

文章编号: 2095-2163(2021)03-0167-05

中图分类号: R391.9

文献标志码: A

# 基于机器视觉的零部件的缺陷检测

齐金龙, 张俊峰, 戴贤萍, 张劲松, 胡 陟

(上海工程技术大学 电子电气工程学院, 上海 201620)

**摘要:** 汽车零部件在实际生产过程中, 打磨、抛光等各种无法确定的因素会在零部件的表面留下缺陷, 从而严重影响车辆的组装、制造, 存在巨大的安全隐患, 因此车辆零部件的缺陷检测十分重要。零部件缺陷检测已从传统的手工分类发展为基于机器视觉的方法。本文基于卷积神经网络对 VGG16 网络模型进行改进, 提高了模型对汽车零部件缺陷的检测精度。首先, 改进了网络模型的全连接层; 其次, 在模型中引入了 AMF-Softmax 损失函数, 在达到更优的聚类效果的同时解决了数据不平衡问题, 最终实现了零部件缺陷的识别与定位。与传统模型的缺陷检测效果比较表明, 改进的 VGG16 网络结构模型测试准确率可以达到 97.59%, 在零部件缺陷检测方面具有优越性。

**关键词:** 机器视觉; 缺陷检测; 卷积神经网络; VGG16

## Machine vision-based component defect detection

QI Jinlong, ZHANG Junfeng, DAI Xianping, ZHANG Jinsong, HU Zhi

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

**[Abstract]** In the actual production process of automobile parts, polishing, polishing and other uncertain factors will leave defects on the surface of the parts, which will seriously affect the assembly and manufacturing of vehicles, and there are huge safety risks, so the defect detection of vehicle parts is very important. Part defect detection has developed from traditional manual classification to machine vision based method. In this paper, the VGG16 network model is improved based on convolutional neural network to improve the detection accuracy of the model for automobile parts defects. Firstly, the full connection layer of the network model is improved. Secondly, AMF-Softmax loss function is introduced into the model, which solves the problem of data imbalance while achieving better clustering effect. Finally, the paper realizes the identification and location of parts defects. Compared with the traditional model, the test accuracy of the improved VGG16 network structure model can reach 97.59%, which has advantages in parts defect detection.

**[Key words]** machine vision; defect detection; Convolutional neural network; VGG16

## 0 引言

基于机器视觉的图像分类是根据图像特征使用算法对其进行识别, 类似于人眼识别的一个过程, 车辆零部件的缺陷检测是图像分类的一种应用。伴随人工智能技术的发展, 提供了准确更高、识别效果更显著的缺陷检测识别分类方法。其中, 2 种最常用的方法是基于机器学习的支持向量机(SVM)和基于深度学习的卷积神经网络法(CNN)<sup>[1]</sup>。

国内外利用机器视觉的方法进行缺陷检测的研究也越来越多, Han 等人<sup>[2]</sup>提出了基于深度学习的目标检测算法对轮毂表面缺陷检测; Liu 等人<sup>[3]</sup>提出了 Haar-Weibull-Variance 模型, 可以对钢体表面存在的缺陷进行检测; 李瑶等人<sup>[4]</sup>利用 CCD 摄像机

机器视觉系统对金属等工业生成板材表面进行在线缺陷检测。基于机器视觉的检测已逐步取代人工, 大大提高了生产效率, 降低了检测过程中漏检、错检的概率。

目前, 在一些汽车制造公司, 机器视觉被广泛用于检测车辆装配过程中的缺陷<sup>[5]</sup>。首先通过相机采集待检测的零部件图像, 进行图像预处理, 再训练处理后的图像, 构建信息模板, 然后采集目标图像, 提取特征, 完成匹配。然而, 在检测车辆零部件缺陷上仍然有许多问题, 包括零部件形状大小不一, 以及缺陷的种类, 面积的各不相同; 获取的汽车零部件图像结构复杂, 还可能存在着诸多不相关因素, 噪声较大, 难以进行特征提取。

