

# 基于深度学习的管道纱线及其颜色检测<sup>①</sup>



李进飞, 李建强, 段玉堂, 任国栋, 史伟民

(浙江理工大学 机械与自动控制学院, 杭州 310018)

通讯作者: 李进飞, E-mail: [201930507022@mails.zstu.edu.cn](mailto:201930507022@mails.zstu.edu.cn)

**摘要:** 为了保证自动换筒系统中的纱线自动打结机能够正常运行, 需要对管道吸取的纱线进行检测. 纱线纤细、种类繁多且颜色各异, 传感器方法难以胜任, 使用图像处理的方式较为合适. 但是对于纱线检测问题传统的图像处理方法复杂且检测准确率低, 难以解决纱线种类多、尺寸不一以及颜色多等问题, 故本文提出了一种基于 Inception v4 中 Inception-ResNet-A 块进行改进的多尺度深度可分离卷积块组成的网络来检测管道中的纱线. 其中改进的多尺度深度可分离卷积块采用  $3 \times 3$  卷积核的深度可分离卷积层代替 Inception-ResNet-A 块中  $3 \times 3$  传统卷积层并去除了其中的一些  $1 \times 1$  卷积层, 简化卷积块的计算量以及参数量, 此外还结合了残差网络 ResNet 的方法进行通道融合, 防止特征丢失. 试验结果表明, 该网络模型具有非常好的泛化能力以及辨识效果.

**关键词:** 管道纱线检测; 深度学习; 卷积神经网络; 图像分类; 分类模型; 数据增强

引用格式: 李进飞, 李建强, 段玉堂, 任国栋, 史伟民. 基于深度学习的管道纱线及其颜色检测. 计算机系统应用, 2021, 30(6):311-315. <http://www.c-s-a.org.cn/1003-3254/7967.html>

## Pipe Yarn and Color Detection Based on Deep Learning

LI Jin-Fei, LI Jian-Qiang, DUAN Yu-Tang, REN Guo-Dong, SHI Wei-Min

(Faculty of Mechanical Engineering & Automation, Zhejiang Sci-Tech University, Hangzhou 310018, China)

**Abstract:** To ensure the normal operation of the automatic knotting machine of yarn in the automatic bobbin changing system, we need to detect the yarn sucked by the pipe. Yarn is detected by image processing instead of sensors because it is thin with diverse types and colors. However, traditional image processing methods are too complex and inaccurate to identify yarn with various types, sizes, and colors. This study proposes a network of multi-scale depth separable convolution blocks modified based on Inception-Resnet-A block of Inception v4 to detect yarn in pipes. The conventional  $3 \times 3$  convolution layers in the Inception-ResNet-A block is replaced with the depth separable convolution layers of the  $3 \times 3$  convolution kernel, and some of the  $1 \times 1$  convolution layers are removed for less parameters of convolution blocks and simpler calculation. In addition, ResNet is employed for channel fusion to prevent feature loss. According to the experimental results, this network model is remarkable in generalization and recognition.

**Key words:** pipe yarn detection; deep learning; Convolutional Neural Network (CNN); image classification; classification model; data enhancement

随着工业 4.0 的快速推进, 机器代替工人的趋势不可逆转, 因此为了提高纺纱<sup>[1]</sup>效率、降低人工成本, 采用机器代替人工进行纱线打结大势所趋. 自动换筒系

统中的纱线自动打结机通过管道吸取纱线, 机械机构利用管道中的纱线进行打结完成人工打结的任务, 大大降低人工成本, 提高了打结可靠性. 但是纱线打结机

① 收稿时间: 2020-10-13; 修改时间: 2020-11-16, 2020-11-24; 采用时间: 2020-12-01; csa 在线出版时间: 2021-06-01

正常运行前提是管道吸取了所需的纱线,因此检测管道是否成功吸取纱线以及该纱线是否是所需颜色的纱线至关重要。

传感器方法如红外开关传感器<sup>[2]</sup>利用物体阻隔发射器与接收器通路的信号接收来判断是否有物体。由于纱线纤细,激光或者红外容易穿透纱线,所以难以对纱线进行检测。相比之下,采用摄像头对管道进行拍摄,利用图像处理的方法进行检测较为容易。传统的图像处理方法可以使用 OpenCV 对图像进行 HSV 颜色分离从而对图片中纱线进行提取,最终确定纱线及其颜色。但是对于一些纱线颜色与背景颜色相差不多以及各种各样的纱线进行检测时,该方法就显得较为复杂,且鲁棒性较差。

