

[自控·检测]

DOI:10.3969/j.issn.1005-2895.2023.01.011

# 基于语义生成与语义分割的机织物缺陷检测方法

马浩然, 张团善, 王峰, 赵浩铭

(西安工程大学机电工程学院, 陕西 西安 710048)

**摘要:**针对织物疵点的语义分割任务中因数据集规模限制,而导致网络出现的严重过拟合问题,课题组提出了针对织物的语义生成网络。语义生成网络使用随机产生的语义标签生成对应的织物缺陷图像,相较于传统数据增强方法,语义生成可生成全新图像,更贴近实际缺陷分布,并且可通过判别器对生成图像进行筛选;课题组将语义生成的图像作为语义分割网络的输入,相应的随机语义标签作为目标,免去标注过程,扩充语义分割网络的训练样本,提升网络性能;对于语义分割网络,提出尺寸自适应Dice损失函数,解决样本不平衡问题,提升网络对小尺寸的检测能力。实验结果表明:尺寸自适应Dice损失函数使得模型精度提高11.1%,使用BEGAN扩充的数据集相较于传统方法扩充的数据集训练得到的模型精度提高7.4%。

**关键词:**机织物缺陷检测;语义分割;语义生成网络;尺寸自适应Dice损失函数;BEGAN;免标注  
**中图分类号:**TS107;TH165 **文献标志码:**A **文章编号:**1005-2895(2023)01-0066-08

## Woven Fabric Defect Detection Method Based on Semantic Generation and Semantic Segmentation

MA Haoran, ZHANG Tuansan, WANG Feng, ZHAO Haoming

(School of Mechanical and Electrical Engineering, Xi'an Polytechnic University, Xi'an 710048, China)

**Abstract:** Aiming at the severe overfitting problems caused by the limitation of data set size in the semantic segmentation task of fabric defects, a semantic generation network for fabric which uses randomly generated semantic labels to generate corresponding fabric defect images was proposed. Compared to traditional data enhancement methods, semantic generation can generate brand new images that are closer to the actual defect distribution, and the generated images can be filtered by discriminators. The semantically generated images were used as the input of the semantic segmentation network, and the corresponding random semantic labels were used as the target, so as to eliminate the labeling process, expand the training samples of the semantic segmentation network and improve the network performance. For the semantic segmentation network, the size-adaptive Dice loss function was proposed to solve the sample imbalance problem and improve the detection ability of the network for small sizes. The experimental results show that the size-adaptive Dice loss function improves the model accuracy by 11.1%, and the accuracy of the model obtained by training with the BEGAN expanded dataset is improved by 7.4% compared with that of the dataset expanded by the traditional method.

**Keywords:** woven fabric defect detection; semantic segmentation; semantic generation network; size-adaptive Dice loss function; BEGAN; annotation-free

纺织业的重要性上至航空航天、生物医疗,下至服装家纺、包装材料。织物缺陷是影响织物可用性的关键因素,具有较多缺陷的织物将被判定为次品,影响销

售甚至被销毁,同时织物缺陷率的上升也能反映出纺织机械故障的出现。所以织物缺陷检测是提升织物质,监测纺机状态的重要步骤,快速、准确地检出缺陷

收稿日期:2022-05-05;修回日期:2022-11-10

基金项目:西安市现代智能纺织设备重点实验室资助项目(2019220614SYS021CG043)。

第一作者简介:马浩然(1998),男,河南洛阳人,硕士研究生,主要研究方向为数字图像处理与机器视觉。E-mail: MaHaoran\_erssa@stu.xpu.edu.cn

对提升生产效率、提高产品质量都有重要意义<sup>[1]</sup>。

目前工厂中织物缺陷检测主要依靠人工进行,这不仅需要工人具有丰富的工作经验,且劳动强度大,对视力损伤严重。即使如此,人工也仅能达到70%的正确率,这意味着高质量的布匹不得不反复检测,这极大地提高了生产成本。

近年来,计算机技术快速发展,深度学习因其较高的检测正确率,已经逐渐成为织物缺陷检测的常用方法,使用较多的任务模式有分类、目标检测和语义分割<sup>[2-4]</sup>。分类任务仅能判别当前整张图像是否含有缺陷,并对缺陷分类,这种方式检测较为粗略,无法提供具体的缺陷位置;目标检测可以提取缺陷的中心位置和外接矩形的大小,能够更加精细地提供缺陷的信息,但无法提供缺陷的轮廓信息;语义分割可以提供织物缺陷的种类、位置、准确轮廓以及像素级精度的缺陷面积,对缺陷的描述较为全面<sup>[5]</sup>。因此语义分割任务在织物缺陷检测中的使用频率越来越高。

