

基于改进部分卷积的瑕疵布匹图像生成算法^①



乐 飞, 宋亚林, 李小艳

(河南大学 软件学院, 开封 475004)

通信作者: 宋亚林, E-mail: syl@vip.henu.edu.cn

摘 要: 针对工业生产中布匹瑕疵自动化检测模型训练时缺少带瑕疵位置信息的瑕疵布匹图像数据集的问题, 本文提出了一种以改进的部分卷积网络作为基本框架的带瑕疵位置信息的瑕疵布匹图像生成模型 EC-PConv. 该模型引入小尺寸瑕疵特征提取网络, 将提取出的瑕疵纹理特征与空白 mask 拼接起来形成带有位置信息和瑕疵纹理特征的 mask, 然后以修复方式生成带有瑕疵位置信息的瑕疵布匹图像. 另外, 本文提出一种结合 MSE 损失的混合损失函数以生成更加清晰的瑕疵纹理. 实验结果表明, 与最新的 GAN 生成模型相比, 本文提出的生成模型的 FID 值降低了 0.51; 生成的瑕疵布匹图像在布匹瑕疵检测模型中查准率 P 和 MAP 值分别提高了 0.118 和 0.106. 实验结果表明, 该方法在瑕疵布匹图像生成上比其他算法更稳定, 能够生成更高质量的带瑕疵位置信息的瑕疵布匹图像, 可较好地解决布匹瑕疵自动化检测模型缺少训练数据集的问题.

关键词: 机器视觉; 部分卷积; 图像生成; 布匹瑕疵检测模型; MSE 损失

引用格式: 乐飞, 宋亚林, 李小艳. 基于改进部分卷积的瑕疵布匹图像生成算法. 计算机系统应用, 2022, 31(12): 187-194. <http://www.c-s-a.org.cn/1003-3254/8851.html>

Image Generation Algorithm for Defective Cloth Based on Improved Partial Convolution

YUE Fei, SONG Ya-Lin, LI Xiao-Yan

(School of Software, Henan University, Kaifeng 475004, China)

Abstract: Given the problem that no image datasets of defective cloth with defect location information are available for the training of the automatic detection model for cloth defects in industrial production, this study proposes an image generation model EC-PConv with defect location information for defective cloth, and it uses an improved partial convolutional network as its basic framework. This model adopts a feature extraction network for small-scale defects, splices the extracted defect texture features with the blank mask to obtain a mask with position information and defect texture features, and generates an image with defect position information in a repaired way for the defective cloth. Furthermore, a hybrid loss function integrating the mean squared error (MSE) loss is proposed to generate clearer defect textures. The experimental results show that compared with the latest generative adversarial network (GAN) generation model, the proposed model reduces the Fréchet inception distance (FID) score by 0.51 and improves the precision P and mean average precision (MAP) values of the generated image of the defective cloth in the cloth defect detection model by 0.118 and 0.106, respectively. This method is more stable than other algorithms in generating images of defective cloth and can generate images of defective cloth that contain defect location information and have higher quality. Therefore, it can effectively solve the problem that no training datasets are available for the automatic detection model for cloth defects.

Key words: machine vision; partial convolution; image generation; cloth defect detection model; MSE loss

^① 基金项目: 河南省科技研发项目 (212102210078); 河南省重大科技专项 (201300210400); 河南省重点研发与推广专项 (科技攻关)(202102210380)

收稿时间: 2022-03-23; 修改时间: 2022-05-16; 采用时间: 2022-05-28; csa 在线出版时间: 2022-07-22

1 引言

随着社会工业化的不断发展,在人们的日常生活中,布匹作为一种生活必需品发挥着不可替代的作用,布匹质量的好坏也越发重要.传统的布匹瑕疵检测方法通过人的肉眼来筛选,但这种方法效率低下.近年来深度学习在工业生产中的应用解决了很多人工无法解决的问题,使得布匹瑕疵检测变得自动化、智能化,消除了传统方法的弊端.YOLO^[1]等使用深度学习训练布匹瑕疵的检测模型往往需要大量带瑕疵位置信息的瑕疵布匹图像做数据集,但数据集获取难度大、标注成本高.因此,一种能够生成带瑕疵位置信息的瑕疵布匹图像的方法十分重要.

随着深度学习的不断发展,许多使用生成对抗网络的算法生成图像越来越逼真^[2-6],其中也有将生成对抗网络与图像纹理等相结合得到的图像生成算法^[7-13],以及使用注意力机制等图稿图像生成质量的算法^[14-16].

传统的图像生成方法通过各种变换操作改变原始图像的表现形态,以产生新图像,主要可以分为几何变换、色域变换、清晰度变换、噪声注入和局部擦除等方法.LeNet-5^[17]采用了平移、缩放、挤压和水平裁剪等几何变换生成新图像且可以保留原始图像数据的标注信息.AlexNet^[18]使用PCA抖动,通过给协方差矩阵增加噪声实现一种图像在视觉表现上的滤镜效果,以此产生新图像.这些方法主要对单幅图像进行操作,生成新图像时可以利用的先验知识很少,仅仅是图像本身的信息和图像对的互信息,无法增加信息量,缺乏多样性.