近年来,深度学习<sup>[3,4]</sup>的卷积神经网络在图像处理领域发展迅速赢得了广泛的关注。相比传统的图像处理方式,深度学习具有诸多优点,如模型具有可迁移性、无需手动设计特征、检测准确度高特性。从卷积网络的出现到现在,已经出现了非常优秀的网络,例如 LeNet<sup>[5]</sup>、VGG<sup>[6]</sup>、AlexNet<sup>[7]</sup>、GoogLeNet<sup>[8]</sup>等,但是这些网络要么就是检测尺度单一,要么就是计算量以及参数量较大。

本文针对纱线粗细不一、种类繁多、检测困难等问题,提出了一种多尺度的深度分离卷积<sup>[9,10]</sup>网络,结合 ResNet<sup>[11,12]</sup>网络的特点,解决纱线检测特征尺度单一、检测效果低、特征丢失、模型计算量与参数量大等问题,使得检测精度能满足一定的应用要求。

## 1 管道纱线图像采集与处理

### 1.1 图像采集与整理

利用海康 MV-CE050-30UC 摄像头对管道进行拍照采集 1200 张图片,经过去除重复以及有缺陷的图片后制作成数据集,其中包括无纱线、黑色纱线、蓝色纱线、粉色纱线、白色纱线、黄色纱线 6 种情况。该数据集在程序运行时,以 4:1 的比例随机划分为训练集和验证集,并将每张图片裁剪成 224×224 像素。

### 1.2 图像数据增强

由于采集的数据量较少,在训练过程中需要对图像进行数据增强来提高数据集的多样性,从而有效地降低过拟合现象。图像增强的方法很多,本论文对样本进行增强采用的方法主要包括:上下翻转、左右翻转、亮度变化等,具体效果如图 1 所示。