深度学习作为有监督学习,最为重要的是数据集的规模,数据集的规模越大、覆盖越广,越能反映出真实状态下的缺陷分布。一个全面的数据集对于获得一个泛用的、鲁棒的模型十分重要。在织物缺陷检测方面,现有的公共数据集数据量较少,且织物的种类也在日益增多,往往需要对拥有的有限的数据进行扩充<sup>[6]</sup>。传统的扩充方式采用翻转、旋转和缩放等方式,但这些方式只是对现有图像的处理,不能产生新的样本图像。

生成对抗网络(generative adversarial networks, GAN)<sup>[7]</sup>常常被用来扩充数据集,传统的GAN输入服从某种分布的噪声,输出生成的图像,这种方式可以产生新的图像,但这些图像仍需人工标注,这增加了工作量。传统的GAN存在难以训练、模式崩溃等问题,作为数据扩充工具,传统GAN在使用上存在困难。

近年来,语义生成作为GAN新的研究方向,广受关注,语义生成采用编码-解码器的方式,输入语义标签,输出与之对应的真实图像<sup>[8]</sup>。语义生成的提出使得语义分割的数据增强不再需要手工标注,仅需要随机产生语义标签,或对现有的语义标签进行数据增强,即可获得语义标签与对应的真实图像的一组数据,免去标注过程,可以在语义分割网络训练过程中实时

增强。

传统GAN、深度卷积生成对抗网络(deep convolution generative adversarial networks, DCGAN)中<sup>[9]</sup>,库尔贝特-莱布勒散度(Kullback-Leibler divergence),简称KL散度,和杰森-香农散度(Jensen-Shannon divergence),简称JS散度,在生成分布与真实分布不重叠时会造成梯度消失,导致网络无法收敛。瓦瑟斯坦生成对抗网络(Wasserstein generative adversarial networks, WGAN)<sup>[10]</sup>使用推土机距离(earth mover's distance, EM距离)替代KL散度和JS散度,EM距离在2分布无重叠时也能良好地衡量真实分布与生成分布的距离。边界平衡生成对抗网络(boundary equilibrium generative adversarial networks, BEGAN)将判别器设置为编码-解码器的形式,通过重构误差的分布的逼近来代替真实分布的逼近,提升了训练速度,并且通过设置生成器与判别器损失的反馈机制来避免判别器过强导致的训练崩溃。

菲利普等<sup>[11]</sup>发现,在GAN的生成器训练中加入一些传统损失函数( $L_1$ 损失函数)可以辅助训练,仅使用判别器的输出作为生成器的损失函数会导致生成的图像缺乏真实,仅使用 $L_1$ 损失函数会导致生成的图像十分模糊,二者融合可以更好地生成图像。

课题组采用BEGAN的网络结构<sup>[12]</sup>,同时在生成器训练时添加加权的 $L_1$ 损失函数作为辅助, $L_1$ 损失函数使用尺寸自适应加权函数进行加权,减少大尺寸缺陷模糊的问题。BEGAN同样采用EM距离评估生成分布和真实分布的差距,这可以避免KL散度和JS散度带来的梯度消失。课题组选用多样性比率 $\gamma$ (diversity ratio),用于控制生成网络在多样性和生成质量中的比例,同时比例控制原理(proportional control theory)的加入用于控制判别器对生成样本和真实样本重构误差期望维持在 $\gamma$ 附近,从而避免判别器对生成样本过于严格而导致的训练崩溃。

## 1 语义生成

生成网络主要包括网络原理的构造和网络结构及参数的选择。生成对抗网络采用BEGAN的形式,生成器与判别器都采用编码-解码器的形式,判别器获取生成样本与真实样本的重构误差,损失函数采用EM

距离,采用  $L_1$  损失函数辅助生成器的训练。

### 1.1 原理及损失函数设计

传统 GAN 的优化目标分为 2 部分:判别器  $D(x_i)$  与生成器  $G(z)$  部分。

1) 判别器部分。

$$\left. \begin{aligned} E_{x_i \sim P_r} \ln D(x_i); \\ E_{x_i \sim P_G} \ln [1 - D(x_i)]。 \end{aligned} \right\} \quad (1)$$