随着深度神经网络的不断发展,通过机器学习生成图像的方法越来越成熟,这种方法使用神经网络学习训练数据中图像的高维分布,再对学习到的分布进行过采样生成新图像,由于将整个训练集作为先验知识,可以生成逼真的图像,同时满足了数据的多样性.VAE (variational autoencoder)^[19]对编码器添加单位高斯分布约束,再从得到的单位高斯分布中进行采样,最后把它传给解码器生成新图像.GAN (generative adversarial networks)^[20]利用零和博弈思想,使生成器和判别器在不断优化中各自提高自己的生成能力和判别能力,最后通过对学习到的数据分布进行过采样来生成图像.这些算法生成的图像获得了很好的视觉效果.但是,现有的图像生成领域中关于带瑕疵位置信息的瑕疵布匹图像生成的研究很少.

本文的主要贡献为提出了一种带瑕疵位置信息的

瑕疵布匹图像生成模型 EC-PCConv,并通过实验确定生成的带瑕疵位置信息的瑕疵布匹图像具有较高的质量以及良好的迁移能力,在瑕疵布匹检测模型中取得了较高的瑕疵检测率.

2 相关理论

2.1 部分卷积网络

现有的基于深度学习的图像修复方法在受损图像上使用标准卷积网络,使用以有效像素以及 mask 中的替换值(通常为平均值)为条件的卷积滤波器响应.但通常会导致颜色不一致和模糊等伪影.可以通过后棋处理减少这类伪影,但成本很高,而且可能会失败.部分卷积网络 (partial convolutional)^[21]使用了部分卷积层设计,部分卷积层由图像卷积和 mask 卷积构成,网络整体使用 U-Net 结构,8 个下采样部分卷积层组成编码网络,8 个上采样部分卷积层组成解码网络,最后将完成修复的图片输出.结构如图 1 所示.其中部分卷积层使用卷积方式对 mask 进行更新,使得在每一个部分卷积层都使用上一层更新后的 mask,随着网络层数的增加,mask 中为 0 的像素越来越少,输出的结果特征图中有效区域的面积越来越大,mask 对整体 loss 的影响也越来越小.网络在训练学习过程中采用从 mask 外到内逐渐缩小的学习机制.底层网络学习 mask 外围,高层网络学习 mask 里面.与整个网络都学习整个 mask 相比具有明显的优势.部分卷积层结构如图 2 所示.

部分卷积网络的损失函数由感知损失、风格损失、L1 损失和 TV 损失构成.感知损失计算原本不带瑕疵的布匹图像和生成的带瑕疵图像在所有或几层特征值之间的 L1 距离.具体感知损失表示如下:

$$L_{\text{perceptual}} = \sum_{p=0}^{p-1} \frac{\|\Psi_p^{\text{out}} - \Psi_p^{\text{gt}}\|_1}{N_{\Psi_p^{\text{gt}}}} + \sum_{p=0}^{p-1} \frac{\|\Psi_p^{\text{comp}} - \Psi_p^{\text{gt}}\|_1}{N_{\Psi_p^{\text{gt}}}} \quad (1)$$

风格损失使用 gram 矩阵计算裁剪下来的瑕疵图像和不带瑕疵的布匹图像的 gram 矩阵之间的风格差异.风格损失表示如下:

$$L_{\text{style}} = \sum_{p=0}^{p-1} \frac{1}{C_p C_p} \left\| K_p \left(\left(\Psi_p^{\text{comp}} \right)^T \left(\Psi_p^{\text{comp}} \right) - \left(\Psi_p^{\text{gt}} \right)^T \left(\Psi_p^{\text{gt}} \right) \right) \right\|_1 \quad (2)$$

L1 损失是为了确保图像整体像素级重建的准确性, 分为瑕疵部分 L_{hole} 和非瑕疵部分的损失 L_{valid} :

$$L_{hole} = \frac{1}{N_{I_{gt}}} \left\| (1 - M) \odot (I_{out} - I_{gt}) \right\|_1$$

$$L_{valid} = \frac{1}{N_{I_{gt}}} \left\| M \odot (I_{out} - I_{gt}) \right\|_1 \quad (3)$$

全局平滑性损失 total variation (TV) 保证了生成的

布匹瑕疵区域与原本不带瑕疵的布匹图像区域相接的位置能够更加平滑. TV 损失表示如下:

$$L_{TV} = \sum_{(i,j) \in R, (i,j+1) \in R} \frac{\|I_{comp}^{i,j+1} - I_{comp}^{i,j}\|_1}{N_{I_{comp}}} + \sum_{(i,j) \in R, (i+1,j) \in R} \frac{\|I_{comp}^{i+1,j} - I_{comp}^{i,j}\|_1}{N_{I_{comp}}} \quad (4)$$

