



文章编号:1001-9081(2019)07-2103-06

DOI:10.11772/j.issn.1001-9081.2018112247

基于深度神经网络的表面划痕识别方法

李文俊^{1,2}, 陈斌^{2,3*}, 李建明^{1,2}, 钱基德^{1,2}

(1. 中国科学院成都计算机应用研究所, 成都 610041; 2. 中国科学院大学, 北京 100049;
3. 中国科学院广州电子技术有限公司, 广州 510070)
(* 通信作者电子邮箱 chenbin306@sohu.com)

摘要:为实现亮度不均的复杂纹理背景下表面划痕的鲁棒、精确、实时识别,提出一种基于深度神经网络的表面划痕识别方法。用于表面划痕识别的深度神经网络由风格迁移网络和聚焦卷积神经网络(CNN)构成,其中风格迁移网络针对亮度不均的复杂背景下的表面划痕进行预处理,风格迁移网络包括前馈转换网络和损失网络,首先通过损失网络提取亮度均匀模板的风格特征和检测图像的知觉特征,对前馈转换网络进行离线训练,获取网络最优参数值,最终使风格迁移网络生成亮度均匀且风格一致的图像,然后,利用所提出的基于聚焦结构的聚焦卷积神经网络对生成图像中的划痕特征进行提取并识别。以光照变化的金属表面为例,进行划痕识别实验,实验结果表明:与需要人工设计特征的传统图像处理方法及传统深度卷积神经网络相比,划痕漏报率低至8.54%,并且收敛速度更快,收敛曲线更加平滑,在不同的深度模型下均可取得较好的检测效果,准确率提升2%左右。风格迁移网络能够保留完整划痕特征的同时有效解决亮度不均的问题,从而提高划痕识别精度;同时聚焦卷积神经网络能够实现对划痕的鲁棒、精确、实时识别,大幅度降低划痕漏报率和误报率。

关键词:亮度不均;复杂纹理背景;表面划痕识别;风格迁移网络;卷积神经网络

中图分类号: TP391.413; TP18 **文献标志码:**A

Surface scratch recognition method based on deep neural network

LI Wenjun^{1,2}, CHEN Bin^{2,3*}, LI Jianming^{1,2}, QIAN Jide^{1,2}

(1. Chengdu Institute of Computer Application, Chinese Academy of Sciences, Chengdu Sichuan 610041, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. Guangzhou Electronic Technology Company Limited of Chinese Academy of Sciences, Guangzhou Guangdong 510070, China)

Abstract: In order to achieve robust, accurate and real-time recognition of surface scratches under complex texture background with uneven brightness, a surface scratch recognition method based on deep neural network was proposed. The deep neural network for surface scratch recognition consisted of a style transfer network and a focus Convolutional Neural Network (CNN). The style transfer network was used to preprocess surface scratches under complex background with uneven brightness. The style transfer networks included a feedforward conversion network and a loss network. Firstly, the style features of uniform brightness template and the perceptual features of the detected image were extracted through the loss network, and the feedforward conversion network was trained offline to obtain the optimal parameter values of network. Then, the images with uniform brightness and uniform style were generated by style transfer network. Finally, the proposed focus convolutional neural network based on focus structure was used to extract and recognize scratch features in the generated image. Taking metal surface with light change as an example, the scratch recognition experiment was carried out. The experimental results show that compared with traditional image processing methods requiring artificial designed features and traditional deep convolutional neural network, the false negative rate of scratch detection is as low as 8.54% with faster convergence speed and smoother convergence curve, and the better detection results can be obtained under different depth models with accuracy increased of about 2%. The style transfer network can retain complete scratch features with the problem of uneven brightness solved, thus improving the accuracy of scratch recognition, while the focus convolutional neural network can achieve robust, accurate and real-time recognition of scratches, which greatly reduces false negative rate and false positive rate of scratches.

Key words: uneven brightness; complex texture background; surface scratch recognition; style transfer network; Convolutional Neural Network (CNN)

0 引言

工业产品从配料到最终成型的整个过程中,由于运输、生

产工艺等意外情况,表面常存在磕伤、划伤、擦伤等造成的损伤性划痕缺陷,因此,各种表面的划痕识别方法一直都是学者和工业界研究的热点,如表面凹坑缺陷检测^[1]、玻璃屏幕表

收稿日期:2018-11-09;修回日期:2019-01-25;录用日期:2019-01-28。

基金项目:广东省重大科技专项(2017B030306017);广东省产学研合作项目(2017B090901040)。

作者简介:李文俊(1994—),男,湖南岳阳人,硕士研究生,CCF 会员,主要研究方向:计算机视觉、深度学习; 陈斌(1970—),男,四川广汉人,研究员,博士生导师,博士,主要研究方向:实时工业图像分析、字符识别、智能视觉控制; 李建明(1989—),男,四川南部人,博士研究生,主要研究方向:目标识别、深度学习; 钱基德(1988—),男,四川大英人,博士,主要研究方向:机器视觉、深度学习。