式中: $x_i$  为采样样本, $P_r$  是真实图像分布, $P_G$  是生成器生成的图像分布。

对于判别器,在输入真实图像时,判别器  $D(x_i)$  的期望输出越大越好,即  $E_{x_i \sim P_r} \ln D(x_i)$  越大越好;在输入生成图像时,判别器  $D(x_i)$  的期望输出越小越好,即  $E_{x_i \sim P_G} \ln [1 - D(x_i)]$  越小越好。

2) 生成器部分。

$$E_{x_i \sim P_G} \ln [1 - D(x_i)]。 \quad (2)$$

对于生成器的优化目标,在对判别器输入为生成图像时,判别器的输出期望越大越好,即  $E_{x_i \sim P_G} \ln [1 - D(x_i)]$  越小越好。

定义  $V(D, G)$  为优化目标,则有:

$$\min_G \max_D V(D, G) = E_{x_i \sim P_r} \ln D(x_i) + E_{x_i \sim P_G} \ln [1 - D(x_i)], x_i \in \prod(P_r, P_G)。 \quad (3)$$

式中: $P_G = G(z), z \sim N(\mu, \sigma^2)$ ;  $V(D, G)$  为关于判别器与生成器的优化函数;  $\prod(P_r, P_G)$  是真实样  $P_r$  与生成样本  $P_G$  的联合分布; $z$  是符合某种分布的噪声,往往是正态分布。

在固定生成器  $G(z)$  时,最优的判别器

$$D^*(x_i) = \frac{P_r(x_i)}{P_r(x_i) + P_G(x_i)}。 \quad (4)$$

当判别器最优时,生成器的优化目标为:

$$\min_G \left( E_{x_i \sim P_r} \ln \frac{P_r(x_i)}{[P_r(x_i) + P_G(x_i)]} + E_{x_i \sim P_G} \ln \frac{P_G(x_i)}{[P_r(x_i) + P_G(x_i)]} \right)。 \quad (5)$$

引入 KL 散度和 JS 散度:

$$\text{div}_{\text{KL}}(P_1 \parallel P_2) = E_{x_i \sim P_1} \ln (P_1/P_2)。 \quad (6)$$

$$\begin{aligned} \text{div}_{\text{JS}}(P_1 \parallel P_2) &= \frac{1}{2} \text{div}_{\text{KL}}(P_1 \parallel \frac{P_1 + P_2}{2}) + \\ &\frac{1}{2} \text{div}_{\text{KL}}(P_2 \parallel \frac{P_1 + P_2}{2})。 \end{aligned} \quad (7)$$

式中  $P_1$  与  $P_2$  为不同的样本分布。

将生成器优化目标写成 JS 散度的形式:

$$2 \text{div}_{\text{JS}}(P_r \parallel P_G) - 2 \ln 2。 \quad (8)$$

JS 散度存在弊端,当真实样本分布  $P_r$  与生成样本分布  $P_G$  分布完全不重叠时,JS 散度恒为  $\ln 2$ 。即无论  $P_r$  与  $P_G$  的分布多近或多远,JS 散度为定值,这导致了梯度消失;而在实际情况中, $P_r$  与  $P_G$  的分布完全不重叠的情况十分常见,JS 散度存在的弊端直接导致了原始的 GAN 难以训练。

EM 距离  $W(P_r, P_G)$  可以更好地衡量 2 个分布的差距。

$$W(P_r, P_G) = \inf_{\sigma \in \prod(P_r, P_G)} E_{(x_r, x_G) \sim \sigma} \|x_r - x_G\|。 \quad (9)$$

对于联合分布样本  $\sigma$ ,从中采样真实样本  $x_r$  和生成样本  $x_G$ ,计算 2 者距离,取下界,定义为 EM 距离。

BEGAN 将样本分布的优化转化为样本重构误差  $L$  的优化,即:

$$L(x_i) = |x_i - D(x_i)|^\eta。 \quad (10)$$

式中  $\eta \in [1, 2]$ 。

除此之外,BEGAN 还引入了多样性比例  $\gamma$  和比例控制原理,使得网络可以调节多样性和生成质量。

$$\left. \begin{aligned} L_D &= L(x_r) - k_t L(x_G); \\ L_G^c &= L(x_G); \\ k_{t+1} &= k_t + \lambda_k (\gamma L(x_r) - L(x_G))。 \end{aligned} \right\} \quad (11)$$

