Abstract

用多阶段的GAN来合成defects到新采集的、没有瑕疵的样本上

- First：一个 texture-conditioned GAN 被训练来生成特定背景的疵点块（这些疵点块最后会被贴到新收集的样本中，然后再丢到二阶段的GAN中去生成一张有疵点的平滑图）
- Second：GAN-based fusion network 融合一阶段生成的疵点块到新样本的特定位置
- Finally：二阶段生成的疵点样本添加到数据集中，训练一个新的semantic segmentation network

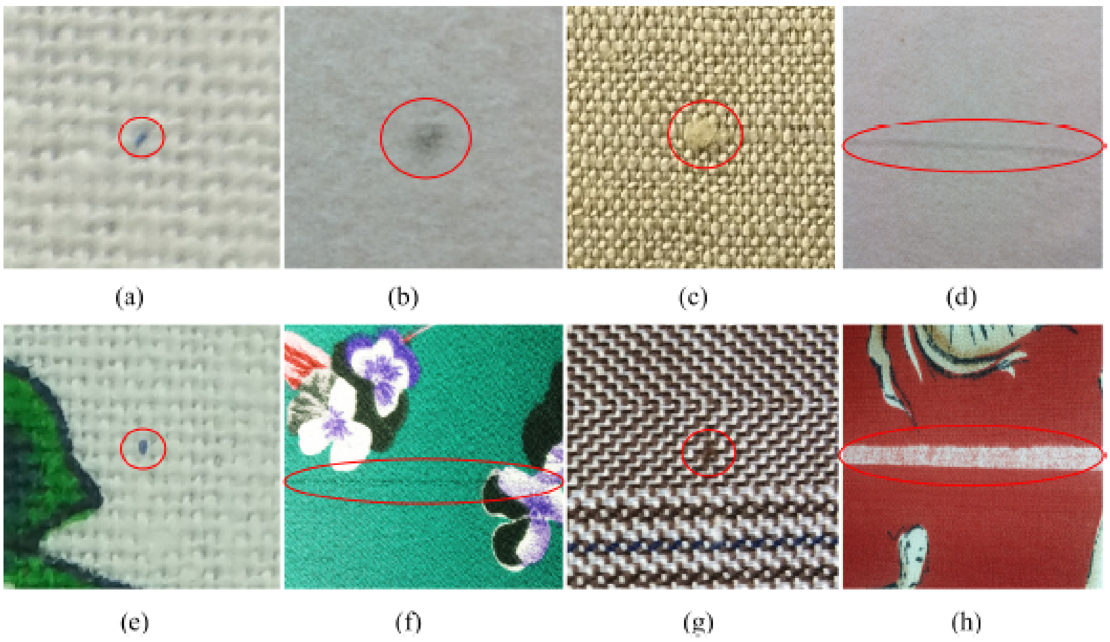

Fig. 1. Defect samples in simple- and complex-textured fabrics. From (a) to (d), the defects are classified as color spots, oil stains, knots, and broken ends in simple texture fabrics. From (e) to (h), the defects are classified as color spots, broken ends, broken yarn, and white strips in complex texture fabrics.

Introduction

介绍一些检测的方法，但是这些方法难以解决复杂背景的布匹

- Fourier transform
- nonlocally centralized sparse representation（NCSR）
- Fisher criterion-based stacked denoising autoencoder（FCSDA）

现实世界中的织物检测方法所面临的困难：

- 大量地标注数据十分消耗时间，织物和织物瑕疵的多样性让收集一个全面的数据集变得十分困难
- 工序和材料的不同，不同瑕疵的外观和特征非常不同

本文

- 假设织物缺陷和背景信息遵循一个条件分布，给定一个织物背景，生成模型的目标是合成defects到织物背景上，因此当处理一种新的织物时，收集和标注数据是不必要的

Related Works

Semantic Segmentation

- CNN采用连续池化和调整步长来增加感受野，从而在特征分辨率降低的同时学习更多abstract features
- FCN用于解决分辨率降低的问题，它用卷积代替了所有的全连接层，然后upsampling 被用在范围减少的特征上从而输出一个密集像素标注图
- SegNet 映射encoder生成的low-resolution的图片给pixelwise predictions，并且使用一个由一些upsampling和一个softmax层组成的decoder在最后。每一个upsamping区拥有一个upsampling layer和数个相继的卷积层，并且他们的最大池化对应encoder的pooling层
- 一些方法集成了全局信息到cnn
  - CRFs
  - atrous convolution
  - feature fusion

Generative Adversarial Networks

- pix2pix for image-to-image translation
- primal GAN aims to explore the mapping relationship from source images to target images
- dual GAN performs the inverse task to primal GAN
- image manipulation tasks:
  - SRGAN for super-resolution
  - ID-CGAN for image de-raining
  - IAN for photo modification
  - Context Encoders for image in-painting.