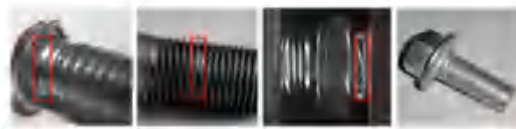
汽车零部件容易产生缺陷的 7 个最常见部位包

**作者简介:** 齐金龙(1994-), 男, 硕士研究生, 主要研究方向: 虚拟现实、手术机器人; 张俊峰(1995-), 男, 硕士研究生, 主要研究方向: 主从控制; 戴贤萍(1996-), 女, 硕士研究生, 主要研究方向: 医用介入机器人; 张劲松(1996-), 男, 硕士研究生, 主要研究方向: 主从控制; 胡 陟(1985-), 男, 博士, 讲师, 主要研究方向: 医用介入机器人。

**通讯作者:** 胡 陟 Email: huzhi26@163.com

收稿日期: 2020-12-10

括:前轴左前和中左螺栓、传动轴、驻车锁、变速器、专用转向工具、右大灯定位螺栓和减震器。本研究基于 VGG16 的改进模型,选取汽车螺栓为缺陷检测对象。3 种典型的螺栓部件缺陷,即顶部裂纹、中部磨损、底部残余、无缺陷,如图 1 所示。



(a) 顶部裂纹 (b) 中部磨损 (c) 底部残余 (d) 无缺陷  
(a) Top crack (b) Middle wear (c) Bottom residue (d) Non defects

图 1 螺栓缺陷分类

Fig. 1 Bolt defect classification

## 1 VGG 模型

卷积神经网络的深度与其性能之间的变化关系是 VGG 模型的主要研究方向,网络结构如图 2 所示,具有以下特点:

(1) 该网络模型使用  $3 \times 3$  的小卷积核作为滤波器,反复堆叠出一个深度为 16~19 层的卷积神经网络。随着网络深度的增加,模型的各方面性能也会随之提升。

(2) VGG16 网络结构含有 5 个卷积段,每段包含 2~3 个卷积层,对原始图片进行等比例缩放。而原始图片的大小需要调整为  $3 \times 224 \times 224$ ,初始卷积核大小为  $3 \times 3 \times 3$ ,中间的池化层采用  $2 \times 2$  的最大池化函数进行特征降维。



图 2 VGG 模型网络结构

Fig. 2 VGG model network structure

$3 \times 3$  卷积核如图 3 所示。由图 3 可知,在无中间池化层中,对 2 个  $3 \times 3$  卷积核进行堆叠,达到的效果能够替代一个  $5 \times 5$  卷积核,即单个  $5 \times 5$  卷积的输出与 2 个  $3 \times 3$  卷积级联输出的尺寸完全一样。除此之外,3 个串联的  $3 \times 3$  卷积层产生的参数数量也较少,更少的参数意味着减少过拟合。最重要的是,3 个  $3 \times 3$  卷积层具有更多的非线性变换。这样的网络结构设计,可以确保 CNN 具有更强的学习功能。VGG 模型中都使用了  $3 \times 3$  卷积内核和  $2 \times 2$  池化内核,可以通过深化网络结构来提高性能。

VGG16 各个卷积段中的卷积核输出数量依次为:64、128、256、512、512。VGG16 模型卷积步骤

为:

(1) 对 2 次 64 个卷积核的卷积进行处理,将处理后的结果进行池化。

(2) 经过 2 次 128 个卷积核的卷积,仍然进行一次池化。

(3) 将结果经过 3 次 256 个卷积核的卷积并池化。

(4) 还要重复 2 次 3 个 512 个卷积核卷积<sup>[6]</sup>。

期间每一次卷积处理都要完成一次池化层,所有处理完成后就是全连接层。在卷积过程中,网络中需要处理的参数量会随着卷积核数目的增加而减少。

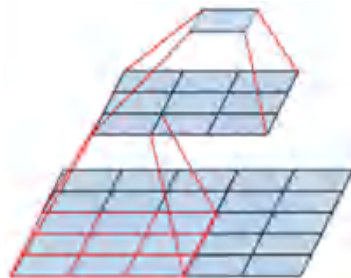


图 3  $3 \times 3$  卷积核

Fig. 3  $3 \times 3$  Convolution kernel

## 2 基于 VGG16 的改进模型

为解决检测目标区域大小不一,图像缺陷特征提取失败问题,提升对螺栓零部件的缺陷识别准确度。本文利用 VGG16 卷积神经网络的特性,在其基础上加以改进,可以提高提取螺栓零部件图像特征和识别精度。