面划痕检测^[2]、塑件划痕检测^[3]、机械零件表面划痕检测^[4]等;但是表面划痕识别检测技术受环境、光照、生产工艺和噪声等多重因素的影响,检测系统的信噪比一般较低,微弱信号难以检出或不能与噪声有效区分,以及由于检测对象多样、表面划痕缺陷种类繁多、形态多样、背景复杂,致使对划痕缺陷的描述不充分,对缺陷的特征提取有效性不高^[5]。2012年,胡文瑾等^[6]针对复杂的藏族唐卡图像上的垂直划痕利用小波变换实现划痕检测,首先通过计算小波模极大值来描述图像中目标的多尺度边界,然后通过投影变换增强划痕中心亮度的极值特性并采用多尺度检测划痕的可能位置,再根据划痕的宽度、高度以及连通分量筛选出垂直划痕;但是该方法无法适应各种形态、方向下的划痕。2015年,周鹏等^[7]针对复杂金属表面提出了将剪切波和小波特征融合的划痕识别方法,剪切波变换可以提取图像在不同尺度、不同方向上的信息,因此对方向性的缺陷识别效果好。2017年,马云鹏等^[8]首先从多个方向对金属表面的灰度图像进行转换,然后对多幅图像分别进行灰度波动分析,自适应地改变阈值与步长对图像进行邻域灰度差分割处理,最后利用主成分分析(Principal Component Analysis, PCA)算法将多幅图像压缩至单幅图像实现划痕分割,但是灰度变化不明显的金属表面图像容易出现细节信息丢失等问题以及分割需要时间太长,无法满足工业实时检测要求。郭皓然等^[9]针对高光噪声缺陷图像,利用高斯函数的一阶导数构建Canny最优边缘检测器,结合全局阈值的最大类间方差法和形态学图像分割法,对亮度不均的表面划痕图像进行分割从而实现识别;但该算法通用性不强,并且无法保留划痕的完整特征。因此,如何构建准确、鲁棒、实时的表面划痕检测系统,以适应光照变化、背景噪声以及其他外界不良环境的干扰,仍然是目前需要解决的问题。

2016年,Li等^[10]针对图像转换中的图像风格迁移问题,提出了由前馈转换网络和损失网络构成的风格迁移网络,使用基于损失网络得到的特征重建损失以及风格重建损失来训练前馈转换网络,生成与模板图像风格一致,同时能够保留原始图像主体特征的转换图像。

为了克服上述利用人工设计特征的划痕识别方法的不足,本文提出基于风格迁移网络将亮度均匀、背景单一的模板图像风格迁移到待检图像中进行预处理得到输出图像,再利用聚焦卷积神经网络对输出图像进行表面划痕识别。待检图像通过前馈转换网络生成图像,利用损失网络对图像特征进行特征提取,计算生成图像与待检图像以及生成图像与模板图像语义特征之间的均方差损失,然后传回前馈转换网络进行反向传播^[11],迭代更新权重参数。风格迁移网络能够融合亮度均匀模板图像风格特征以及待检检测图像的知觉特征生成亮度均匀保留完整划痕特征的图像,然后利用基于聚焦结构的多尺度卷积神经网络对生成图像中划痕特征进行识别,由于训练集的亮度均匀且风格特征一致,因此网络能更好地针对不同形态、尺寸的划痕特征进行提取进而识别。

1 风格迁移网络

本文基于Li等^[10]提出的风格迁移网络对亮度不均的复杂背景纹理进行预处理。如图1所示,风格迁移网络包括前馈转换网络 fw 以及损失网络 ϕ 。

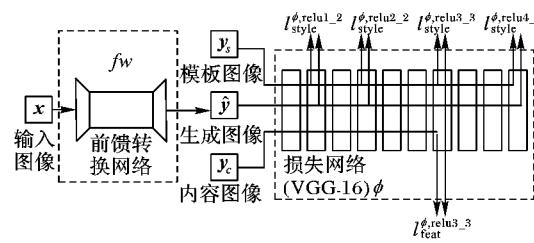


图1 风格迁移网络结构

Fig. 1 Structure of style transfer network

前馈转换网络为一个深度残差网络^[12],参数是权重 w ,网络将输入图像 x 通过映射 $\hat{y} = fw(x)$ 转换成输出图像 \hat{y} ,每一个损失函数计算一个标量值 $l_i(\hat{y}, y_i)$ 用来衡量输出图像 \hat{y} 以及一个目标图像之间的差异。前馈转换网络使用随机梯度下降(Stochastic Gradient Descent, SGD)^[13]优化方法来最小化一组加权为 λ_i 的损失函数 W^* ,可以写作:

$$W^* = \arg \min_w E_{x, |y_i|} \left[\sum_{i=1} \lambda_i l_i(fw(x), y_i) \right] \quad (1)$$

