

基于深度学习 GoogLeNet 网络的日用陶瓷缺陷检测方法研究

胡渊

景德镇陶瓷大学机电学院 江西省景德镇市 333403

【摘要】目前,日用瓷的缺陷主要由人工检测和判断,因此只能识别一些较为明显的缺陷,如大裂纹,严重变形等,缺陷识别模糊。而且这种方法会导致生产效率低下,并且加重工人的劳动负担,同时可靠性也不是很高,影响企业的生产效益。针对这种情况,本文提出利用 GoogLeNet 网络对日用陶瓷缺陷进行检测,进而提高对于日用陶瓷产品缺陷检测的准确率,减轻工人的负担,大大提高了企业的生产效率。

【关键词】深度学习; GoogLeNet 网络; 日用陶瓷; 缺陷检测

引言

日用陶瓷产品作为我国重要的工业产品,其质量检测是生产过程中的关键环节,但由于人工分拣的成本较高而且效率不高,因此不断降低成本和提高检测效率逐步成为国内外研究的重点领域。李业富等检测后将裂纹缺陷经骨架提出处理后,再由长度统计得到了裂纹长度值,并通过统计气孔缺陷区占有的像素数,计算得到气孔面积值^[1]。石志君等较系统地比较了采用 Canny 算子对二值图像进行的边缘提取,并对日用瓷产品边缘做中心对称检测,以判断日用瓷产品是否存在缺口^[2]。

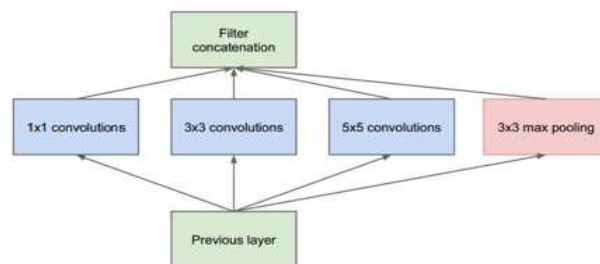
随着科技的不断发展,神经网络的层数不断增加,结构越来越复杂,而所能实现的功能也变得更加强大。在语音识别、图像处理、目标检测等领域的准确性也越来越高。深度学习的思想最早源于 Geoffrey Hinton 等人在 Science 上发表的文章^[3],深度学习不需要人工提取特征,而是让机器自动提取特征,它具有以下优点:学习能力强;覆盖范围广,适应性好;数据驱动,上限高;可移植性好。随着 CPU 和 GPU 性能的不断发展和卷积神经网络的层数也不断加深,比如出现了 AlexNet 网络, VGG-16 网络, VGG-19 网络, ResNet 网络, GoogLeNet 网络等。本文利用 GoogLeNet 网络对日用陶瓷产品进行缺陷检测。

1 GoogLeNet 网络

GoogLeNet 在 2014 年由 Google 团队提出,将 L 大写是为了致敬 LeNet 斩获当年 ImageNet 竞赛中分类任务

第一名^[4]。该网络中有以下亮点:

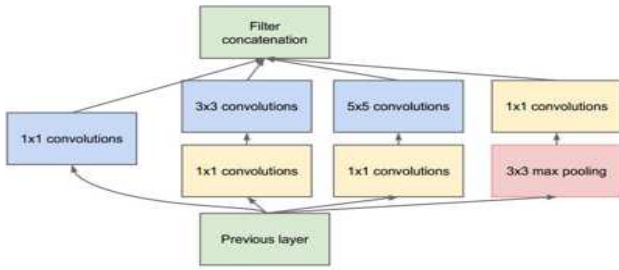
- (1) 引入了 Inception 结构(用于融合不同尺度的特征消息);
- (2) 使用 1×1 的卷积核进行降维以及映射处理;
- (3) 添加了两个辅助分类器帮助训练;
- (4) 丢弃全连接层,使用平均池化层从而大大减小模型参数。



