

基于双线性 CNN 与 DenseBlock 的导光板 标记线缺陷检测^①



张家瑜¹, 周迪斌¹, 魏东亮¹, 金秉文², 解利军³

¹(杭州师范大学 杭州国际服务工程学院, 杭州 311121)

²(杭州利珀科技有限公司, 杭州 311300)

³(浙江大学 航空航天学院, 杭州 310027)

通讯作者: 张家瑜, E-mail: 1483956702@qq.com

摘 要: 导光板标记线检测是导光板制造品控中的一个重要步骤, 但在使用传统图像算法进行检测的过程中, 有大量的气泡、严重污染和无标记线的情况存在. 因有大量气泡、严重污染和无标记线的情况造成人工特征难以设计, 因此, 使用基于卷积网络的方法来代替人工特征设计进行缺陷检测. DenseNet 卷积神经网络较其他分类神经网络具有参数较少, 梯度收敛稳定等特点. 因 DenseNet 卷积神经网络中使用特征融合的思想, 保证了图片分类准确率. 通过迁移学习的方法, 将训练得到的 DenseNet 网络权重迁移到 Bilinear-CNN 算法进行训练, 提升卷积神经网络局部注意力, 提高图像分类准确率. 通过实现结果表明, 所提方法具有可行性, 相比于 V2-ResNet-101 网络结构, 准确率提升至 95.53%, 参数减少了 97.2%, 平均单张图像检测时间减少 25%.

关键词: 深度学习; DenseNet; 迁移学习; Bilinear-CNN

引用格式: 张家瑜, 周迪斌, 魏东亮, 金秉文, 解利军. 基于双线性 CNN 与 DenseBlock 的导光板标记线缺陷检测. 计算机系统应用, 2020, 29(7): 152-159. <http://www.c-s-a.org.cn/1003-3254/7467.html>

LGP Marker Defect Detection Algorithm Based on Bilinear-CNN and DenseBlock

ZHANG Jia-Yu¹, ZHOU Di-Bin¹, WEI Dong-Liang¹, JIN Bing-Wen², XIE Li-Jun³

¹(Hangzhou Institute of Service Engineering, Hangzhou Normal University, Hangzhou 311121, China)

²(Hangzhou LEAPER Science and Technology Co. Ltd., Hangzhou 311300, China)

³(School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China)

Abstract: Light Guide Plate (LGP) marker detection is a crucial procedure in LGP manufacturing quality control but a mass of bubbling, polluted and marker-free cases may arise out of detection in traditional image algorithms. Because the mass of bubble, polluted and unmarked lines makes people difficult to design features. Compared with other classification neural networks, DenseNet Convolution Neural Network (CNN) has fewer parameters and stable gradient convergence. Because DenseNet CNN uses the idea of feature fusion, the accuracy of image classification is guaranteed. Through the transform learning method, the weights of the trained DenseNet network are transferred to the bilinear-CNN algorithm for training, which improves the local attention of the convolutional neural network and improves the accuracy of image classification. The implementation results show that the proposed method is feasible. Compared with the V2-ResNet-101, the accuracy of proposed approach is increased to 95.53%, while parameter number is decreased by 97.2%, and average single image detection time drops by 25% in the proposed network structure.

Key words: deep learning; DenseNet; transfer learning; Bilinear-CNN

① 基金项目: 国家自然科学基金面上项目 (11772301); 浙江省自然科学基金 (LY17F020016); 杭州师范大学第二轮专业学位研究生课程教学案例库建设 (1115B20500159)

Foundation item: General Program of National Natural Science Foundation of China (11772301); Natural Science Foundation of Zhejiang Province (LY17F020016); Second Round of Teaching Case Database Construction for Graduate Courses of Hangzhou Normal University (1115B20500159)

收稿时间: 2019-11-20; 修改时间: 2019-12-16; 采用时间: 2020-01-02; csa 在线出版时间: 2020-07-03

导光板是一种亚克力材质的塑料板材,经过花纹烫印后,被广泛用于显示器背光模组核心组件。对导光板标记线的检测,则是验证当前烫印工艺是否合格的一个重要标准。使用数字图像算法检测导光板标记线虽然可以有效地对图片进行检测,但是面对特征更加复杂的大面积气泡、严重污染和无标记线等情况无法得到满意的结果。使用深度学习方法进行工业检测,是学术界和工业界共同探索的课题。使用深度学习的方法对小数据集、低对比、高纹理、高相似的导光板标记图片进行缺陷检测,导光板缺陷板材示意图如图1所示。