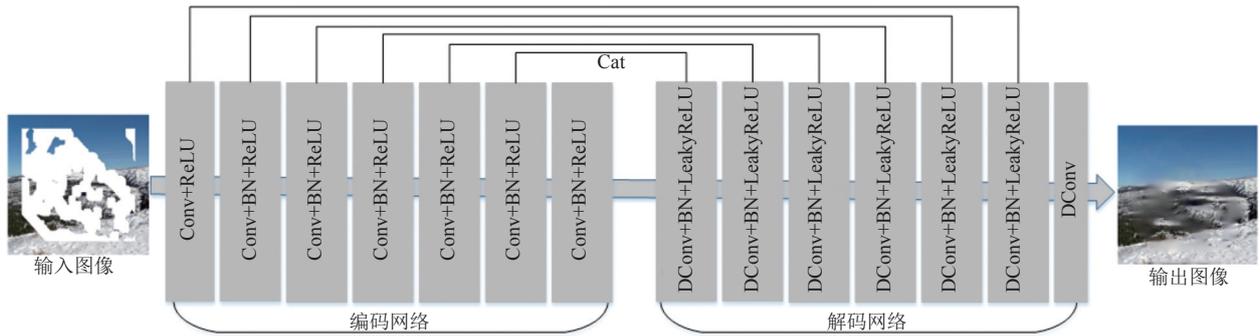


图1 部分卷积网络结构图

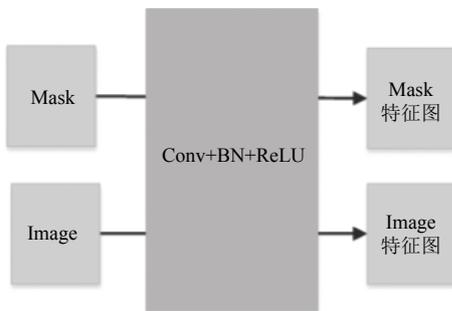


图2 部分卷积层网络结构



图3 CGAN 生成的手写数字图像

2.2 条件生成对抗网络

基础的 GAN 模型中使用一种分布直接进行采样, 从而达到生成图像完全逼近真实图像的效果, 这也是 GAN 最大的优势. 然而, GAN 生成的图像是随机的, 不能控制生成图像属于何种类别. 为了解决这个问题, 一个很自然的想法是给 GAN 加一些约束, CGAN^[22] 提出了一种带条件约束的 GAN, 在生成模型 (D) 和判别模型 (G) 的建模中均引入条件变量 y , 使用额外信息 y 对模型增加条件, 可以指导数据生成过程. 这些条件变量 y 可以基于多种信息, 例如类别标签, 用于图像修复的部分数据. 如果条件变量 y 是类别标签, 可以看做 CGAN 是把纯无监督的 GAN 变成有监督模型的一种改进. 如图 3 所示, 根据 MNIST 上以类别标签为条件 (one-hot 编码) 训练 CGAN, 生成手写数字图像.

3 瑕疵布匹图像生成模型 EC-PConv

3.1 EC-PConv 网络结构

本文对部分卷积网络进行了改进, 添加了瑕疵特征提取分支网络, 使得改进后的部分卷积网络可以生成带瑕疵位置信息的瑕疵布匹图像. 改进后的网络结构如图 4 所示.

CGAN 通过添加条件信息控制生成图像的种类, 鉴于此, 本文在部分卷积网络修复过程中通过瑕疵特征提取网络给 mask 添加上条件信息, 将原本的空白 mask 变为带有瑕疵位置信息和瑕疵特征的 mask, 网络使用带有瑕疵特征的 mask 将待修复部分修复成瑕疵, 并通过控制 mask 位置可以修复出位置信息已知的瑕疵, 最终达到生成带瑕疵位置信息的瑕疵布

匹图像的效果. 本文对部分卷积网络中 mask 掩膜修复添加条件, 利用卷积网络提取布匹瑕疵的深层特征作为网络中 mask 掩膜修复的条件, 网络在修复过程中逐步将待修复部分的图像修复为布匹瑕疵, 从而在原本没有瑕疵的布匹图像上生成瑕疵.

3.2 瑕疵特征提取网络

瑕疵特征提取网络利用神经网络的特性来提取裁剪出来的布匹瑕疵的特征, 如图 5 所示. 首先, 对裁剪出来带位置信息的瑕疵执行一次卷积-归一化-ReLU (Conv-BN-ReLU) 操作, 将 512×512 大小的布匹瑕疵

图像下采样为 256×256 大小 16 通道的布匹瑕疵特征图, 接着再执行一次 Conv-BN 运算, 将 256×256 大小 16 通道布匹瑕疵特征图下采样为 128 大小 32 通道的布匹瑕疵特征图. 最后, 将 256×256 大小 32 通道布匹瑕疵特征图输入到 MaxPooling 层, 通过进一步对特征的提取减少无关特征, 提取出布匹瑕疵的纹理特征图, 输出 64×64 大小 64 通道布匹瑕疵特征图. 由于布匹的瑕疵尺寸很小, 因此采用两层卷积层加一层最大池化层的结构来确保能够提取出瑕疵的主要特征.