为了提高模型在零部件缺陷检测的精度,本文在样本训练过程中,改变固定层的深度,从而达到对全连接层的改进。同时,通过改进损失函数,减少类内特征的差异性,解决螺栓零部件数据库不均衡问题。本文基于 VGG16 设计的 VGG-2CF 网络,通过增加卷积特性图中包含信息数量,从而提高识别能力。

其主要流程如下:

先是采集螺栓零部件数据集,挑选出需要识别的类别图像,作为训练集与测试集使用;由于采集的螺栓零部件图像中存在较多噪声,在螺栓图像进行特征提取前,需要进行滤波、去噪、增强等预处理操作;接着,将整体平均池化后的 Conv3 和 Conv5 经过 L2 归一化方法解决并实现连接;再将改良完成的生成特征图当作 ROI 池化层的输入;此后,将 ROI 池化层的结果输入到全连接层;在此基础上,建立改进模型,调节超参数;稍后,则由回归模型得到候选区中缺陷目标的准确位置,调节缺陷区域的边界范

围;继而,将训练改进模型;进一步地,在分类层中,由改进的分类损失函数对候选边界范围内的缺陷类别进行判断;最后,使用测试集,测试和验证模型。

VGG16的预练模式是通过 ImageNet 数据库的 1 000个类别图片集训练,可以对 1 000 个类进行分类。本研究以 VGG16 为基础,对螺栓零件缺陷的图像数据进行分类,需要用 4 个标签 softmax 分类层来替换原网络上的分类层,同时用 ReLU 作为一个激活函数。

全连接层改进步骤如下:全连接层层数和单元数确定,VGG16 全连接层为固定的 3 层结构,即:4 096×4 096×4;对全连接层进行改进,目标是将全连接层设定为 2 层,网络层结构为  $a \times 4$ ,其中  $a$  为首层单元数,分别设  $a$  为 512,1 024,2 048,3 069,4 096;对第 2 层全连接层的超参数进行设定,第 2 层全连接层结构为  $4 096 \times b \times 4$ ,其中  $b$  为第 2 层单元数,此时分别设定  $b$  为 0,100,1 024 和 2 048,当第 2 层单元层设定为 0 时,收敛最快而且稳定。由以上得出最佳的全连接网络层为  $4 096 \times 4$ 。

VGG-2CF 网络设计,首先将 VGG 的 Conv3 输出图像通过 2 次滤波,调整为  $2 \times 2$  整体平均池化;其次,Conv3 和 Conv5 通过 L2 归一化方法设置;最终,通过归一化方法连接 Conv3 与 Conv5 的输出特征,并且构建特征图。改进的 VGG-2CF 和 VGG 的共用卷积层,有效降低了计算开销,大大提高了网络收敛有效性和缺陷的识别精度。

## 2.1 ReLU 非线性激活函数

在神经网络中数据间的关系通常为非线性关系,本研究使用 ReLU 作为激活功能,来提高网络的表征能力。当输入值小于零时,输出值为零。如果输入大于或等于零,则输出等于输入。当输入值为正时,导数为 1。其定义可写为:

$$f(x) = \max(0, x). \quad (1)$$

在训练网络时,与 Sigmoid 函数相比,ReLU 函数的优势在于:网络的收敛速度要快得多;只需要一个阈值即可获取激活值;不会对模型的泛化准确性产生重大影响。

## 2.2 基于 AM-Softmax 的改进损失函数

损失函数表示实际和期望值二者间的偏离度,可以促进模型的学习。

Softmax 是对神经网络获得的多个值进行归一化处理,值在  $[0, 1]$  之间。交叉熵损失计算公式可写为:

$$L_i = - \log \frac{e^{z_{y_i}}}{\sum_j e^{z_j}} = -z_{y_i} + \log \sum_j e^{z_j}, \quad (2)$$

Softmax 总样本的丢失公式可进一步表示为:

$$L(X, Y) = - \frac{1}{N} \sum_i \sum_j 1\{j = y^{(i)}\} \log(p_{i,j}), \quad (3)$$