METHODOLOGY

三种主要方法的细节

- segmentation network
- multi-stage GAN
- fine-tuning strategy

A. Framework Description

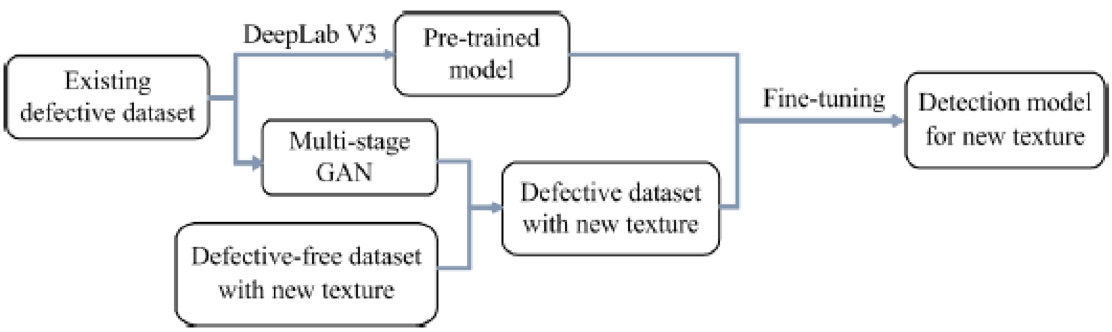

Fig. 2. The architecture of our proposed system for fabric defect detection.

- 现存有的织物数据被用来训练segmentation network（DeepLab V3，修改了一些atrous convolutional layers 的atrous rate，并且加入了一个wegiht parameter）
- 基于现有的数据，合成更多不同织物的缺陷样本

B. Segmentation Network

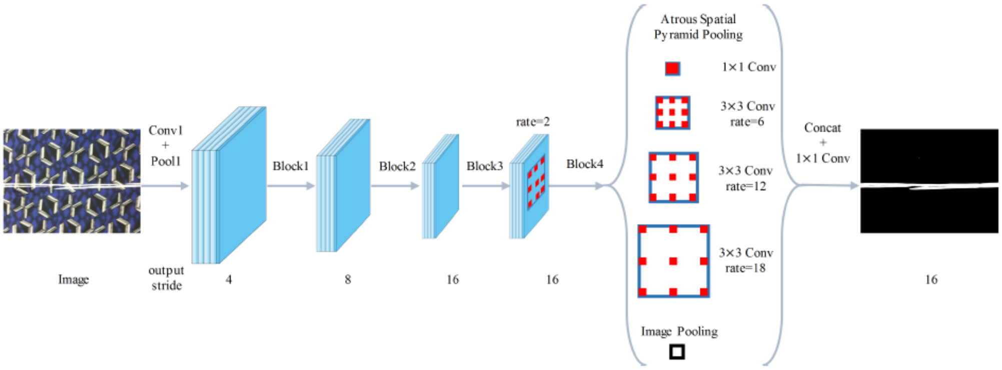

Fig. 3. The architecture of DeepLab V3. Image-level features augmented by parallel atrous convolution modules (ASPP) are included.

- 为了在更深层保留feature maps 的大小，DeepLab 使用atrous convolution 而不是striding
- 使用ResNet-50作为backbone network，然后atrous convolution被用来替代连续的striding，并且atrous rates根据需要的输出步值而设置
- 部署一个 atrous spatial pyramid pooling (ASPP)函数在feature map 的顶部。 函数包含了四个平行的、带有不同atrous rate的 atrous convolutions。为了提取global features, 一个方法是设置一个非常大的atrous rate，但是这样有可能造成filter被下降得更小，从而可能没有捕获到global information。因此，DeepLab给ASPP增加了一个全局信息分支，包含了average pooling， 1 * 1 conv， upsampling。
- 织物检测的三个问题：
    - output stride 不能设置得过大。本文设置为了8而不是16
    - 没有应用CRFs来对边缘进行提炼
    - modifacations to the loss function. add different weights to the defect segmentation to the defect segmentation and the background segmentation($\alpha_1 = 1, \alpha_2 = 9$)

$$L = \alpha_1 L_{background} + \alpha_2 L_{defect}$$

C. Synthesizing Novel Defective Samples Using a Multistage GAN

stage1

- ‘style label’代表新织物样本的背景信息。收集新织物补丁作为stage 1的输入，他们从不带有defect的区域随机裁剪下来
- 一个 pretrained 的 VGGNet被用来提取这些补丁的信息，然后，根据特征计算出Gram 矩阵，Gram矩阵被用作style labels(本文stage1采用的是condition-GAN，CGAN编码了类别信息，将特征当作类别输入CGAN)
    - Gram :

$$G^{\phi}(x) = \frac{\psi\psi^T}{CHW}$$

- 将从原始图裁剪下来的real defect和它的style label 作为一个real pair，而一个生成的defect和它的style label 作为一个fake pair

$$\underset{G}{min}\underset{D_s}{max}L(D_s, G) = \text{Æ}_{x\sim p_{x|G^\phi}}\left[logD_s(x, G^\phi)\right]$$
$$+\text{Æ}_{z\sim p_z}\left[log(1 - D_s(G(z), G^\phi))\right]$$