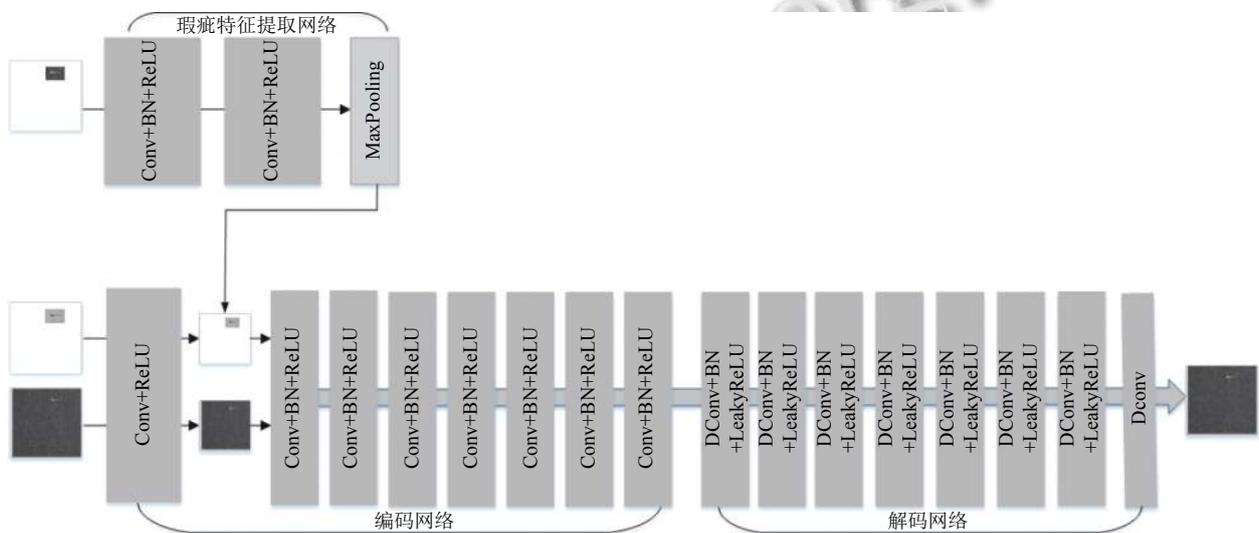


图 4 EC-PConv 网络结构

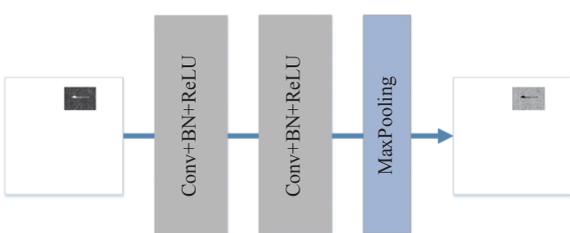


图 5 瑕疵特征提取网络结构

瑕疵特征提取网络包含两个卷积层和一个最大池化层, 由于破洞瑕疵面积较小, 两层卷积层的设计既可以提取破洞瑕疵的特征, 又不会因为网络过深导致破洞瑕疵特征丢失过多, 最大池化层可以保留卷积后的有效特征, 还能够压缩数据和参数的数量, 减小过拟合, 并且提高了模型的容错性.

网络中的部分卷积层将卷积分为了输入图片的卷积和输入 mask 的卷积. 部分卷积层的形式如下:

$$\begin{cases} x'_{(i,j)} = \begin{cases} W^T(X_{(i,j)} \odot M_{(i,j)})r_{(i,j)}, & \|M_{(i,j)}\|_1 > 0 \\ 0, & \text{otherwise} \end{cases} \\ \text{where } r_{(i,j)} = \frac{\|1_{(i,j)}\|_1}{\|M_{(i,j)}\|_1} \end{cases} \quad (5)$$

其中, M 代表 mask, mask 覆盖的区域带有提取出的瑕疵特征, 即待修复区域. 每次卷积计算的结果都受到瑕疵特征的影响, 随着卷积核对应图像区域内元素的增多, 通过 r 这个比例因子来调控有效信息的缩放. 每次进行完部分卷积操作后, 需要对 M 进行更新, 更新公式如下:

$$m'_{(i,j)} = \begin{cases} 1, & \|M_{(i,j)}\|_1 > 0 \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

从式 (2) 中可以看到, 由于卷积核所对应的 mask 区域带有瑕疵特征, 那么卷积核中心位置对应的 mask 就会在更新过程中根据瑕疵特征进行修复. 通过多次

迭代, mask 的区域会越来越趋向于瑕疵, 最终将待修复区域修复成 mask 中给定的带有瑕疵特征的瑕疵, 并且修复出的瑕疵还带有位置信息。

瑕疵提取网络用于将布匹瑕疵图像转为 64 通道的瑕疵特征图, 再将提取到的瑕疵特征图拼接经过编码网络的第一个部分卷积层之后的 64 通道 mask 上, 拼接后的 mask 不再是空白 mask, 而是带有瑕疵特征的 mask; 再将经过编码网络第一个部分卷积层, 得到的 64 通道待修复不带瑕疵的布匹图像特征图, 拼接后 64 通道 mask 特征图输入编码网络的第 2 到第 8 个部分卷积层, 得到 3×3 大小 512 通道的特征图; 解码网络则将编码网络提取到的 3×3 大小 512 通道的通特征图作为输入, 经过 8 个上采样层最终输出带瑕疵的瑕疵布匹图像。另外, 作为瑕疵特征提取网络输入的瑕疵图像, 是经过已有的带瑕疵位置信息的瑕疵布匹图像裁剪得来, mask 也是由相应的带瑕疵位置信息的瑕疵布匹图像生成得来, 因此, 瑕疵特征提取网络输入的瑕疵图像与 mask 图像都带有瑕疵的位置信息, 经过本文的 EC-PConv 网络模型在原本不带瑕疵的布匹图像上生成的瑕疵也就有了位置信息, 并且本文提出的带瑕疵位置信息的瑕疵布匹图像生成的 EC-PConv 网络模型可以在裁剪瑕疵图像与生成 mask 时控制瑕疵的位置。最终, 通过对部分卷积网络的改进实现了带瑕疵位置信息的瑕疵布匹图像的生成。