相比逐像素计算损失函数的方式,损失网络 ϕ 定义了一个特征重建损失 l_{feat}^{ϕ} 以及一个风格重建损失 l_{style}^{ϕ} ,分别用来衡量输出图像 \hat{y} 与内容图像 y_c 以及模板图像 y_s 之间的差异。在训练时,对于每一张输入图像 x ,都对应有一个内容图像 y_c 以及一张模板图像 y_s 。对于风格迁移而言,内容图像 y_c 就是输入图像 x ,而输出图像 \hat{y} 应该是内容图像 y_c 与模板图像 y_s 的结合。对于一张模板图像,可以训练得到一个风格迁移网络。在测试时,输入图像 x 直接通过前馈转换网络得到输出图像 \hat{y} 。

1.1 前馈转换网络

前馈转换网络为一个深度残差网络,网络设计基本遵循Radford等^[14]提出的指导性意见,没有使用任何的池化层,而是利用步幅卷积在网络中进行上采样或者下采样操作。前馈转换网络由五个残差块组成,使用ResNet网络结构^[15],所有的非残差卷积层之后都使用了一个空间性的批标准化(Batch Normalization, BN)^[16]以及非线性激活函数修正单元(Rectified Linear Unit, ReLU)^[17],最后一层除外,最后一层使用一个缩放的tanh非线性函数来确保输出图像的像素值在[0, 255]区间。除了第一层和最后一层使用 9×9 大小的卷积核之外,其余卷积层都使用 3×3 大小的卷积核。

输入和输出:对于风格迁移而言,输入和输出都是灰度图,大小为 200×200 。网络使用了两个步长为2的卷积对输入进行下采样,紧接着是几个残差块,然后是两个步长为 $1/2$ 的卷积层来进行上采样。

1.2 损失网络

损失网络采用的网络模型为利用ImageNet数据集预训练的VGG16^[18],并且定义了两个损失函数用来衡量两张图像在高层语义以及感知上的差别。

1) 特征重建损失 l_{feat}^{ϕ} 。没有用到传统的逐像素对比损失,而是用VGG16网络模型来得到高层特征来表征图像内容,公式可以写作:

$$l_{\text{feat}}^{j,f}(\hat{y}, y_c) = \frac{1}{C_j H_j W_j} \| \phi_j(\hat{y}) - \phi_j(y_c) \|_2^2 \quad (2)$$

其中, C, H, W 分别为特征图的通道数及尺寸大小,下述相同。式(2)寻找一个输出图像 \hat{y} 与内容图像 y_c 对应的第 j 层的特征图 $\phi_j(\hat{y})$ 和 $\phi_j(y_c)$ 使得低维语义层之间特征损失最小,使得输出图像能够产生与内容图像 y_c 视觉上不太能区分的图



像,如果用高维语义层来进行特征重建,内容和全局结构会被保留下来,但是颜色纹理以及精确的形状信息将丢失,因此使用一个特征重建损失 l_{feat}^{ϕ} 来训练前馈转换网络让输出图像 \hat{y} 非常接近内容图像 y_c ,但并不是让它们完全匹配。

2) 风格重建损失 l_{style}^{ϕ} 。特征重建损失 l_{feat}^{ϕ} 使输出图像 \hat{y} 的内容更加接近内容图像 y_c ,同时,也希望通过风格重建损失来让输出图像 \hat{y} 与模板图像 y_s 在颜色、纹理、通用模式等方面更加接近。为了实现这一点,定义了 Gram 矩阵来表征上述方面特性,令 $\phi_j(x)$ 表示对于输入图像 x 在损失网络 ϕ 中第 j 层的激活响应,特征图的形状为 $C_j \times H_j \times W_j$,大小为 $C_j \times C_j$ 的 Gram 矩阵 $G_j^{\phi}(x)$ 中的元素定义可以写作:

$$G_j^{\phi}(x)_{c,\hat{c}} = \frac{1}{C_j \times H_j \times W_j} \sum_{h=1}^{H_j} \sum_{w=1}^{W_j} \phi_j(x)_{h,w,c} \phi_j(x)_{h,w,\hat{c}} \quad (3)$$