(a) Inception module, naïve version

图 1

图 1 为 Inception 初始结构, AlexNet 和 VGG 网络都是串型结构,将一系列的卷积层和最大池化采样层进行串联; Inception 出现了并行的结构,在上一层输入之后,将得到的特征矩阵同时输入到四个分支中处理,然后将得到的四个分支的特征矩阵按深度进行拼接,就得到了输出特征矩阵。第一个分支是大小为 1×1 的卷积核;第二个分支是大小为 3×3 的卷积核;第三个分支是大小为 5×5 的卷积核;第四个分支是大小为 3×3 的池化核(最大池化采样)。通过这四个分支可以得到不同尺度的特征矩阵。并且每个分支所得到的特征矩阵的高和宽必须相等,否则无法延深度方向进行拼接。



(b) Inception module with dimension reductions
图 2

图 2 为 Inception 结构加上了降维的功能，多出来的三个 1×1 的卷积层起到了降维的作用。假设有一个深度为 512 的特征矩阵，在不使用 1×1 卷积核降维的情况下，用 64 个 5×5 的卷积核进行卷积（可以求得最后结果，即需要的参数非常多）。

2 辅助分类器 (图 3 为 GoogLeNet 图的一部分)

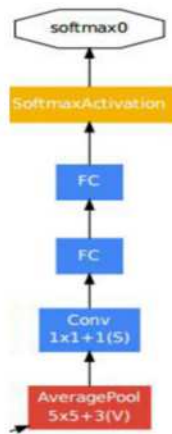


图 3

$$out_{size} = \frac{(in_{size} - F_{size} + 2P)}{S} + 1$$

第一个辅助分类器的大小为 14×14×512， in_{size} 为 14，池化核大小为 5，步长 (stride) 为 3，填充 (padding) 为 0，根据公式可以计算得到结果为 4。池化层不改变特征矩阵的深度，同理，第二个辅助分类器通过平均池化下采样操作之后，特征矩阵大小为 4×4×528。然后再采用 128 个大小为 1×1 的卷积核进行卷积，用来降低维度，并且使用 ReLU 激活函数。接着采用结点个数为 1024 的全连接层和 ReLU 激活函数，也就是所对应的第一个全连接层。然后在全连接层之间采用 dropout 函数，它以 70% 的比例随机失活神经元。最后输出层的结点个数对应类别个数，由于 ImageNet 数据集有 1000 个类别，因此结点个数为 1000，通过 softmax 激活函数得到概率分布。

根据文献中的图可以得到 GoogLeNet 使用了 9 个

Inception 模块，将原来 CNN 的串行结构改为并行，共 22 层；使用平均池化层代替全连接层；参数仅为 AlexNet 的 $\frac{1}{12}$ ；使用 softmax 获取平均结果，为了便于微调 (finetune)，最后还是加上了全连接层，最终的性能比 AlexNet 要好。

3 实验环境与方法

3.1 实验材料与系统配置

本次实验采用日用陶瓷产品作为研究对象，运用工业相机对日用陶瓷产品图像进行采集。共采集到 2350 幅图像，并且对于这些图像进行筛选，剔除掉其中由于曝光时间过长等因素造成的质量不符合实验要求的图像。最终得到 2125 幅图像，包括 855 幅正常图像和 1270 幅具有缺陷的图像。通过数据增强，得到 16750 幅 5472×3648 的日用陶瓷产品缺陷图像，将十分之一的样本用于测试集，其余的样本用于训练集。

实现本次模型构建的实验环境为：

- (1) 硬件配置：Intel-i7-5820K; GeForce-GTX-1080Ti
- (2) 操作系统：Linux
- (3) 编程语言：Python3.6
- (4) 深度学习框架：Tensorflow1.8

3.2 实验数据预处理及数据增强

运用直方图均衡化对样本进行预处理



图 4(预处理前)

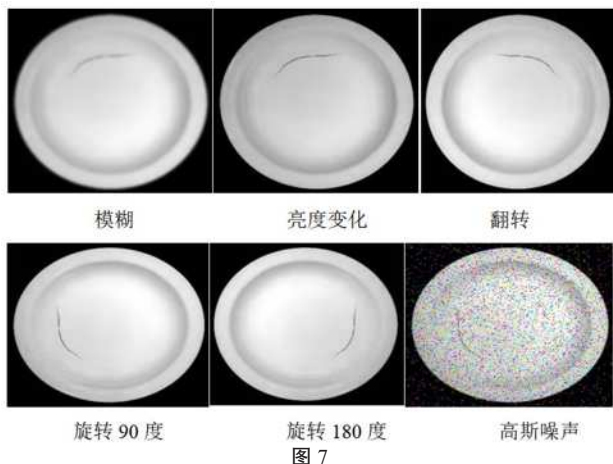


图 5(预处理后)