3.3 网络损失函数

本文在部分卷积网络工作的基础上结合了 MSE 损失, 以此构造一个新的混合损失函数。部分卷积网络中采用 L1 损失分别计算孔洞和非空洞区域的损失, 即 mask 区域与非 mask 区域的像素损失, 但原本的部分卷积网络是以修复图像为目的, 需要比较 mask 遮盖区域与周边没有 mask 遮盖区域的像素差来更好地修复图像。本文提出的 EC-PConv 网络是在原本没有瑕疵的图像上生成带位置信息的破洞和三丝两种类型的瑕疵, 而破洞和三丝这两种瑕疵都是小尺寸的瑕疵, 它们的特征主要是纹理特征。原本的布匹图像由于不带瑕疵, 纹理和像素没有隔断的异常点, 本文在原本没有瑕疵的布匹图像上生成瑕疵, 在原本正常的布匹纹理上加入了异常点, 这些异常点就是破洞和三丝两种瑕疵的纹理, 由于 MSE 损失对异常点极为敏感, 为了更好地保留瑕疵的纹理以及衡量生成的瑕疵与真实瑕疵间的差异, 本文在部分卷积网络原有的损失函数基础上

增加了 MSE 损失用于减少生成的瑕疵纹理的损失, 保证生成的瑕疵纹理完整。MSE 损失可以根据式 (7) 计算:

$$L_{\text{MSE}} = \frac{\sum_{n=0}^{n-1} (\Psi_n^{f_{\text{comp}}} - \Psi_n^{f_{\text{gt}}})^2}{n} \quad (7)$$

其中, n 表示瑕疵的像素总数, f_{comp} 和 f_{gt} 分别表示真实的瑕疵与生成的瑕疵, Ψ_n 表示每个像素的值。

网络的总损失函数由 MSE 损失、感知损失、风格损失、L1 损失和 TV 损失共同构成。MSE 损失计算生成的瑕疵与真实瑕疵之间的纹理损失, L1 损失分为瑕疵部分图像像素和非瑕疵部分布匹图像像素的损失, 风格损失分为重建的带瑕疵位置信息的瑕疵布匹图像与生成的带瑕疵位置信息的瑕疵布匹图像之间的损失 $L_{\text{style}_{\text{out}}}$ 和带瑕疵位置信息的瑕疵布匹图像与原本不带瑕疵布匹图像之间的损失 $L_{\text{style}_{\text{comp}}}$, 感知损失 $L_{\text{perceptual}}$ 计算原本无瑕疵的布匹图像与生成的带瑕疵位置信息的瑕疵布匹图像在高层的特征值之间的 L1 距离。在对 100 张验证图像进行实验的基础上, 设置了用于控制每个损失项重要性的超参数。网络的最终损失函数 L_{total} 如下所示:

$$L_{\text{total}} = L_{\text{MSE}} + L_{\text{valid}} + 6L_{\text{hole}} + 0.05L_{\text{perceptual}} + 120(L_{\text{style}_{\text{out}}} + L_{\text{style}_{\text{comp}}}) + 0.1L_{\text{TV}} \quad (8)$$

4 实验结果

实验中, 为了有效地评估本文提出的布匹图像生成模型 EC-PConv 生成的瑕疵布匹图像的质量, 使用大小为 512×512 的灰度图像作为测试图像。实验中使用的是互联网上由华为举办布匹图像瑕疵检测天池大赛的公开数据集。共有 9 576 张布匹图像, 其中有瑕疵图片 5 913 张, 无瑕疵图片 3 663 张, 包含的瑕疵布匹图像都是从工业生产中采集并标注好瑕疵位置信息。本文从瑕疵布匹图像数据集中随机选取 700 张破洞和三丝这两种类型的瑕疵作为训练集; 再从剩余的图片中随机选取 700 张同样类型的瑕疵布匹图像样本对作为测试集。并对该数据集进行了预处理。实验数据集样例图如图 6 所示。

网络模型使用的是深度学习框架 PyTorch, 使用的实验设备是单块英伟达 2080Ti 显卡, 训练中使用 CUDA 进行加速。网络使用 Adam 优化器, 初始学习率设定为 2×10^{-4} , 每隔 50 代对学习率进行一次更新, 更新后的学习率是上一次的 0.9 倍, 训练总代数为 3 000 代。图像

生成模型与识别分类任务不同,不同图片之间相对独立,因此本文实验设置 $\text{batch size}=1$,生成的带瑕疵位置标注的瑕疵布匹图像彼此独立更利于实验中的瑕疵检测.