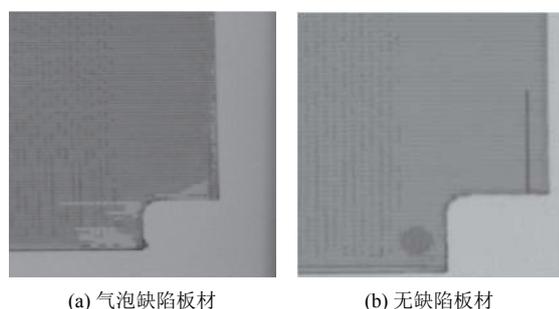


图1 缺陷板材示意图

目前越来越多的中外学者致力于机器视觉的研究,但却少有导光板检测的成熟方法以及学术论文可被参考,通过借鉴液晶屏缺陷检测的研究方法进行研究。在数字图像算法检测方法中,文献[1]中将图像作为像素矩阵,利用奇异值分解(SVD)检测缺陷,首先去除图像背景,然后利用剩余奇异值重建图像并显示缺陷区域。但是这种方法中没有提供消除各种缺陷类型时必须的奇异值数,以获得良好的实验结果。文献[2]中,利用奇异值对图像进行分类,并将缺陷图片分为粗缺陷和细缺陷两类,采用二维离散小波系数与区域增长相结合的方法提取缺陷区域,该算法在时间复杂度上表现较为优异,但是在应对导光板复杂特征时表现欠佳。文献[3]中扩展了原霍夫变换的点到线的距离差,在非平稳灰度梯度的缺陷检测中得到了有效的应用。然而不同的距离差值与相同的控制参数组合将会导致得到不同的结果,文中并未提到此自动检测方法是否需要人工干预来设置参数。在神经网络的检测算法中,文献[4]中作者为Mura缺陷的定量评估设计了一个模糊神经网络,然而对于低对比度的Mura缺陷,如何准确地提取到缺陷特征是一个挑战。文献[5]中作者为预测Gap Mura缺陷使用4种神经网络,分别是BP神经网络、径向基函数、学习向量量化1和学习向量量化

2进行训练,但是BP神经网络本身具有收敛速度慢、权值收敛局部极小化等问题,并不能提供稳定的解决方案。文献[6]中作者针对Mura缺陷采用监督学习网络,将观察到的Mura缺陷标记为特征,并被输入到一个卷积神经网络提取所选择的特征。文中所使用的网络需要提前表及特征,过程繁杂,工作量大,且文中使用网络结构简单单一,捕获特征较少,无法应对复杂纹理的图像分类任务。文献[7]中作者使用神经网络光学模型和遗传算法等方法优化设计棱镜型液晶导光板,以求在现有条件下获得优异的亮度均匀性。但是使用的网络BP神经网络训练方面问题多多,无法保证使用不同的数据集时的检测效果。文献[8]中作者融合STN和DenseNet卷积神经网络组成新型网络结构ST-DenseNet用于树种叶片数据集分类。文献[9]中作者通过改进AlexNet完成遮挡猕猴桃目标识别任务。文献[10]中作者使用VGG16分类卷积神经网络对血液细胞进行诊断分类研究。在理论研究方面,人们在探索分类卷积神经网络时注意到了注意力这一概念,例如:文献[11]中简单的形状检测,文献[12-15]中基于数字的任务,文献[16]中的捉住的视觉控制游戏;文献[17]中的正面人脸检测等,都是面向分类卷积神经网络中注意力机制的探索和研讨。

上述分类卷积神经网络的分类效果虽然表现不错,但是由于网络参数较多,运行时卷积神经网络的时间复杂度较大。因工业产品检测需要考虑时间成本和资金成本,无法直接使用上述分类网络进行分类检测。由于目前用于工业检测的卷积神经网络案例较少,查阅相关文献后,选择对现有卷积神经网络进行改进,完成导光板分类检测算法。针对工业检测需求,更改DenseNet网络中的卷积结构,结合Bilinear-CNN算法提高卷积神经网络局部特征注意力。最终在900张的测试图片组成的测试数据集上达到了95.53%的准确率,平均单张图像检测时间为40.1 ms。

1 相关理论方法

卷积神经网络(Convolutional Neural Networks, CNN)是一种特殊的人工神经网络,其主要的特点是卷积运算操作。卷积神经网络在进行分类或者回归的时候,大致都分为两个部分,他们分别是前馈运算和反馈运算。前馈运算是产生运算数据的过程,其主要通过卷积操作、汇合操作和非线性激活函数映射等一系列操

作共同完成. 反向传播主要是计算预测值和真实值之间的误差或损失, 凭借反向传播算法将损失逐层向前反馈, 更新每一层的权重和偏置, 通过周而复始的进行前馈运算和反馈运算, 最终得到期望的模型.

假设某网络有 L 层, 用 X_0 表示整个网络的输入, X_{L-1} 表示第 $L-1$ 层的输出. 这里的层并不是单指某一个操作, 它可以指卷积操作、汇合操作、非线性映射或其他变化中的任意单独的一个操作, 也可以是不同形式变化的任意组合. 这里使用 $H_l(\cdot)$ 作为第 L 层的非线性变化.