式中: $L_D$  为判别器损失; $L_G^c$  为生成器损失; $k_t$  为第  $t$  次循环中的  $k$  值, $k_0 = 0$ ;  $z_D$  为训练判别器  $D$  时生成器的输入; $\lambda_k$  为  $k_t$  的调节率; $\gamma \in [0, 1]$ ,通过调整  $\gamma$  值来调节生成多样性和生成质量。

在训练中,添加加权的  $L_1$  损失函数用以辅助生成器训练。若织物存在一些小尺寸缺陷,如破洞、棉球等,而  $L_1$  损失函数在所有像素上取均值,这极大地稀释了小尺寸缺陷的损失值占比,导致小尺寸的缺陷不易训练。故课题组采用了一个尺寸自适应函数对  $L_1$  损失函数进行加权,并添加了一个调节权重  $\alpha$  用于调节  $L_1$  损失函数  $L_G^L$  占比。

$$\begin{aligned} R_d &= \frac{y_d}{y_d + y_b}; \\ \mu &= \begin{cases} 0.275 \cdot \sin \left( 10 \cdot \pi R_d - \frac{\pi}{2} \right) + 0.425, & 0 \leq \mu \leq 0.1, \\ 0.7, & 0.1 < \mu \leq 1.0; \end{cases} \end{aligned}$$

$$L_C^1 = \alpha(\mu |g_d - y_d| + (1 - \mu) |g_b - y_b|).$$

式中: $R_d$  为标签中缺陷像素占比, $y_d$  为缺陷样本, $y_b$  为背景样本, $\mu$  为尺寸自适应权重, $g_d$  为生成缺陷样本, $g_b$  为生成背景样本, $\alpha$  为调节权重。

## 1.2 网络结构

生成器与判别器均采用编码-解码器的形式。编码器的网络结构如图 1 所示,输入经过纯卷积网络进行下采样,最终到  $15 \times 15$  像素大小。传统的最大池化下采样不适用于语义生成任务,抑制低激活值不利于

网络推导织物纹理。下采样部分采用卷积进行,卷积核  $4 \times 4$ ,步长为 2。

解码器的作用主要是还原图像分辨率以及产生纹理。传统网络上采样为避免训练初期出现的周期性噪声而回避使用反卷积,转而采用双线性插值进行上采样。而在织物缺陷生成任务中,训练初期反卷积的周期性噪声的出现有利于织物底纹的推理,故采用反卷积进行上采样,同时串联一个  $3 \times 3$  的卷积进一步产生纹理。解码器网络结构如图 2 所示。