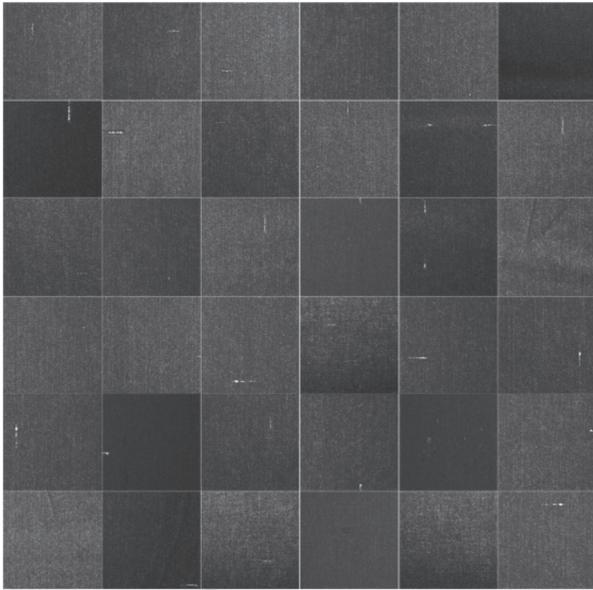


图6 实验数据集样例图

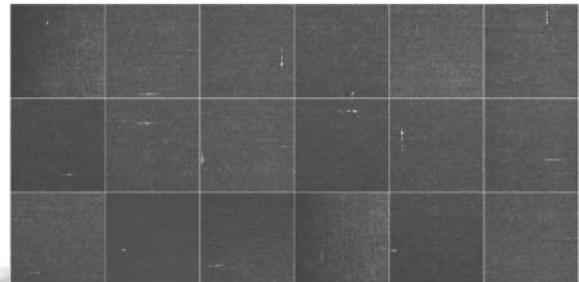
4.1 结果分析

目前,图像生成算法都从两个方面进行分析,一是图片自身的质量,生成的图像是否逼真,是否清晰,内容是否完整;二是多样性,在图像生成任务中,生成网络所产生的图片是多种多样的,如果只产生一种或者几种类型的图像,将会出现模式崩溃 (mode collapse) 的现象.在进行图像评价时,一般使用 IS (inception score) 和 FID (Frechet inception distance) 两种常用的评价指标.但 IS 评价标准的 Inception V3 权重是在 ImageNet 下训练出来的,而本文中使用的其他数据集生成图像,这种跨数据集地计算 IS 是错误的,所以在利用 ImageNet 以外的数据集计算 IS 时,得出的结果没有意义.为了克服 IS 评价的缺点,本实验中使用 FID 进行评价.

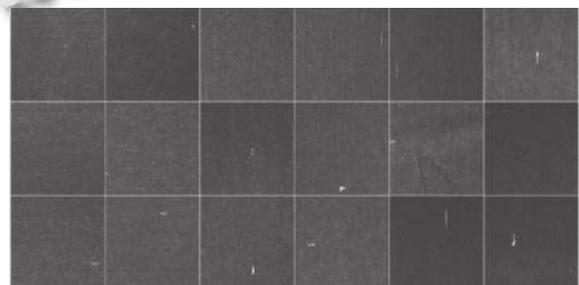
图 7(a) 为真实的瑕疵布匹图像,图 7(b) 为生成的瑕疵布匹图像.从图 7 中可以看出,本文的方法生成的瑕疵布匹图像纹理比较清晰,图像识别度也比较高,与真实图片差距较小.

在图 8 中可直观地看到不同模型生成的瑕疵布匹图像,图 8(a) 是 EC-PConv 生成的瑕疵布匹图像,图 8(b) 是 WGAN-GP^[23] 生成的瑕疵布匹图像,图 8(c) 是 BigGAN^[24] 生成的瑕疵布匹图像.实验结果表明,由 WGAN-GP、BigGAN 和本文提出的方案生成的瑕疵

布匹图像在整体视觉外观上没有显著差异,但从细节上仍然可以看出,由 WGAN-GP 和 BigGAN 生成的瑕疵布匹图像瑕疵部分会有细微的色差,图像整体较模糊等问题.相反,本文的模型生成的结果没有颜色偏差,图像整体较清晰.