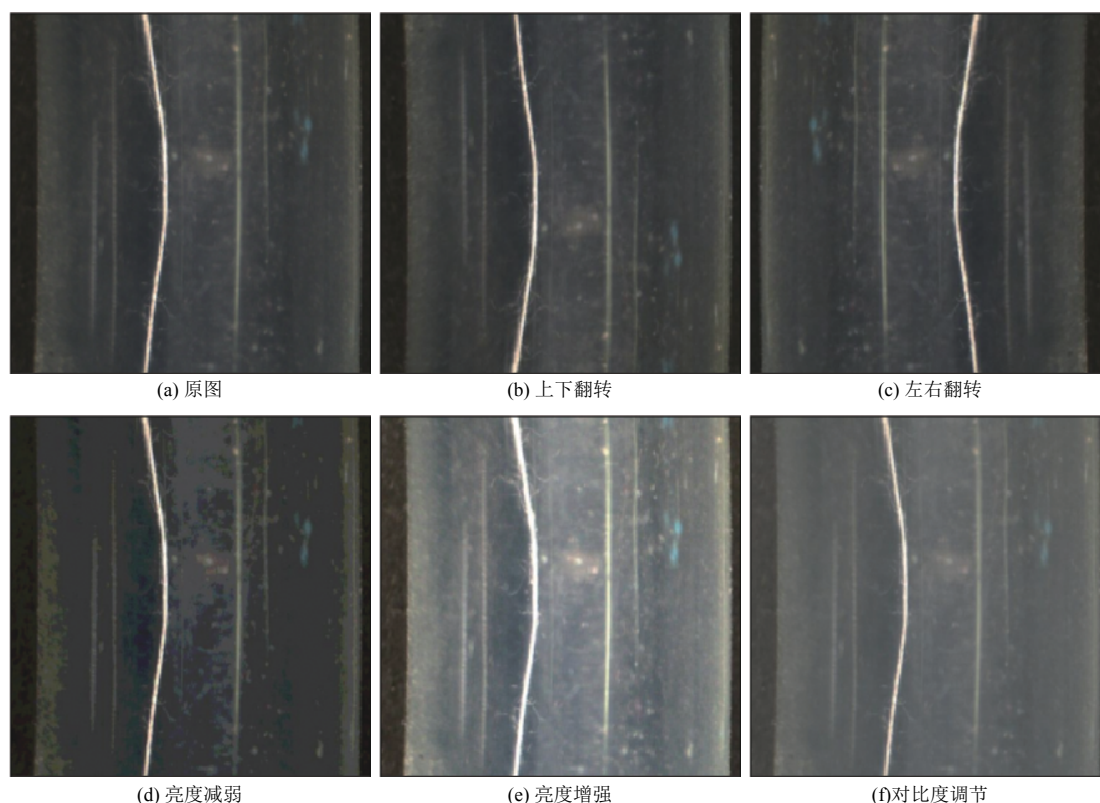


图 1 图像增强效果图

## 2 改进的多尺度深度可分离卷积模型

### 2.1 改进的多尺度深度可分离卷积块

改进的多尺度深度可分离卷积块是基于 Inception v4<sup>[13-15]</sup> 中的 Inception-ResNet-A 块提出并改进的,从图 2 可知,输入特征图  $x$  分别经过  $1 \times 1$  卷积核卷积、一个  $3 \times 3$  的深度可分离卷积核卷积以及两个  $3 \times 3$  的深度可分离卷积核卷积,得到输出特征图后,再经过  $1 \times 1$  卷积核进行多尺度特征融合得到  $F(x)$ ,从而保证输出特征图的通道数.经过最后的  $1 \times 1$  卷积核卷积后,采用 ResNet 的短接线方式,进行输出  $F(x)$  与输入特征图之间的特征融合,解决特征梯度消失问题,其中输入特征与输出特征之间的关系如式 (1) 所示:

$$y = F(x) + x \quad (1)$$

与 Inception-ResNet-A 块相比,在结构上改进的多尺度深度可分离卷积块少了两个  $1 \times 1$  的卷积核,因为 Inception-ResNet-A 块在设计上希望通过  $1 \times 1$  卷积核减少输入通道数量,但是本论文所需的模型不大,所以去除这两个卷积核,从而降低网络的计算量以及参数量.其次,因为深度可分离卷积与传统卷积相比计算量和参数量更少,所以改进的多尺度深度可分离卷积块采用  $3 \times 3$  卷积核的深度可分离卷积代替传统的  $3 \times 3$  卷积核卷积,从而使得该块的计算量和参数量更少,网络更轻量化.其中,深度可分离卷积与传统卷积的比较<sup>[16]</sup>如下所示.

设输入与输出特征图大小相同时,传统卷积的计算量以及参数量如下:

$$D_k \times D_k \times M \times D_f \times D_f \times N \quad (2)$$

$$D_k \times D_k \times M \times N \quad (3)$$

式中,式 (2) 为传统卷积计算量的计算公式,式 (3) 为传统卷积参数量的计算公式.其中卷积核大小为  $D_k \times D_k$ ,输出特征图大小为  $D_f \times D_f$ , $M$  为输入特征图数量, $N$  为输出特征图数量.

与之相比,深度可分离卷积计算量与参数量为逐

通道卷积和逐点卷积的两者之和,如式 (4)、式 (5) 所示:

$$D_k \times D_k \times M \times D_f \times D_f + M \times D_f \times D_f \times N \quad (4)$$

$$D_k \times D_k \times M + M \times N \quad (5)$$

其中,式 (4) 和式 (5) 前相乘部分为逐通道卷积计算量和参数的计算,后面相乘部分为逐点卷积计算量与参数的计算.深度可分离卷积与传统卷积计算量与参数的比值如式 (6) 和式 (7) 所示.

$$\frac{D_k \times D_k \times M \times D_f \times D_f + M \times D_f \times D_f \times N}{D_k \times D_k \times M \times D_f \times D_f \times N} = \frac{1}{N} + \frac{1}{D_k^2} \quad (6)$$

$$\frac{D_k \times D_k \times M + M \times N}{D_k \times D_k \times M \times N} = \frac{1}{N} + \frac{1}{D_k^2} \quad (7)$$

从式 (6) 和式 (7) 可以看出深度可分离卷积与传统卷积的计算量以及参数的差异.当采用  $5 \times 5$  卷积核时,则  $D_k^2$  为  $1/25$ ,则深度可分离卷积与传统卷积相比计算量以及参数的比值接近  $1/25$ .从中可以看出,深度可分离卷积更轻量化.

从上可知,本文的多尺度深度可分离卷积块采用  $3 \times 3$  卷积核的深度可分离卷积代替  $3 \times 3$  卷积核的传统卷积,从计算量以及参数量上看,更优于 Inception-ResNet-A 块.