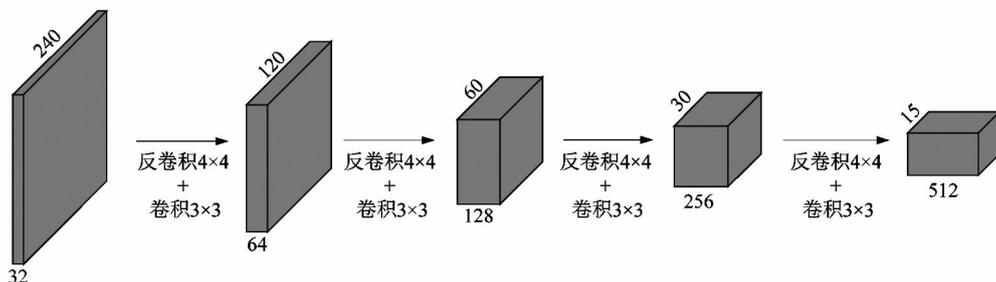


图 1 编码器网络结构

Figure 1 Encoder structure

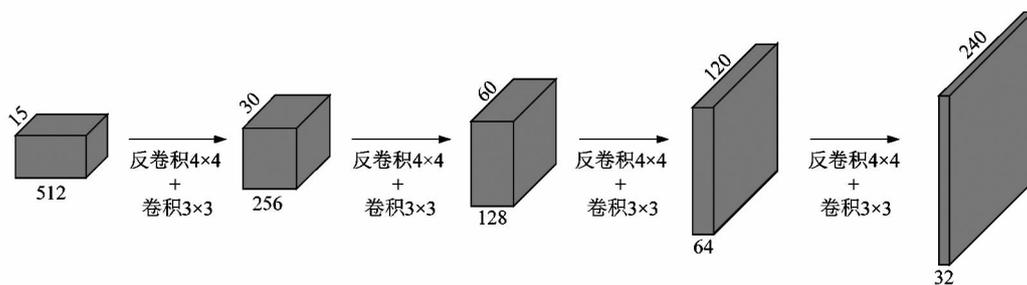


图 2 解码器网络结构

Figure 2 Decoder structure

生成器与判别器采用相同的结构,唯一不同的是生成器的输入维度  $m = 6$ ,而判别器的输入维度  $m = 3$ 。模型添加了大跨度的特征融合,编码器的特征与其尺寸匹配的解码器特征相加。每个卷积层后添加 BN 层,使用 ELU 作为激活函数。整体采用大跨度的特征连接用于补充特征损失,辅助轮廓推导。图 3 所示为整体网络结构。

## 1.3 参数选择

在 BEGAN 的训练过程中,采用 Adam 优化器, $\beta_1 = 0.500, \beta_2 = 0.999$ 。为确保生成样本的多样性,在 BEGAN 的损失函数中,多样性比例  $\gamma = 1, \lambda_k = 0.001$ 。

在生成器训练过程中,图 4 中展示了不同  $\alpha$  下的生成结果。在  $\alpha = 10.0$  时,生成图像清晰度低,真实性较差;在  $\alpha = 1.0$  时,清晰度提高; $\alpha = 0.1$  时,生成图像清晰,真实性好。过高的  $L_1$  损失函数权重  $\alpha$  将导致生成网络多样性降低,且生成的图像较为模糊,故取  $\alpha = 0.1$ 。

## 2 语义分割

语义分割网络同样采用编码-解码器的形式,采用带空洞卷积的空间池化金字塔(atrous spatial pyramid pooling, ASPP)结构<sup>[13]</sup>,最小下采样尺寸为  $30 \times 30$  像素,这是保持织物缺陷不被缩放的最小下采样尺寸。设置尺寸自适应函数对 Dice 损失函数进行加权。



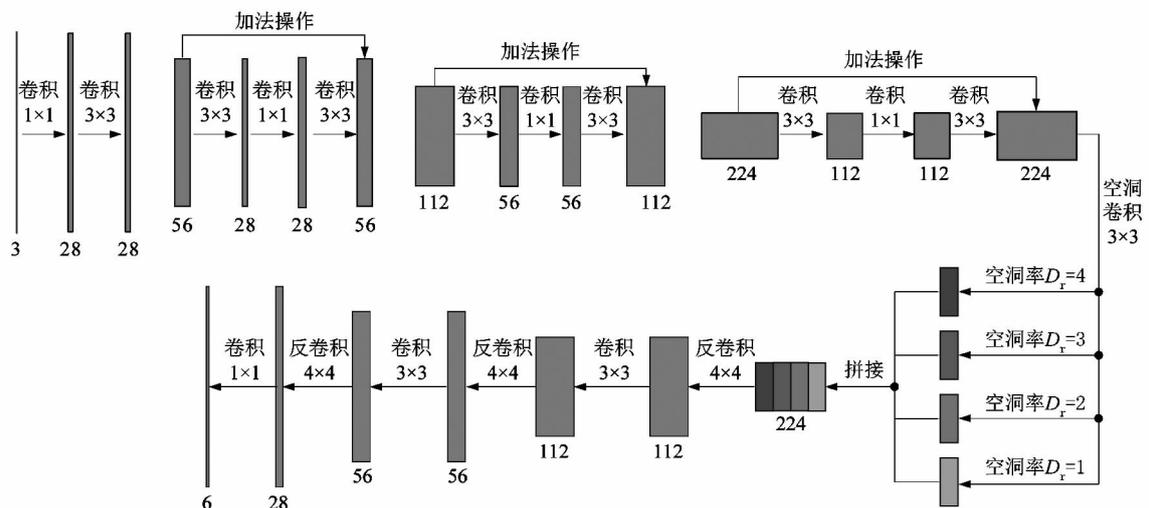


图 5 语义分割网络结构

Figure 5 Semantic segmentation network structure

### 3 实验结果

BEGAN 在数据集上使用尺寸自适应  $L_1$  损失函数与原始  $L_1$  损失函数各训练了 1 800 次,使用测试集并进行数据扩充,增加缺陷多样性。语义分割网络在原始数据集上,使用原始 Dice 损失函数与尺寸自适应 Dice 损失函数各训练了 1 000 次。在使用 BEGAN 扩充的数据集上,使用尺寸自适应 Dice 损失函数训练 1 000 次。将结果进行对比分析。