由于螺栓零部件数据集各个类别的样本数目相差大,数据不平衡,使用 Softmax 函数导致识别准确度不高。由于螺栓零部件数据集各个类别的样本数目相差大,缺陷特征类间差异性较小,及类内差异较大以及数据不平衡问题,使用 Softmax 函数导致识别准确度不足和分类错误,因此 Softmax 损失函数不适用螺栓零部件检测。

为了减小缺陷识别的差错率,本文引入 AM-Softmax 损失函数。具体计算公式如下:

$$L_{AMS}^{(n)} = - \frac{1}{n} \sum_{i=1}^n \log \frac{e^{s \cdot (\cos \theta_{y_i} - m)}}{e^{s \cdot (\cos \theta_{y_i} - m)} + \sum_{j=1, j \neq y_i}^c e^{s \cdot \cos \theta_j}} = - \frac{1}{n} \sum_{i=1}^n \log \frac{e^{s \cdot (W_j^T f_i - m)}}{e^{s \cdot (W_j^T f_i - m)} + \sum_{j=1, j \neq y_i}^c e^{s \cdot W_j^T f_i}} \quad (4)$$

其中,  $n$  表示现阶段的迭代次数;  $i$  表示迭代变量;  $W_j$  和  $f_i$  分别表示末尾全连接层的参数矩阵  $W$  第  $j$  列以及第  $i$  个样本输入;  $\theta_j$  表示  $W_j$  与  $f_i$  角度值;  $m$  为难易程度;  $s$  为缩放参数,将  $s$  表示为一个数值,有利于提高损失函数的收敛速度。

由于螺栓零部件图像中,各个缺陷之间的差异较小,缺陷种类的样本数据不平衡等问题,本文在 AM-Softmax 损失函数中加入  $F$ -measure, 构建成 AMF-Softmax 损失函数<sup>[7]</sup>。 $F$ -measure 表示为查全率与准确率的折衷均值,当迭代次数为  $n$  时,得到的计算公式为:

$$F_{measure}^{(n)} = \frac{(1 + \beta) P^{(n)} R^{(n)}}{\beta^2 P^{(n)} + R^{(n)}}, \quad (5)$$

AMF-Softmax 损失函数有效抑制了训练效率的降低和图像信息的遗漏。相应数学公式为:

$$L_{oss}^{(n)} = \bar{L}_{AMS}^{(n)} - \bar{F}_{measure}^{(n)}, \quad (6)$$

其中:

$$\begin{cases} \bar{L}_{AMS}^{(n)} = (L_{AMS}^{(n)} - \mu) / \sigma, \\ \bar{F}_{measure}^{(n)} = (F_{measure}^{(n)} - \mu_F) / \sigma_F. \end{cases} \quad (7)$$

由于  $F_{measure}^{(n)}$  对损失函数带来的影响,即使得在 FasterR-CNN 训练中,可以降低 RPN 输出的目标区域特征与其相契合特征间的距离,增加类间特征间



的不同,网络能够进一步提高待识别区域的目标缺陷的分类准确率。

### 3 实验分析

#### 3.1 实验环境

本文的实验环境是 Python3.5,使用 TensorFlow 框架完成对零部件的特征提取、参数设置与网络训练等。

#### 3.2 实验数据

本文的零部件缺陷数据来源于汽车生产线,包括4种缺陷,分别是:顶部裂纹、中部磨损、底部残余、无缺陷。

每一种采集样本照片600张,总共2400张原始图片,训练集每种500张,总共2000张;测试集每种100张,总共400张。因为深度学习的样本数量不足会出现过拟合,必须有足够数量的样本,方可确保识别的准确率。为了减小过拟合对样本采集增加难度,本文通过翻转、缩放、色度调整等手段处理图片,从而增加了样本照片的数量。

图像旋转的常用方法是顺时针旋转图像的中心,当平面绕圆心旋转 $270^\circ$ 时,平面上像素的坐标将发生变化。旋转前的坐标见式(8):

$$x = r \cos \alpha, y = r \sin \alpha, \quad (8)$$

坐标的旋转变换公式可分别表示为:

$$x' = r \cos(\alpha + \theta) = r \cos \alpha \cos \theta - r \sin \alpha \sin \theta = x \cos \theta - y \sin \theta, \quad (9)$$

$$y' = r \sin(\alpha + \theta) = r \sin \alpha \cos \theta + r \cos \alpha \sin \theta = x \sin \theta + y \cos \theta, \quad (10)$$