数据增强的方法主要有亮度变化、翻转、模糊、高斯噪声、旋转等。



图 6(原图像)



3.3 模型设计

陶瓷产品图像中的缺陷特征通过神经网络中的卷积层进行提取，而池化层则进行特征降维，从而优化卷积层中所提取到的缺陷特征。卷积神经网络的训练模型由前向传播函数和反向传播函数（即损失函数）所组成，特征向量由前向传播函数利用激活函数得到，损失函数通过不断优化更新权重以逐渐取得最小值，构造如下函数表达式：

$$C_j^l = f(z^l) = f\left(\sum_{i \in M_j} C_i^{l-1} k_{ij}^l + b_j^l\right) \quad (3-1)$$

$$L_i = \frac{1}{m} \left[\sum_{j=1}^n \sum_{l=1}^m 1\{y^l = j\} \log \frac{\exp(w_j^T x^{(i)})}{\sum_{l=1}^n \exp(w_l^T x^{(i)})} \right] + \frac{\lambda}{2} \sum_{j=1}^n W_j^2 \quad (3-2)$$

$$W_j \leftarrow W_j - \alpha \nabla_{W_j} L(W; x^{(i)}, y^{(i)}) \quad (3-3)$$

$$v \leftarrow \varepsilon v - \alpha \nabla_{W_j} L(W; x^{(i)}, y^{(i)}) \quad (3-4)$$

$$W_j \leftarrow W_j + v \quad (3-5)$$

式 3-1 表示前向传播函数，式 3-2 表示损失函数；式 3-3~3-5 表示在反向传播过程中权重参数的更新。梯度下降的方向作为 W_j 的方向， α 和 ε 分别表示学习率和冲量单元，通过式 3-4 可以得到更新速度 V ，最后再利用式 3-5 来更新权重。由于较小的学习速率可以加快网络的收敛速度，因此在实验过程中 α 和 ε 的值分别取 0.001 和 0.5。

实验结果分析：

本次实验在 Tensorflow 框架下进行，学习率 α 取 0.001，batch size 大小选取 15，学习轮次为 100，经过约 4 小时训练后得出如下结果，图 8 表示训练集和测试集的损失(loss)值,图9表示训练集和测试集的精确度(acc),表 1 为 CNN 模型的测试结果。

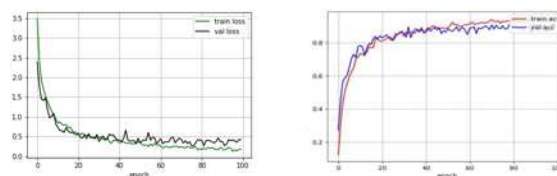


图 8 图 9

表 1 CNN 模型测试结果

CNN 模型	训练准确率 /%	测试准确率 /%	AUC
GoogLeNet	100	92.76	0.8542

本次实验 GoogLeNet 模型的 AUC 为 0.8542，说明该模型具有较强的泛化能力。GoogLeNet 模型通过自身的优势，使得训练集和测试集的准确率均较高，从而省去了人工提取特征的时间，提高了生产效率。

4 结束语

随着工业生产不断向智能化的方向发展，缺陷检测方法成为科研人员的一个重要研究方向，由于深度学习的精度较高并且稳定性较好，受到了越来越广泛的关注。本文使用 GoogLeNet 深度学习模型对日用陶瓷产品进行缺陷检测，对样本进行扩充后，测试准确率达到 92.76%，说明该模型的泛化能力较好，对于今后的生产实践活动具有借鉴意义。

【参考文献】

[1] 李业富, 赵玉刚, 姜文革. 基于图像处理的陶瓷产品缺陷识别研究[J]. 现代制造工程, 2014, 05: 109-112.
 [2] 石志君. 基于机器视觉的陶瓷圆环件缺口检测与尺寸测量[D]. 青岛: 中国海洋大学, 2013.
 [3] Hinton G E, Salakhutdinov RR. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(5786): 504-507.
 [4] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.