其中 c, \hat{c} 分别代表特征图中任意两个通道,如果将对应特征图 $\phi_j(x)$ 看成一个 C_j 维度的特征,每个特征的形状为 $H_j \times W_j$,式(3)中 $G_j^{\phi}(x)$ 与 C_j 维度的特征的非中心的协方差成正比,每一个网格位置都可以看作一个独立的样本,从而获取能够影响其他信息的特征,Gram 矩阵能够通过调整 $\phi_j(x)$ 的维度为 $C_j \times H_j \times W_j$ 的矩阵 ψ ,那么 $G_j^{\phi}(x) = \psi\psi^T / C_j H_j W_j$ 。

3) 风格重建损失 l_{style}^{ϕ} 。输出图像 \hat{y} 与模板图像 y_s 的 Gram 矩阵的弗罗贝尼乌斯范数(Frobenius norm)的差的平方,可以写作:

$$L_{style}^{\phi,j}(\hat{y}, y_s) = \| G_j^{\phi}(\hat{y}) - G_j^{\phi}(y_s) \|^2 \quad (4)$$

即使 \hat{y} 与 y_s 的尺寸不同,风格重建损失 l_{style}^{ϕ} 仍然能够被很好地定义,因为它们的 Gram 矩阵的形状都是相同的。优化风格重建损失 l_{style}^{ϕ} 能够保留模板图像 y_s 的风格化特征,但是没有保留它的空间结构信息,为了在一系列层 J 中而不是单层 j 中执行风格重建,定义 $l_{style}^{\phi,J}(\hat{y}, y_s)$ 为每一层 $j \in J$ 的损失之和。

2 聚焦卷积神经网络

由于划痕种类、长度、宽度、形状等的多样性使得传统划痕检测方法适应性比较差,因此利用卷积神经网络来进行识别,因为卷积神经网络是一种学习的方式,在训练样本数量足够大、质量足够高、种类足够多的情况下,能够设计足够多层次的带有大量参数的卷积神经网络来自动学习到具有代表性的划痕的本质特征,凭借网络中大量激活函数的使用,也提升了泛化能力,使得算法对于划痕识别的适应性大幅度增强。本文在 8 层卷积核大小都为 3×3 的卷积层以及 3 层全连接层的 VGG11^[19] 网络模型上进行改进,提出了基于聚焦结构的聚焦卷积神经网络来进一步解决如图 2 工业环境下产生的短长度碎屑对于划痕识别的误报问题。

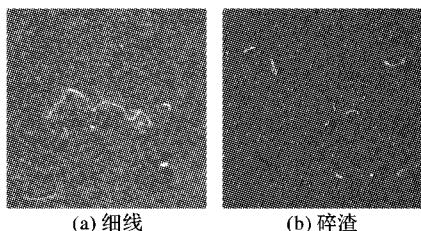


图 2 短长度碎屑
Fig. 2 Debris of short length

2.1 聚焦结构

图像聚焦表现为焦点区域分辨率增大变得更加清晰,非焦点区域分辨率减小变得模糊。因为聚焦卷积神经网络能更好地提取识别形态、大小、对比度各异的划痕特征,基于聚焦结构 VGG11 网络结构进行改进提出了聚焦卷积神经网络。

聚焦结构中的卷积核使用如图 3 所示的空洞卷积^[20], d 代表扩张率, F_d 表示利用卷积核内部实值的间距,空洞卷积的感受野范围与扩张率 d 的关系可以写作:

$$F_d = (2^{d+1} - 1) \times (2^{d+1} - 1) \quad (5)$$

如图 3(a)为标准卷积核,其扩张率为 1,感受野范围为 3×3 ,图 3(c)中扩张率为 3,其感受野范围为 15×15 。

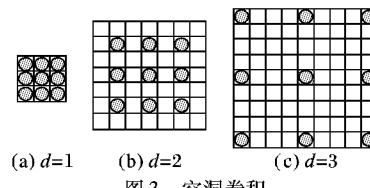


图 3 空洞卷积

Fig. 3 Dilated convolution

聚焦结构由如图 4 所示的聚焦块和对焦块组成。聚焦结构的主要设计思想在于利用不同扩张率的空洞稀疏卷积结构来聚焦到一个多尺度高度相关的单元中来近似表达一个优化的局部稀疏结构,好处在于计算复杂度未达到计算爆炸时,能够增加网络的宽度和深度,增加网络的表达能力,即类似于在图像尺寸大小不变的情况下,图像中非焦点区域的分辨率减小转移至焦点区域,使焦点区域内分辨率增加进而提高目标特征显著度及辨识度。