(a) 真实的瑕疵布匹图像

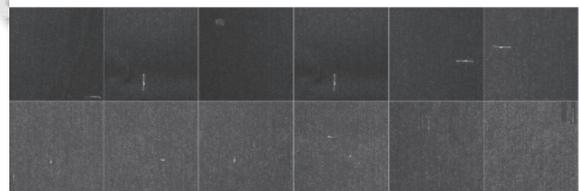


(b) 生成的瑕疵布匹图像

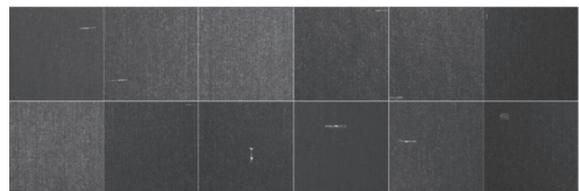
图7 真实和生成的瑕疵布匹图像



(a) EC-PConv



(b) WGAN-GP



(c) BigGAN

图8 不同生成模型生成瑕疵布匹图像

本实验中使用 FID 进行评价不同生成模型在生成瑕疵布匹图像任务上的表现. FID 是计算真实瑕疵布

匹图像和生成瑕疵布匹图像的特征向量距离的一种度量,特征向量由 Inception V3 network 所得到,它使用的是 Inception V3 network 分类网络倒数第 2 个全连接层的输出的 $1 \times 1 \times 2048$ 维图像特征向量进行距离度量.这个距离可以衡量真实图像和生成图像之间的相似程度, FID 值越小,相似程度越高,在 FID=0 时,两个图像相同.

从表 1 的实验结果可以看出,在同样的瑕疵布匹图像数据集作为训练集的情况下,本文提出的改进的部分卷积网络的带瑕疵位置信息的瑕疵布匹图像模型 EC-PConv 的 FID 分数较好.这表明,该模型在瑕疵布匹图像生成领域优于现有的先进图像生成模型.现有的生成模型都是通过神经网络学习图像特征分布,再从分布中进行采样来生成新的图像,因此生成的图像往往有一定程度的模糊,又因为瑕疵布匹图像中的瑕疵尺寸较小,在经过较深的网络进行卷积后,瑕疵特征会丢失,导致生成图像时生成的瑕疵部分模糊不清等问题.本文提出的基于改进的部分卷积网络的带瑕疵位置信息的瑕疵布匹图像生成框架,使用较浅的网络提取出瑕疵的特征,再利用部分卷积层更新 mask 的特性将瑕疵“修复”到原本不带瑕疵的布匹图像上,而且已知瑕疵的位置信息.

表 1 不同生成模型生成瑕疵布匹图像的 FID 值

方法	FID
GAN	8.95
IGAN	7.87
WGAN-GP	8.03
BigGAN	7.64
EC-PConv	7.59

在实际的应用场景中,只是生成瑕疵布匹图像无法满足基于深度学习的布匹瑕疵检测模型的需求,生成的瑕疵布匹图像没有标注瑕疵位置信息,依旧需要人工对瑕疵进行标注.与现有生成模型不同,本文提出的模型 EC-PConv,不仅能够生成瑕疵布匹图像,而且生成的图像带有瑕疵位置信息,无需进行人工标注. DeblurGAN^[25] 中引入一种新的评价基准,即辅助提升其他任务在模糊图像上的性能(如目标检测).因此,为了评价本文提出的模型 EC-PConv 生成的带瑕疵位置信息的瑕疵布匹图像的质量,使用基于 YOLO 的布匹瑕疵检测模型来辅助评价生成图像的质量.分别将生成的瑕疵布匹图像,以及真实瑕疵布匹图像作为验证集,输入到经过真实瑕疵布匹图像训练过的布匹瑕疵检测模型中,从表 2 中可以看出,生成的带瑕疵位置信息的瑕疵布匹图像查准率 P (precision) 以及 MAP

(mean average precision) 值比起真实瑕疵布匹图像分别提升了 0.118 和 0.106,而查全率 R 和真实图像相差较小,可能是因为本文提出的基于改进的部分卷积网络的带瑕疵位置信息的瑕疵布匹图像生成模型目前能够生成的瑕疵布匹图像只有破洞和三丝两种,在进行查全率计算时无法对应所有类型的瑕疵,从实验数据分析可以得出 EC-PConv 模型生成的瑕疵布匹图像逼真程度非常高.

表 2 生成图像作验证集对比分析

图像	P	R	MAP
真实	0.807	0.767	0.707
生成	0.925	0.759	0.813

再将真实瑕疵布匹图像作为验证集,分别输入到经过真实瑕疵布匹图像训练过的布匹瑕疵检测模型,以及经过生成瑕疵布匹图像训练过的布匹瑕疵检测模型中.从表 3 中可以看出,使用不同瑕疵布匹图像数据集训练的布匹瑕疵检测模型查准率 P,查全率 R 以及 MAP 值分别提升了 0.019、0.027 和 0.085,实验数据表明 EC-PConv 模型生成的瑕疵布匹图像可以作为工业生产中布匹瑕疵检测模型训练集进行使用.实验结果数据表明了本研究在带瑕疵位置信息的瑕疵布匹图像生成问题上初步取得了较好的效果,生成的图像具有良好的迁移能力.

表 3 生成图像作训练集对比分析

模型	P	R	MAP
真实图像训练	0.896	0.785	0.742
生成图像训练	0.915	0.812	0.827

5 结论与展望

文中针对深度学习中布匹瑕疵检测网络训练时,带瑕疵位置信息的瑕疵布匹图像不足的问题,基于改进的部分卷积网络提出了一种带瑕疵位置信息的瑕疵布匹图像生成网络模型 EC-PConv.通过对 mask 添加条件,使网络能够在原本没有瑕疵的布匹图像上,生成位置信息已知的破洞和三丝这两种瑕疵,达到生成带瑕疵位置信息的瑕疵布匹的目的.布匹的瑕疵种类较多,但破洞和三丝具有较为明显的纹理特征,本文提出的瑕疵特征提取网络能够最大程度地提取出瑕疵原本的纹理特征,而类似于污渍等特征不明显的瑕疵,还不能很好地提取出对应的主要特征.