#### 3.1 BEGAN 实验结果

图 6 所示为不同训练阶段下的缺陷图片生成图像,可以明显地观察到:使用尺寸自适应损失函数的生成结果优于原始  $L_1$  损失,使用尺寸自适应  $L_1$  损失函数的图像更加清晰,缺陷部分更加真实、清晰。

#### 3.2 语义分割实验结果

图 7 所示为在原始 Dice 损失与尺寸自适应 Dice 损失下,模型训练中验证集精度的对比。由图 7 可以

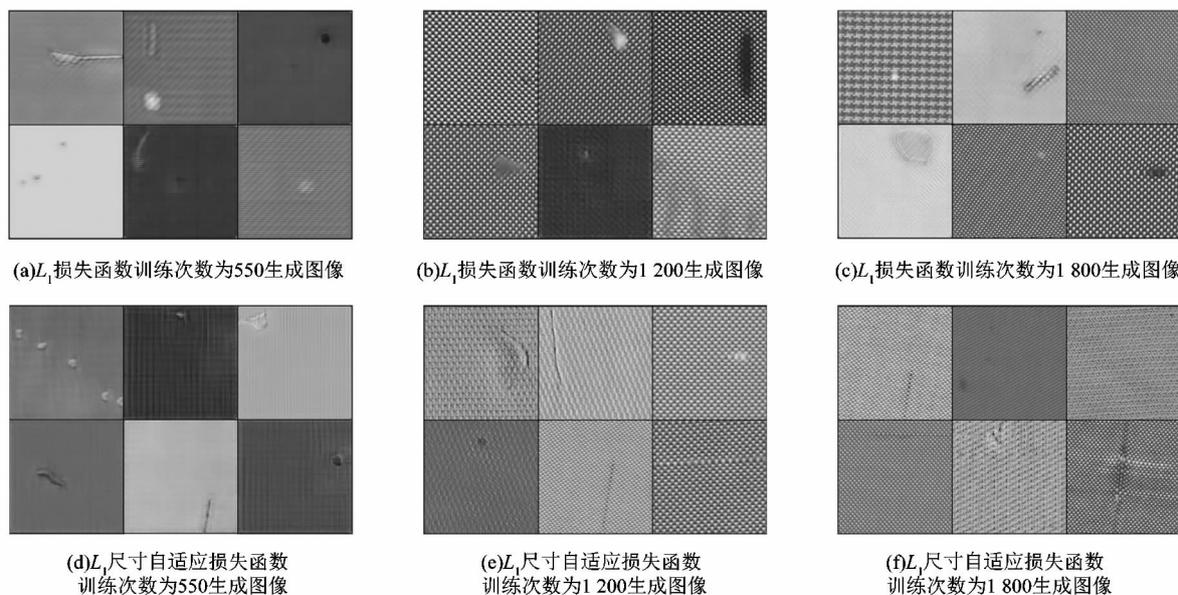


图 6 原始  $L_1$  损失函数与尺寸自适应  $L_1$  损失函数生成图像对比

Figure 6 Comparison of original  $L_1$  loss function and size-adaptive  $L_1$  loss function generated images

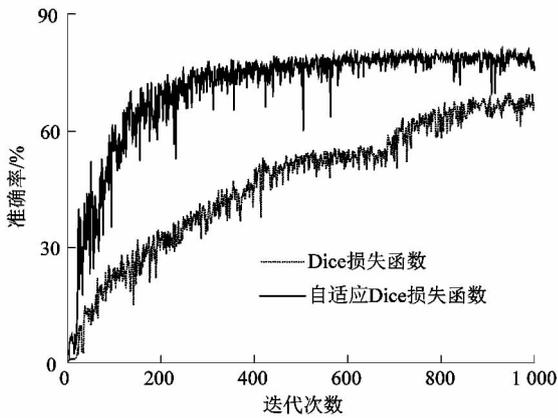


图 7 Dice 损失函数与尺寸自适应 Dice 损失函数模型准确率对比

Figure 7 Accuracy comparison of Dice loss function and size-adaptive Dice loss function models