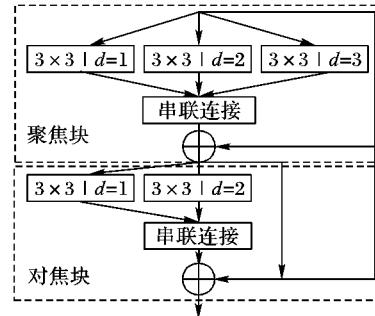


图 4 聚焦结构

Fig. 4 Focus structure

聚焦块由扩张率分别为 1、2、3 的三种空洞卷积对上层相关统计信息卷积得到多尺度聚焦信息,以便于下一个阶段能够同时提取不同尺寸下的特征,然后在通过对焦块,进一步聚焦上一阶段提取到的多尺度聚焦处理结果,使用扩张率分别为 1、2 的两种空洞卷积来最终表达一个稠密且压缩的信息形式。为了提升聚焦结构内部以及多层之间的梯度传播能力,在聚焦块、对焦块以及整个聚焦结构中都加入了与 ResNet-18 中典型残差结构相同的快捷连接方式。在每一个卷积结束之后,都使用了标准化以及非线性激活函数 ReLU,标准化能够去除特征之间的相关性,使所有特征具有相同的均值和方差,能有效提高反向传播的效率,还具有参数正则化的效果;同时结合非线性激活函数 ReLU,能够大幅度提升网络的泛化性能。聚焦卷积神经网络中使用的聚焦结构中的标准化为批标准化(BN)。



2.2 网络结构

网络结构主要基于 VGG11 模型进行构建,这是一个对于工业环境中产生的短长度碎屑也能很好地识别,与划痕特征区别开的满足工业条件实时处理的网络框架。

网络输入为 200×200 的灰度图像,对 VGG11 网络结构,保留第一层网络结构,倒数第二层全连接层替换为全局均值池化层 (Global Average Pooling)^[19],剩下层都替换为聚焦结构。网络参数如表 1 所示,其中 Focus1 等为聚焦结构,包含聚焦块 (a) 及对焦块 (b),Global_avg 表示全局均值池化层,在每一层卷积层之后都使用批标准化 (BN) 以及 ReLU 非线性激活函数来提升网络的泛化能力以及表达能力,网络优化方式使用自适应矩估计 (Adaptive moment estimation, Adam)^[20] 梯度下降优化算法,设置初始学习率为 0.001,权重衰减为 0.0005,一阶矩估计的指数衰减率为 0.9,二阶矩估计的指数衰减率为 0.999。

表 1 聚焦卷积神经网络参数

Tab. 1 Parameters of focus convolutional neural network

卷积层	核尺寸	不同扩张率卷积核特征维度		
		$d=1$	$d=2$	$d=3$
Conv1	3	96	/	/
Focus1(a)	3	64	16	16
Focus1(b)	3	96	64	/
Focus2(a)	3	96	32	32
Focus2(b)	3	160	96	/
Focus3(a)	3	128	64	64
Focus3(b)	3	384	128	/
Focus4(a)	3	256	128	128
Focus4(b)	3	512	256	/
Global_avg	13	768	/	/
Fc (Full connection)	1	2	/	/

3 实验结果及分析

3.1 实验设置

1) 数据集。使用来自 43 套不同型号、不同尺寸、不同涂层工艺的服务器金属表面提取的划痕及背景图像作为数据集,将 200×200 大小的表面划痕图像作为正样本,同样大小的背景图像作为负样本。训练集包含 11 套不同服务器的 5246 张金属表面划痕图像和 5502 张背景图像,验证集包含不同于训练集的 11 套不同服务器的 5134 张划痕图像和 5133 张背景图像,测试集包含剩下 21 套服务器表面随机采样的 10312 张划痕图像和 10298 张背景图像。对训练集用随机旋转、翻转和裁剪来增加划痕图像数据并随机采样背景图像数据,经过数据增强后,训练集包含 35006 张划痕图像和 35020 张背景图像。

2) 评估和实施。对于评价性能指标,采用准确率 (Accuracy, Acc)、单幅图像耗时 (Time)、模型参数 (Size)、灵敏度 (Sensitivity, Sen)、特异性 (Specificity, Spec),其中划痕漏报率 (Missing Report Rate, MRR) 为灵敏度,背景误报率 (False Alarm Rate, FAR) 为特异性。

对于不同卷积神经网络模型,未经过风格迁移网络进行预处理,在测试集中得到的分类结果分别表示为 VGG11^[18]、MobileNetV1^[21]、ShuffleNet^[22]、ResNet18^[15] 和 FocusNet,而经

过风格迁移网络进行预处理,在测试集中得到的分类结果分别表示为 VGG11-Pre、MobileNetV1-Pre、ShuffleNet-Pre、ResNet18-Pre、FocusNet-Pre。

所有神经网络训练的实验参数都设置为利用 Adam 梯度下降优化算法,初始学习率为 0.001,权重衰减为 0.0005,一阶矩估计的指数衰减率为 0.9,二阶矩估计的指数衰减率为 0.999,批数量为 50 张,最大迭代次数为 40000。