### 2.2 多尺度深度可分离卷积网络的构建

考虑到本次研究所用的数据集并不大以及检测并非复杂,所以构建结构较小的网络,如图 3 所示.该网络输入图片大小为  $224 \times 224$ ,经过传统卷积层卷积,得到 32 通道的特征图,经过 5 次多尺度深度可分离卷积块的特征提取(其中步长均为 2),得到特征图的输出为  $7 \times 7 \times 512$ .紧接着该特征图输入到全局平均池化层,得到的输出经过 Dropout 层后再输入到全链接层,最后经过 Softmax 函数计算得到各类输出的概率,具体结构如表 1 所示.其中网络结构中涉及到的卷积层包括多尺度深度可分离卷积块中的卷积层均会使用归一化层以及将 ReLU 函数作为激活函数.从图 2 中可知,该网络并不大,足够处理纱线检测的问题.

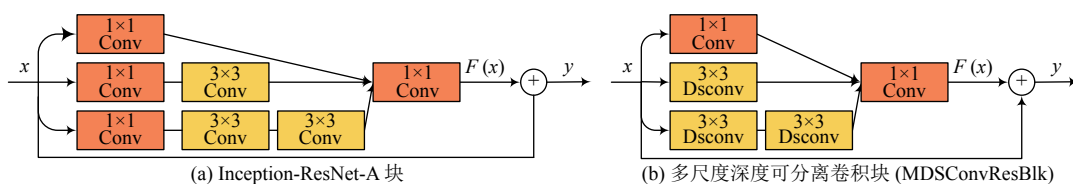


图 2 Inception-ResNet-A 块与多尺度深度可分离卷积块

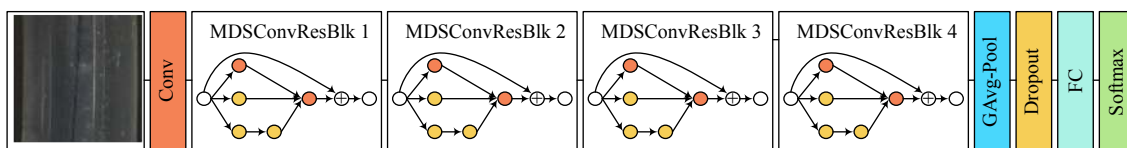


图3 多尺度深度可分离卷积网络结构

表1 基于多尺度深度可分离卷积网络结构

网络层(块)类型	核尺寸/步长	输出尺寸
Conv(卷积)	3×3/2	112×112×32
MDSConvResBlk 1(块)	1×1, 3×3/2	56×56×64
MDSConvResBlk 2(块)	1×1, 3×3/2	28×28×128
MDSConvResBlk 3(块)	1×1, 3×3/2	14×14×256
MDSConvResBlk 4(块)	1×1, 3×3/2	7×7×512
GAvg-Pool(全局平均池化)		1×1×512
Dropout		1×1×512
FC(全连接层)		1×1×6
Softmax		1×1×6

### 3 实验

#### 3.1 实验环境

训练网络的过程是基于 64 位 Windows 10 的系统下进行的,其中计算机的配置为 Inter(R) Cele(R) G4900 CPU@3.10 GHz, RAM 为 8.00 GB, 显卡为 NVIDIA GeForce GTX 1060 6 GB.

#### 3.2 实验设置以及结果分析

网络训练过程中采用 Adam 优化器,初始学习率设置为  $3e-4$ ,学习率变化指数设置为 0.3. 在训练过程中监视验证集的准确率,当统计该数值在 3 次没有降低时,学习率以 0.7 的倍数进行衰减(其中新的学习率  $lr_{new} = 0.3 * lr_{old}$ ). 其次损失评估采用的是交叉熵损失函数来衡量网络输出与真实结果的差距,最后网络模型训练的 30 轮情况如图 4 所示.

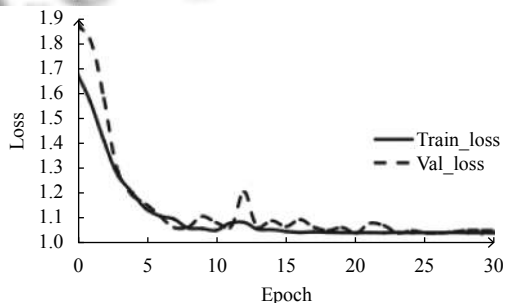
从图 4 中可知,网络初始化完成时尚未能够进行纱线检测,图片分类能力非常差.随着网络训练的不断进行,网络开始快速收敛,在接近 25 轮训练时,网络开始接近稳定,而此时验证集的识别精度能够达到 100%.随着训练次数的增加,网络训练时的验证集识别精度已经相对稳定,损失值逐渐接近于 1.