看出在使用原始 Dice 损失时,模型的收敛出现了明显的滞后。这是因为在原始 Dice 损失下,小尺寸缺陷的损失被稀释,模型难以训练,负样本优先于正样本收敛,最终导致模型训练时间过长。尺寸自适应 Dice 损失函数能很好地解决这一问题,正、负样本损失分开计算可以减轻负样本对正样本的稀释。尺寸自适应函数对不同尺寸缺陷进行自适应权重分配,提高小尺寸缺陷的损失占比,由此解决正样本间不均衡问题。使用尺寸自适应 Dice 损失函数使得模型的检测精度提高了 11.1%。

图 8 所示为使用 BEGAN 扩充数据集并且适应尺寸自适应 Dice 损失函数训练的模型的检测结果。模型可检测多种颜色及纹理的机织物,对缺陷的分割以及分类效果较好,能划分复杂形状缺陷轮廓,对于孔洞、断纱等小尺寸缺陷的检测效果优异。

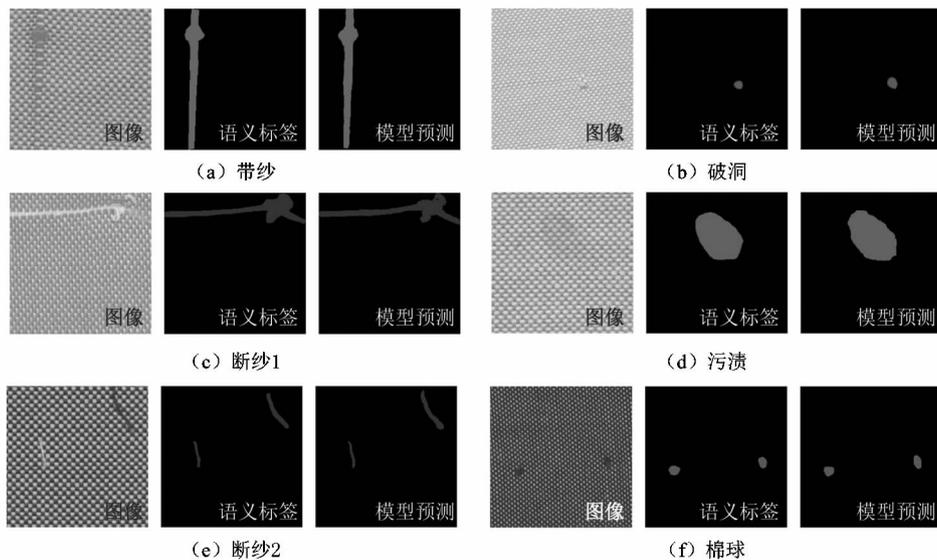


图 8 语义分割网络检测结果

Figure 8 Semantic segmentation network detection results

图 9 所示为使用原始数据集训练出的语义分割网络与使用 BEGAN 扩充过的数据集训练的语义分割网络的性能对比。使用传统数据增强相较于使用原始数据集的提升较小,因为传统的数据增强只是对现有图片进行旋转、缩放、翻转等操作,本质上并未产生新的样本。使用 BEGAN 扩充的数据集更接近缺陷的真实分布,并且生成样本经过判别器筛选,扩充的缺陷图像皆是从随机语义标签重新生成的,缺陷图像更加多样,

使得训练出的模型鲁棒性更好。

均交并比 (mean intersection over union, MIoU) 为语义分割的标准度量,计算所有类别交集和并集之比的平均值:

$$\varphi_{MIoU} = \frac{1}{n + 1} \sum_{i=0}^n \frac{T_p}{F_N + F_p + T_p} \circ$$