本实验主要是在开源框架 PyTorch 下实现,所有实验都在 CPU Intel Core i7 8700K、GPU NVIDIA GeForce GTX 1080Ti 和 16 GB RAM 的计算机上进行。

3.2 本文方法与人工设计特征方法的对比实验

表 2 列出了上述传统人工设计特征的方法^[6-9],基于小波分析 (Wavelet Analysis, WA) 的方法、基于剪切波与小波 (Shear Wave and Wavelet, SWW) 的方法、自适应分割 (Adaptive Segmentation, AS) 及全局阈值自适应 (Global Threshold Adaptive, GTA) 的方法,与本文提出的方法 (FocusNet-Pre) 在测试集上进行比较。

表 2 与传统人工设计特征方法的对比结果

Tab. 2 Comparison with traditional artificial design feature methods

方法	FAR/%	MRR/%	Acc/%	运行时间/ms
WA	4.67	20.78	68.71	24.30
SWW	28.12	20.13	75.74	62.50
AS	22.08	15.26	82.79	34.70
GTA	9.88	14.71	89.63	695.32
FocusNet-Pre	0.55	8.54	96.66	1.90

可以看出本文方法比传统人工设计特征方法在各类评价指标上的表现都要更好,具体来说,本文方法相比人工设计特征方法中的最好结果,误报率降低了 9.33 个百分点,漏报率降低了 6.17 个百分点,准确率提升了 7.03 个百分点,并且运行时间大幅度降低,单张图像效率高了 12 倍左右,仅需要 1.90 ms。这表明深度卷积特征比低级的人工特征描述符具有更加强大的表达能力、泛化性能,各方面大幅度超越传统人工设计特征方法,能够很好地提取并识别复杂背景下亮度不均的形态、大小、种类各异的表面划痕。

3.3 风格迁移网络效果论证实验

为了验证风格迁移网络能够有效解决划痕识别中遇到的亮度不均、复杂背景的情况,因此在不同深度神经网络模型下进行是否利用风格迁移网络对待检输入图像进行预处理的实验,来观察风格迁移网络的作用。图 5 中展示风格迁移网络对于输入图像利用单一背景且亮度均匀的模板图像进行预处理的效果。

表 3 列出了本文提出聚焦卷积神经网络 (FocusNet) 及其他神经网络模型在是否利用风格迁移网络进行预处理,并在测试集得到的分类对比结果。可以看出,在不同的网络架构中,没有经过风格迁移网络进行预处理的分类表现比经过风格迁移网络进行预处理的分类结果表现都要差。具体来说,在 VGG11 前经过风格迁移网络预处理后,漏报率降低了 5.46 个百分点,准确率提高了 2.24 个百分点;在 MobileNetV1 中,漏报率降低了 4.83 个百分点,准确率提高了 2.21 个百分点;在 ShuffleNet 中,漏报率降低了 5.80 个百分点,准确率提高了 2.20 个百分点;在 ResNet18 中,漏报率降低了 5.20 个百分



点,准确率提高了 2.10 个百分点;在本文方法 FocusNet 中,漏报率降低了 5.03 个百分点,准确率提高了 2.45 个百分点。

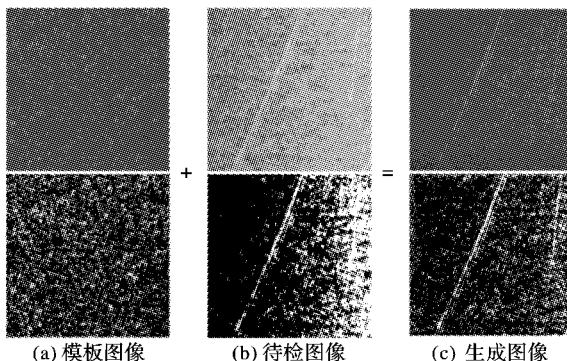


图 5 风格迁移网络处理效果
Fig. 5 Style transfer network effect

表 3 不同方法的实验结果对比

Tab. 3 Experimental results comparison of different methods

方法	FAR/%	MRR/%	Acc/%
VGG11	0.71	16.67	92.34
MobileNetV1	0.62	15.46	93.02
ShuffleNet	0.53	15.02	93.35
ResNet18	0.51	13.89	93.93
FocusNet	0.46	13.57	94.21
VGG11-Pre	0.69	11.21	94.58
MobileNetV1-Pre	0.61	10.63	95.23
ShuffleNet-Pre	0.51	9.22	95.55
ResNet18-Pre	0.49	8.69	96.03
FocusNet-Pre	0.44	8.54	96.66