#### 3.3 模型评估

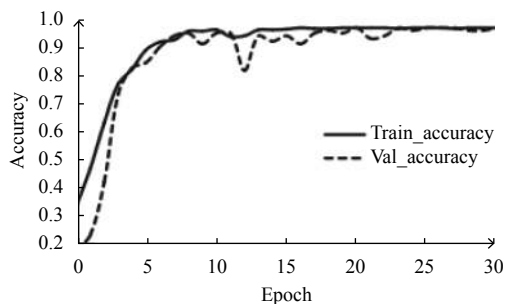
深度学习网络分类模型常有的评估方式有精确度、召回率、准确率等,本文进行模型评估采用的是准确率.准确率(Precision)是指正确分类的样本数与总样本总数之比,计算公式如式(8)所示.

$$Precision = \frac{TP + TN}{TP + FP + FN + TN} \quad (8)$$

式中, TP 指的是实际为正样本网络模型检测结果为正样本的样本数量, TN 指的是实际为负样本网络模型预测结果为负样本的样本数量,此外 FP 指的是实际为负样本网络模型预测结果为正样本的样本数量, FN 指的是实际为正样本网络模型检测结果为负样本的样本数量.



(a) 损失值曲线



(b) 准确率曲线

图4 实验结果曲线

模型评估时采用的是另外采集的 250 张图片进行测试,在测试过程中将网络模型输出概率最大的类别作为图片预测的最终结果.经过测试以及计算得到该模型的准确率为 99.6%,其具体结果如表 2 所示.

表2 模型测试结果

类别	测试数量	正确数量	准确率(%)
无纱线	52	52	100
黑色纱线	43	42	97.67
蓝色纱线	37	37	100
粉色纱线	40	40	100
白色纱线	40	40	100
黄色纱线	38	38	100
总计	250	249	99.6

## 4 结论

为了实现管道中的纱线及其颜色进行检测,使用海康威视的工业摄像头进行数据采集,经过筛选并去除重复以及有缺陷的图片后,利用剩余的图片构建数据集.此外网络主要是由多尺度深度可分离卷积块构成的,该结构块是利用 Inception v4 网络中的 Inception-ResNet-A 模块在结构上进行改进得到的,改进后的多尺度深度可分离卷积块大大降低网络的计算量以及参数量.最后在网络训练过程中对图片进行数据增强,扩大数据集的多样性,从而降低网络过拟合现象,最终实现管道中的纱线检测功能.实验结果表明,检测准确率高达 99.6%,因此该方法可以能够满足一定的实际应用需求.

### 参考文献

- 1 景慎全. 中国色纺纱行业的技术进步和发展趋势. 纺织导报, 2020, (6): 22–27.
- 2 范文东, 张鹏琴, 彭康旭, 等. 基于红外光电开关的智能空调控制系统. 计算机产品与流通, 2020, (6): 154.
- 3 LeCun Y, Bengio Y, Hinton G. Deep learning. Nature, 2015, 521(7553): 436–444. [doi: 10.1038/nature14539]
- 4 LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. Neural Computation, 1989, 1(4): 541–551. [doi: 10.1162/neco.1989.1.4.541]
- 5 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]
- 6 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2014.
- 7 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, CA, USA. 2012. 1097–1105.
- 8 Szegedy C, Liu W, Jia YQ, et al. Going deeper with convolutions. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 1–9.
- 9 Howard AG, Zhu ML, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv: 1704.04861, 2017.
- 10 Sandler M, Howard A, Zhu ML, et al. Mobilenetv2: Inverted residuals and linear bottlenecks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 4510–4520.
- 11 He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770–778.
- 12 张振亚, 倪红军. 基于深度学习的发票图像分类方法. 南通职业大学学报, 2020, 34(2): 79–83.
- 13 Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning. Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, CA, USA. 2017. 4278–4284.
- 14 Chollet F. Xception: Deep learning with depthwise separable convolutions. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 1800–1807.
- 15 蔡之灵, 翁谦, 叶少珍, 等. 基于 Inception-V3 模型的高分遥感影像场景分类. 国土资源遥感, 2020, 32(3): 80–89.
- 16 毛远宏, 贺占庄, 刘露露. 目标跟踪中基于深度可分离卷积的剪枝方法. 西安交通大学学报, 2020, 55(1): 52–59.