其中, $\theta=180^\circ$ 。由此得到式(11):

$$\begin{pmatrix} \hat{x}' \\ \hat{y}' \end{pmatrix} = \begin{pmatrix} \hat{x} \cos \theta - \hat{y} \sin \theta \\ \hat{x} \sin \theta + \hat{y} \cos \theta \end{pmatrix} = \begin{pmatrix} 0 & \hat{x}' \\ \hat{y}' & 0 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}. \quad (11)$$

无论是水平,还是垂直镜像变换,图像的高度和宽度都不会改变。经过一定的数据扩充后,增加了30000张图片,总共有32400张。这里的正样本有15600张,负样本有16800张。测试时训练数据量与测试数据量设置为8:2,即25920张进行训练,6480张图像进行测试。训练数据库中的25920张图片,又包括训练集和验证集。数据扩充后的样本图像如图4所示。

由于输出图像较为粗糙,背景还有多处小的噪点,要从图片中提取出目标物体,后续还要进行开运算及二值化处理,这样能够显著提高模型预测的精度。经处理后的清晰目标物体照片如图5所示。

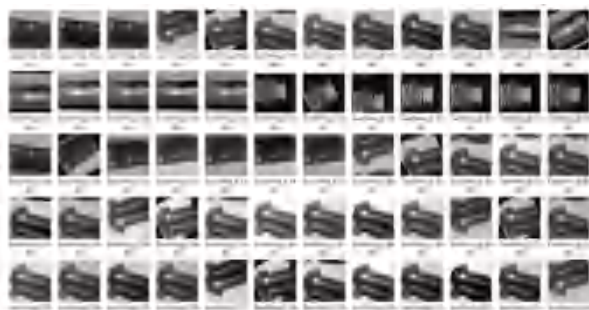


图4 数据扩充

Fig. 4 Data expansion



图5 开运算及二值化处理

Fig. 5 Open arithmetic and binarization

经过图像预处理后,改进的VGG16模型经过训练测试后得到多个缺陷图片的检测结果,如图6所示。

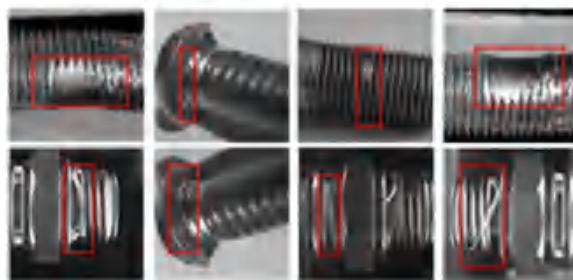


图6 缺陷检测结果

Fig. 6 Defect detection results

#### 3.3 结果分析

在改进的模型上对训练集进行训练,再对训练好的模型进行测试,通过测试集检验改进的模型准确率。训练集通过SF-VGG卷积神经网络进行了200次训练后,改进的VGG16网络模型训练以及测试时准确和损失率的变化关系分别如图7、图8所示。其中,绿色曲线表示训练集,红色曲线表示测试集。