图 6 表示不同卷积神经网络模型在训练过程中的损失对比,图中 VGG11、MobileNetV1、ShuffleNet、ResNet18、FocusNet 代表没有经过风格迁移网络预处理的表现结果,VGG11-Pre、MobileNetV1-Pre、ShuffleNet-Pre、ResNet18-Pre、FocusNet-Pre 表示经过风格迁移网络进行预处理的表现效果。可以看出各类卷积神经网络的表现差异不大,但是相比没有经过风格迁移网络进行预处理的训练收敛曲线,经过风格迁移网络预处理后的训练收敛速度更快,收敛曲线更加平滑。由此说明,综上所示,不同的神经网络模型,经过风格迁移网络预处理后能够有效增加分类结果的准确性,并且减小过拟合程度,提升网络模型的泛化性能。

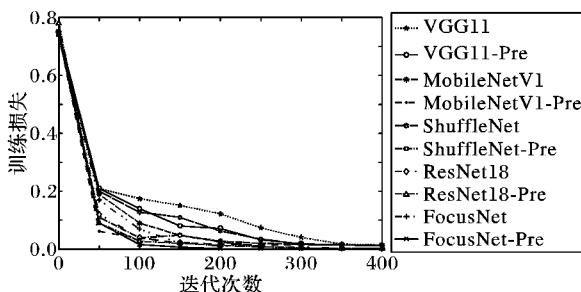


图 6 不同网络模型的训练损失
Fig. 6 Training loss of different network models

3.4 本文方法与其他卷积神经网络方法的对比实验

为了验证本文提出的聚焦卷积神经网络(FocusNet)能够更加有效地提取待检图像中划痕特征以及相比其他神经网络

模型对于表面划痕检测更加有效,将不同的神经网络模型在相同训练集下经过相同风格迁移网络预处理后进行训练,并将测试集得到的各种评价分类结果中的指标进行对比。从表 4 实验结果可以看到,FocusNet-Pre 提取到的深层特征优于 VGG11-Pre、MobileNetV1-Pre 和 ShuffleNet-Pre 中提取的深层特征,与 ResNet18-Pre 差异不大,但模型参数减少了 67% 左右,单幅图像耗时减少了 63.46%,漏报率仅为 8.54%,误报率为 0.44%,同时准确率高达 96.66%,说明聚焦卷积神经网络不仅能够识别不同种类、长度、宽度、形状等的划痕,并且具有较好的鲁棒性和泛化性能;同时误报率也有所降低并且模型运行时间短,单张图像耗时仅仅需要 1.90 ms,完全满足工业生产线上的实时性的要求。

表 4 不同卷积神经网络的对比结果
Tab. 4 Comparison results of different convolutional neural networks

网络	FAR/%	MRR/%	Acc/%	耗时/ms	Size/MB
VGG11-Pre	0.69	11.21	94.58	1.90	74.10
MobileNetV1-Pre	0.61	10.63	95.23	2.40	25.80
ShuffleNet-Pre	0.51	9.22	95.55	4.00	7.20
ResNet18-Pre	0.49	8.69	96.03	5.20	89.40
FocusNet-Pre	0.44	8.54	96.66	1.90	29.50

4 结语

1)提出基于深度神经网络的表面划痕识别方法,该方法一方面将亮度不均、复杂背景的待检图像通过风格迁移网络进行亮度均匀化及背景单一化,同时保留完整的划痕特征;另一方面利用本文提出的聚焦卷积神经网络实现表面划痕深度特征提取并识别,能够鲁棒地识别形态、大小、种类各异的表面划痕。

2)通过实验证明了所提算法在解决亮度不均问题同时生成保留完整的划痕特征图像和有效地识别不同形态、大小的表面划痕方面具有更好的效果。另外,实验数据表明,基于所提算法在不同深度卷积神经网络模型对复杂背景下的划痕均有不错的识别效果,进一步证明了该算法的有效性。

参考文献 (References)

- [1] 宋迪,张东波,刘霞. 基于 Gabor 和纹理抑制的手机配件划痕检测[J]. 计算机工程,2014,40(9):1-5.(SONG D, ZHANG D B, LIU X. Scratch detection for mobile phone accessories based on Gabor and texture suppression[J]. Computer Engineering, 2014, 40 (9):1-5.)
- [2] 韩芳芳,段发阶,张宝峰,等. 单线阵 CCD 系统的表面凹坑缺陷检测方法[J]. 哈尔滨工业大学学报,2012,44(4):116-120.(HAN F F, DUAN F J, ZHANG B F, et al. Study and modeling for surface pit defect detection based on linear array CCD system[J]. Journal of Harbin Institute of Technology, 2012, 44(4):116-120.)
- [3] 崔炽标,李阳,毛霆,等. 基于预处理与卷积神经网络的塑件划痕检测[J]. 模具工业,2017,43(9):1-6.(CUI Z B, LI Y, MAO T, et al. Scratch detection of plastics based on preprocessing and convolutional neural network[J]. Die and Mould Industry, 2017, 43(9): 1-6.)
- [4] 李克斌,余厚云,周申江. 基于形态学特征的机械零件表面划痕检测[J]. 光学学报,2018,38(8):815027-1-815027-7.(LI K B,



