defect detection method using PRAN-net [J]. *Applied Sciences*, 2020, 10(23): 8434.
- [23] Tianchi. Smart Diagnosis of Cloth Flaw Dataset [DB/OL]. (2020). <https://tianchi.aliyun.com/dataset/79336>
- [24] SMITH L N. Cyclical Learning Rates for Training Neural Networks [C]. 2017 *IEEE Winter Conference on Applications of Computer Vision (WACV)*. March 24-31, 2017, Santa Rosa, CA, USA. IEEE, 2017: 464-472.
- [25] TAN M, LE Q. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks [C]. International conference on machine learning. PMLR, 2019: 6105-6114.
- [26] HUANG G, LIU Z, VAN DER MAATEN L, *et al.* Densely Connected Convolutional Networks [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 2261-2269.
- [27] XIE S N, GIRSHICK R, DOLLÁR P, *et al.* Aggregated Residual Transformations for Deep Neural Networks [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 5987-5995.
- [28] ZAGORUYKO S, KOMODAKIS N. Wide residual networks [J]. *arXiv preprint arXiv:1605.07146*, 2016.
- [29] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al.* An image is worth 16x16 words: Transformers for image recognition at scale [J]. *arXiv preprint arXiv:2010.11929*, 2020.
- [30] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. *arXiv preprint arXiv:1409.1556*, 2014.
- [31] ZHENG H L, FU J L, MEI T, *et al.* Learning Multi-Attention Convolutional Neural Network for Fine-Grained Image Recognition [C]. 2017 *IEEE International Conference on Computer Vision (ICCV)*. 22-29, 2017, Venice, Italy. IEEE, 2017: 5219-5227.
- [32] FU J L, ZHENG H L, MEI T. Look Closer to See Better: Recurrent Attention Convolutional Neural Network for Fine-Grained Image Recognition [C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 21-26, 2017, Honolulu, HI, USA. IEEE, 2017: 4476-4484.
- [33] HU T, QI H, HUANG Q, *et al.* See better before looking closer: Weakly supervised data augmentation network for fine-grained visual classification [J]. *arXiv preprint arXiv:1901.09891*, 2019.
- [34] ZHENG H L, FU J L, ZHA Z J, *et al.* Looking

- for the Devil in the Details: Learning Trilinear Attention Sampling Network for Fine-Grained Image Recognition[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 5007-5016.
- [35] CHEN Y, BAI Y L, ZHANG W, *et al.* Destruction and Construction Learning for Fine-Grained Image Recognition[C]. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 15-20, 2019, Long Beach, CA, USA. IEEE, 2020: 5152-5161.
- [36] HE J, CHEN J N, LIU S, *et al.* TransFG: a transformer architecture for fine-grained recognition [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, 36(1): 852-860.

通讯作者:



刘 颖(1983—),女,吉林长春人,2015年获得长春理工大学电子信息工程系博士学位,现为长春理工大学电子与信息工程学院讲师,主要从事姿态估计、动作识别、人/车重识别相关的计算机视觉方面的研究。E-mail: liuying02@cust.edu.cn

作者简介:



姜 威(1998—),女,吉林长春人,在读硕士,2020年于吉林师范大学获得电子信息工程学士学位,现为长春理工大学电子与信息工程学院硕士研究生,主要从事语义分割和目标检测方面的研究。