Stage 2

- 训练集的defective 区域被挖掉，留下一个空白区域。然后生成的defect被resize然后贴到这些空白区域中，从而得到一张不完美的图
- 接下来缺陷融合网络被训练来融合生成的defect到相应的背景中
- 设计三个loss微调网络
  - 由于没有理由假设生成的带缺陷图片与原始图是一样的，因此使用了一个Hinge reconstruction loss

$$\mathcal{L}_{rec} = max(0, ||y - T(x)|| - m)$$

  - impose a constraint on the feature map of fused patches extraced by a pretrained VGGNet and penalize it to be similar to the corresponding real patch. The feature reconstruction loss between the features of real and fake images：

$$l^{\phi,j}_{feat}(T(X), y) = \frac{1}{C_j H_j W_j}||\phi(T(X)) - \phi(y)||^2_2$$

  - adversarial-loss(T()代表defect-fusing network、x和y分别代表不完美和训练样本)

$$\underset{T}{min}\underset{D_T}{max}L(D_t, G) = \text{Æ}_{y\sim p_y}\left[logD_T(y)\right]$$
$$+\text{Æ}_{x\sim p_x}\left[log(1 - D_T(T(x)))\right]$$

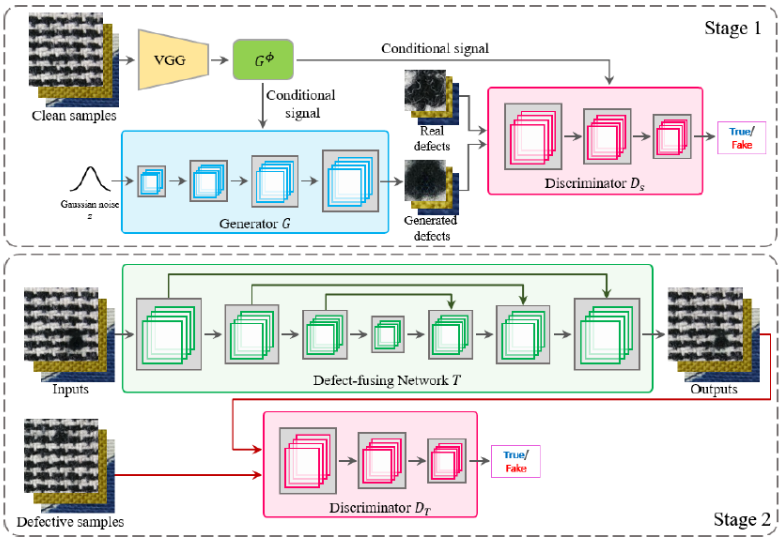

Fig. 4. Schematic of a multistage GAN. Stage 1: Generating defects based on the given texture; Stage 2: Fusing the generated defect into the defect-free sample.

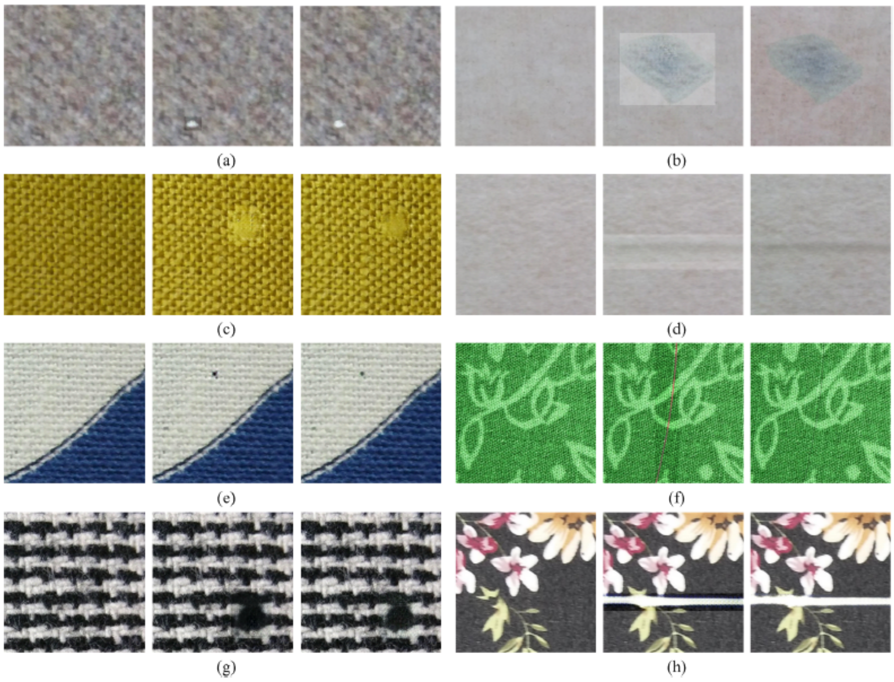

Fig. 6. Samples of synthesized defective fabric samples. Defect types are (a) color spot, (b) oil stain, (c) knot, and (d) broken end in simple-textured fabric and (e) color spot, (f) broken end, (g) broken yarn, and (h) white strip in complex-textured fabric.