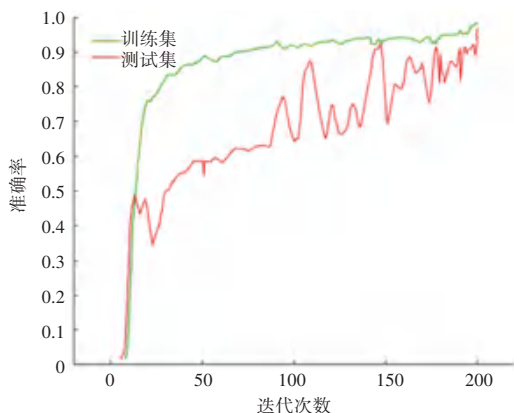


图7 准确率函数曲线

Fig. 7 Accuracy function curve

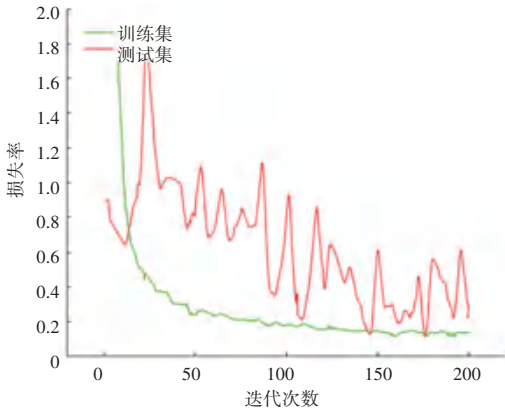


图 8 损失率函数曲线

Fig. 8 Loss rate function curve

由图 7、图 8 可以看出, 无论是训练集、还是测试集, 改进的 VGG16 的准确率都是整体上升的趋势, 损失率都是整体下降的趋势。因为测试集的数据小于训练集, 曲线有所波动, 但准确率和损失率都能够达到要求。

在此基础上, 研究得到的改进模型和传统模型的准确率对比值见表 1。由表 1 可知, 改进的 VGG16 网络模型对于螺栓零部件图像分类在训练集上的准确率可以达到 98.78%, 测试集的精度为 97.59%。而传统 SVM 方法对汽车零部件的缺陷检测的准确率只有 83.10%, 未改进的 VGG16 网络模型对汽车零部件的缺陷检测和识别的准确率为 95.26%。这表明改进的 VGG16 网络结构比传统的分类方法具有更高的识别率和鲁棒性, 提高了网络模型的检测效率和准确性, 可以对汽车零部件的缺陷进行有效的检测和分类。同时, 改进的 VGG16 网络提取方法易于实现, 训练方法简单, 计算复杂度低。

表 1 不同模型下的准确率对比

Tab. 1 Comparison of accuracy rates under different models %

模型	训练集准确率	测试集准确率
KNN	86.70	85.29
AlexNet	84.26	79.68
SVM	83.10	80.35
VGG16	95.26	93.25
Improved VGG16	98.78	97.59

## 4 结束语

本文提出一种改进的 VGG16 模型来解决汽车零部件的缺陷检测问题。基于 AM-Softmax 的改进损失函数, 结合  $F - measure$  指标, 增大类间特征差异性, 对待检测物体缺陷区域进行正确分类, 提高缺陷识别精度。在 VGG16 基础上重新设计网络结构, 增强特征图像的信息量, 提高对较小目标区域的识别能力。改进的 VGG16 网络模型, 测试集精度从 93.25% 提高到 97.59%, 识别精度大大提高。本文只针对车辆螺栓零部件的 4 种类别缺陷进行了识别和检测, 将来有必要增加零部件和缺陷种类, 考虑由于光照、遮挡和图像不清晰等诸多因素所造成的识别准确率不高的问题, 尝试使用更多改进模型, 进一步提高缺陷检测的效果。

## 参考文献

- [1] 吴鹏, 李婷, 全冲, 等. 基于 OCC 模型和 LSTM 模型的财经微博文本情感分类研究[J]. 情报学报, 2020, 39(1): 81-89.
- [2] HAN Kai, SUN MUYI, ZHOU Xiaoguang, et al. A new method in wheel hub surface defect detection: Object detection algorithm based on deep learning [C]//2017 International Conference on Advanced Mechatronic Systems. Xiamen, China; IEEE, 2017: 335-338.
- [3] LIU Kun, WANG Heying, QU Erqing, et al. Steel surface defect detection using a new Haar - Weibull - Variance model in unsupervised manner [J]. IEEE Transactions on Instrumentation and Measurement, 2017, 66(10): 2585-2596.
- [4] 李瑶, 王快社, 刘长瑞. 金属板材表面缺陷检测发展现状[J]. 材料导报, 2011, 25(2): 238-241.
- [5] 傅文军. 基于 5G 机器视觉实现某品牌汽车零部件 AI 缺陷质检案例[J]. 中国仪器仪表, 2020, 348(3): 24-26.
- [6] 田佳鹭, 邓立国. 基于改进 VGG16 的猴子图像分类方法[J]. 微型机与应用, 2020, 39(5): 6-11.
- [7] 吉卫喜, 杜猛, 彭威, 等. 基于改进的 Faster R-CNN 的齿轮外观缺陷识别研究[J]. 系统仿真学报, 2019, 31(11): 24-31.