- YU H Y, ZHOU S J. Scratch detection for the surface of mechanical parts based on morphological features [J]. *Acta Optica Sinica*, 2018, 38(8): 815027-1 – 815027-7.)
- [5] 汤勃, 孔建益, 伍世虔. 机器视觉表面缺陷检测综述[J]. 中国图象图形学报, 2017, 22(12): 1640 – 1663. (TANG B, KONG J Y, WU S Q. Review of surface defect detection based on machine vision [J]. *Journal of Image and Graphics*, 2017, 22(12): 1640 – 1663.)
- [6] 胡文瑾, 李战明, 刘仲民. 一种基于小波分析的唐卡图像划痕检测[J]. 光学技术, 2012, 38(6): 751 – 755. (HU W J, LI Z M, LIU Z M. Scratch detection algorithm based on wavelet analysis for Thangka image [J]. *Optical Technique*, 2012, 38(6): 751 – 755.)
- [7] 周鹏, 徐科, 刘顺华. 基于剪切波和小波特征融合的金属表面缺陷识别方法[J]. 机械工程学报, 2015, 51(6): 98 – 103. (ZHOUPENG, XU K, LIU SHUNHUA. Surface defect recognition for metals based on feature fusion of shearlets and wavelets [J]. *Journal of Mechanical Engineering*, 2015, 51(6): 98 – 103.)
- [8] 马云鹏, 李庆武, 何飞佳, 等. 金属表面缺陷自适应分割算法[J]. 仪器仪表学报, 2017, 38(1): 245 – 251. (MA Y P, LI Q W, HE F J, et al. Adaptive segmentation algorithm for metal surface defects [J]. *Chinese Journal of Scientific Instrument*, 2017, 38(1): 245 – 251.)
- [9] 郭皓然, 邵伟, 周阿维, 等. 全局阈值自适应的高亮金属表面缺陷识别新方法[J]. 仪器仪表学报, 2017, 38(11): 2797 – 2804. (GUO H R, SHAO W, ZHOU A W, et al. Novel defect recognition method based on adaptive global threshold for highlight metal surface [J]. *Chinese Journal of Scientific Instrument*, 2017, 38(11): 2797 – 2804.)
- [10] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution [C]// Proceedings of the 2016 European Conference on Computer Vision. Berlin: Springer, 2016: 694 – 711.
- [11] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating error [J]. *Nature*, 1986, 323(6088): 533 – 536.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2016: 770 – 778.
- [13] BOTTOU, CURTIS F E, NOCEDA J, et al. Optimization methods for large-scale machine learning [J]. *SIAM Review*, 2016, 60(2): 223 – 311.
- [14] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [J]. *ArXiv Preprint*, 2016, 2016: 1511.06434.
- [15] LIM B, LEE K M. Deep recurrent ResNet for video super-resolution [C]// Proceedings of the 2018 IEEE Conference on Asia-Pacific Signal and Information Processing Association Summit. Washington, DC: IEEE Computer Society, 2018: 643 – 648.
- [16] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2015: 448 – 456.
- [17] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines [C]// Proceedings of the 2010 International Conference on Machine Learning. New York: ACM, 2010: 807 – 814.
- [18] SIMONVAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]// Proceedings of the 2015 International Conference on Learning. Washington, DC: IEEE Computer Society, 2015: 687 – 699.
- [19] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [C]// Proceedings of the 2016 International Conference on Learning. Washington, DC: IEEE Computer Society, 2016: 511 – 524.
- [20] KINGMA D, BA J. Adam: a method for stochastic optimization [C]// Proceedings of the 2014 International Conference on Learning. Washington, DC: IEEE Computer Society, 2014: 248 – 263.
- [21] HOWARD A G, ZHOU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2017: 1056 – 1065.
- [22] ZHANG X, ZHOU X, LIN M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2017: 563 – 572.

This work is partially supported by the Major Scientific and Technological Project of Guangdong Province (2017B030306017), the Industry-University-Research Cooperation Project of Guangdong Province (2017B090901040).

LI Wenjun, born in 1994, M. S. candidate. His research interests include computer vision, deep learning.

CHEN Bin, born in 1970, Ph. D., research fellow. His research interests include real time industrial image analysis, character recognition, intelligent visual control.

LI Jianming, born in 1989, Ph. D. candidate. His research interests include object recognition, deep learning.

QIAN Jide, born in 1988, Ph. D. His research interests include machine vision, deep learning.