本文在自制的瑕疵布匹图像数据集上进行训练和测试,使用参考图像评价指标对实验结果进行评价,生成

的破洞与三丝这两种类型的瑕疵布匹图像的 FID 值低于现有的图像生成模型,且生成的破洞和三丝两种瑕疵布匹图像在瑕疵布匹检测模型中 MAP 值和 P 值相较于真实瑕疵布匹图像有了提升,证明了本文方法的有效性。

后续工作中,模型还需要扩充瑕疵布匹图像的种类,继续对其他类型的瑕疵布匹图像生成进行研究,提高模型的泛用性和实用性。

参考文献

- 1 Tian YN, Yang GD, Wang Z, *et al.* Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Computers and Electronics in Agriculture*, 2019, 157: 417–426. [doi: 10.1016/j.compag.2019.01.012]
- 2 Isola P, Zhu JY, Zhou TH, *et al.* Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 5967–5976.
- 3 Ghosh A, Kulharia V, Namboodiri V, *et al.* Multi-agent diverse generative adversarial networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 8513–8521.
- 4 Karras T, Aila T, Laine S, *et al.* Progressive growing of gans for improved quality, stability, and variation. 6th International Conference on Learning Representations. Vancouver: ICLR, 2017. 1–26.
- 5 Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. 4th International Conference on Learning Representations. San Juan: ICLR, 2015. 1–16.
- 6 Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 4396–4405.
- 7 张德浩, 王佳松, 陈禹平, 等. 基于 VAE-MSGAN 网络的复杂细节图像生成方法. *机电工程技术*, 2021, 50(6): 29–33. [doi: 10.3969/j.issn.1009-9492.2021.06.008]
- 8 肖儿良, 林化溪, 简献忠. 基于生成对抗网络探索潜在空间的医学图像融合算法. *信息与控制*, 2021, 50(5): 538–549. [doi: 10.13976/j.cnki.xk.2021.0576]
- 9 杨祖义, 张俊松. 融合主题和视觉特征的图片拼贴画合成方法. *计算机辅助设计与图形学学报*, 2019, 31(8): 1403–1409.
- 10 张兴园, 黄雅平, 邹琪, 等. 基于草图纹理和形状特征融合的草图识别. *自动化学报*, 2020, 45: 1–10. [doi: 10.16383/j.aas.c200070, 2020-05-18.]
- 11 肖晨晨, 陈乐庚, 王书强. 融合残差和对抗网络的跨模态 PET 图像合成方法. *计算机工程与应用*, 2022, 58(1): 218–223. [doi: 10.3778/j.issn.1002-8331.2008-0325]
- 12 Chaki J, Dey N. Texture feature extraction techniques for image recognition. Singapore: Springer, 2020.
- 13 Gatys LA, Ecker AS, Bethge M. Texture synthesis using convolutional neural networks. *Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal: ACM, 2015. 262–270.
- 14 Zheng K, Cheng Y, Kang XJ, *et al.* Conditional introspective variational autoencoder for image synthesis. *IEEE Access*, 2020, 8: 153905–153913. [doi: 10.1109/ACCESS.2020.3018228]
- 15 谈馨悦, 何小海, 王正勇, 等. 基于 Transformer 交叉注意力的文本生成图像技术. *计算机科学*, 2022, 49(2): 107–115.
- 16 韦哲, 李从利, 沈延安, 等. 基于两阶段模型的无人机图像厚云区域内容生成. *计算机学报*, 2021, 44(11): 2233–2247. [doi: 10.11897/SP.J.1016.2021.02233]
- 17 LeCun Y, Bottou L, Bengio Y, *et al.* Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, 86(11): 2278–2324. [doi: 10.1109/5.726791]
- 18 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe: ACM, 2012. 1097–1105.
- 19 Kingma DP, Welling M. Auto-encoding variational Bayes. 2nd International Conference on Learning Representations. Banff: ICLR, 2014. 1–14.
- 20 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. *Proceedings of the 27th International Conference on Neural Information Processing Systems*. Montreal: ACM, 2014. 2672–2680.
- 21 Liu GL, Reda FA, Shih KJ, *et al.* Image inpainting for irregular holes using partial convolutions. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 89–105.
- 22 Mirza M, Osindero S. Conditional generative adversarial nets. arXiv:1411.1784, 2014.
- 23 Gulrajani I, Ahmed F, Arjovsky M, *et al.* Improved training of wasserstein gans. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: ACM, 2017. 5769–5779.
- 24 Brock A, Donahue J, Simonyan K. Large scale GAN training for high fidelity natural image synthesis. 7th International Conference on Learning Representations. New Orleans: ICLR, 2019. 1–35.
- 25 Kupyn O, Budzan V, Mykhailych M, *et al.* Deblurgan: Blind motion deblurring using conditional adversarial networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 8183–8192.

(校对责编: 孙君艳)