式中:  $n$  为预测类别数量;  $T_p$  表示模型预测为正例,实际也为正例,预测正确;  $F_N$  表示模型预测为反例,实际

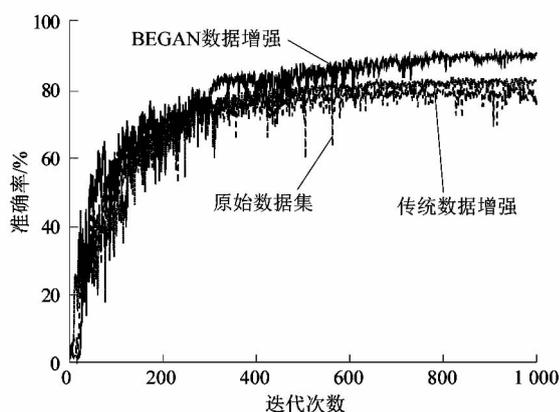


图9 不同数据增强方法对比

Figure 9 Comparison of different data enhancement methods

为正例,预测错误; $F_p$ 表示模型预测为正例,实际为反例,预测错误。

表1所示为不同数据增强方法下,模型的检测精度。使用BEGAN的语义生成方法获得了最好的效果,在测试集上获得90.8%的MIoU均值。

表1 数据增强方法对比

Table 1 Comparison of data enhancement methods

数据增强方法	测试图像张数	$\varphi_{MIoU}/\%$
BEGAN	150	90.8
传统方法	150	83.4
原始数据	150	80.8

## 4 结语

课题组针对织物缺陷检测中的数据缺失设计了语义生成网络,使用BEGAN平衡生成器与判别器,加入尺寸自适应的 $L_1$ 损失函数辅助训练,提升缺陷图像生成质量。语义生成的图像经过判别器筛选后进入扩充数据集,使得缺陷图像真实性更好。相较于传统数据增强方法,语义生成的图像更加真实,且能产出全新样本;相较于传统GAN进行的数据扩充,语义生成可以指定位置、形状和类别进行针对性生成,免去标注过程。

在语义分割方面,提出了尺寸自适应Dice损失函数,该损失函数能够解决样本不均衡问题,使得网络对小缺陷更加敏感,提高检出率。尺寸自适应Dice损失

函数相较于原始的Dice损失函数收敛更快,且最终得到的模型检测精度提高了11.1%。

使用BEGAN的免标注的数据集扩充,与传统数据增强方法进行对比,使用尺寸自适应Dice损失函数对模型进行训练,避免模型收敛滞后。最终在测试集上得到均交并比为90.8%,在GTX 1080Ti的运行速度可达到99帧/s。

## 参考文献:

- [1] 黄乾玮,张团善,周玲,等.基于彩色图像处理与EDLines的数码印花缺陷检测系统[J].轻工机械,2021,39(4):74-79.
- [2] 吕文涛,林琪琪,钟佳莹,等.面向织物疵点检测的图像处理技术研究进展[J].纺织学报,2021,42(11):197-206. DOI:10.13475/j.fzxb.20200702710.
- [3] 王思芝,张团善,刘亚.基于改进Yolo v5的织物缺陷检测方法[J].轻工机械,2022,40(2):54-60.
- [4] 李兰,奚舒舒,张才宝,等.基于DSSD模型的机械加工工件表面缺陷检测研究[J].机电工程,2021,38(2):234-238.
- [5] JING J F, WANG Z, RATSCH M, et al. Mobile-unet: an efficient convolutional neural network for fabric defect detection[J]. Textile Research Journal, 2022, 92(1/2):30-42.
- [6] LIU J H, WANG C Y, SU H, et al. Multistage GAN for fabric defect detection[J]. IEEE Transactions on Image Processing, 2019, 29:3388-3400.
- [7] 陈佛计,朱枫,吴清潇,等.生成对抗网络及其在图像生成中的应用研究综述[J].计算机学报,2021,44(2):347-369.
- [8] 郜仲元,余灵婕,章玉铭,等.基于改进Pix2PixGAN的织物疵点检测算法[J].毛纺科技,2022,50(3):9-13. DOI:10.19333/j.mfkj.20211204805.
- [9] FANG W, ZHANG F H, SHENG V S, et al. A method for improving CNN-based image recognition using DCGAN [J]. Computers, Materials and Continua, 2018, 57(1):167-178.
- [10] 梁俊杰,韦舰晶,蒋正锋.生成对抗网络GAN综述[J].计算机科学与探索,2020,14(1):1-17.
- [11] 张媛媛,张红英.基于pix2pix网络结构的单曝光HDR图像生成方法[J].计算机工程与应用,2022,58(16):242-248.
- [12] 王硕诚,苟刚,葛梦园.基于条件的边界平衡生成对抗网络[J].计算机应用研究,2020,37(5):1514-1517. DOI:10.19734/j.issn.1001-3695.2018.10.0842.
- [13] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs [J]. IEEE transactions on pattern analysis and machine intelligence, 2018, 40(4):834-848.