在 CNN 中, 前一个卷积层的输出是后一个卷积层的输入.

$$x_l = H_l(x_{l-1}) \tag{1}$$

在 ResNet 中加入了一个跳跃连接, 它绕过了具有恒等函数的非线性转换. 第 l 层的输出是 $l-1$ 层的输出与 $l-1$ 层非线性变化后的结果相加.

$$x_l = H_l(x_{l-1}) + x_{l-1} \tag{2}$$

在 DenseNet 中, 与 ResNet 最大的不同则是在于, DenseNet 中的每个 DenseBlock 将所有特征计算单元 (BN-ReLU-Conv (1×1) -BN-ReLU-Cov (3×3)) 的结果以密集连接的方式沿某一维度进行压缩^[18,19], 代替了简单的求和. 一个具有 i 层的 DenseBlock 结构 (结构如图 2

所示), 其密集连接数目为 $i \times (i + 1) / 2$. 这让后续的层可以访问前面任何层所学习到的结果, DenseNet 网络结构如图 3 所示.

DenseBlock 中特征计算表达式如下所示:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \tag{3}$$

公式中的 $[x_0, x_1, \dots, x_{l-1}]$ 分别代表在 DenseBlock 中的 $H_l(\cdot)$ 的结果张量.

$$P(y = i|x) = \frac{e^{h(x,y_i)}}{\sum_{j=1}^n e^{h(x,y_j)}} \tag{4}$$

其中, $P(y = i|x)$ 表示样本 x 属于第 i 类的概率.

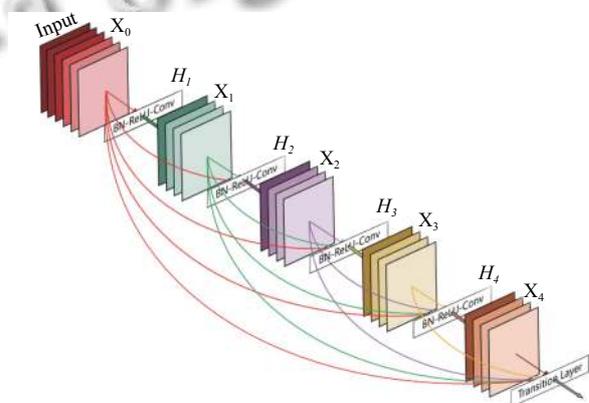


图 2 DenseBlock 结构示意图

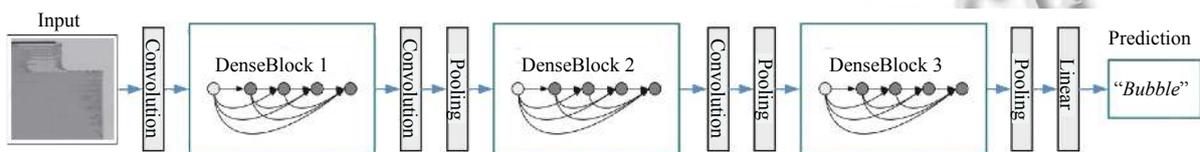


图 3 DenseNet 网络结构

2 基于双线性 CNN 和 DenseBlock 的分类检测框架

本文算法分上下两步, 第 1 步, 基本权重训练: 通过修改 DenseNet 网络, 改进 DenseBlock 中特征计算单元的结构, 减少卷积神经网络卷积层数量, 避免过拟合, 以及减少卷积神经网络参数数量. 第 2 步, 权重模型优化: 使用 Bilinear-CNN 分类算法组成混合深度模型进行迁移学习, 将使用第 1 步训练得到的权重迁移到第 2 步的 Bilinear-CNN 分类网络重新训练, 通过捕捉特征通道之间的成对相关性, 提高分类准确率, 各步网络结构如图 4 所示. 通过对实验的卷积神经网络算法的

权重进行分析发现, 本文所设计的卷积神经网络算法仅需要极少的参数参与运算.

2.1 改进的神经网络模型介绍

与传统的为提高识别精度, 通过增加卷积层数量或使用金字塔特征融合的方法不同. 卷积神经网络在工业检测应用的场景, 主要运行在工控机或者普通 PC 机上. 因此设计卷积神经网络时, 首先考虑减少卷积神经网络卷积层, 然后再考虑提高卷积神经网络分类精度.

2.1.1 减少卷积神经网络参数

减少 DenseBlock 中的卷积层数, 通过减少卷积操作次数和卷积数量, 提高卷积神经网络运行速度. 减少

卷积层数量可以在减少运行时间的同时减少网络参数数量. 但是随之而来的副作用, 则是无法获得足够的特征, 使得特征激活映射区域无法正确映射到目标物体上.

减少卷积操作中的通道数量, 随着卷积操作中通

道数量的增加, 网络参数数量也随之增加, 使得网络在运行时需要申请更多的显存和内存空间. 减少卷积操作中的通道数量, 可以有效的降低卷积神经网络的空间复杂度.

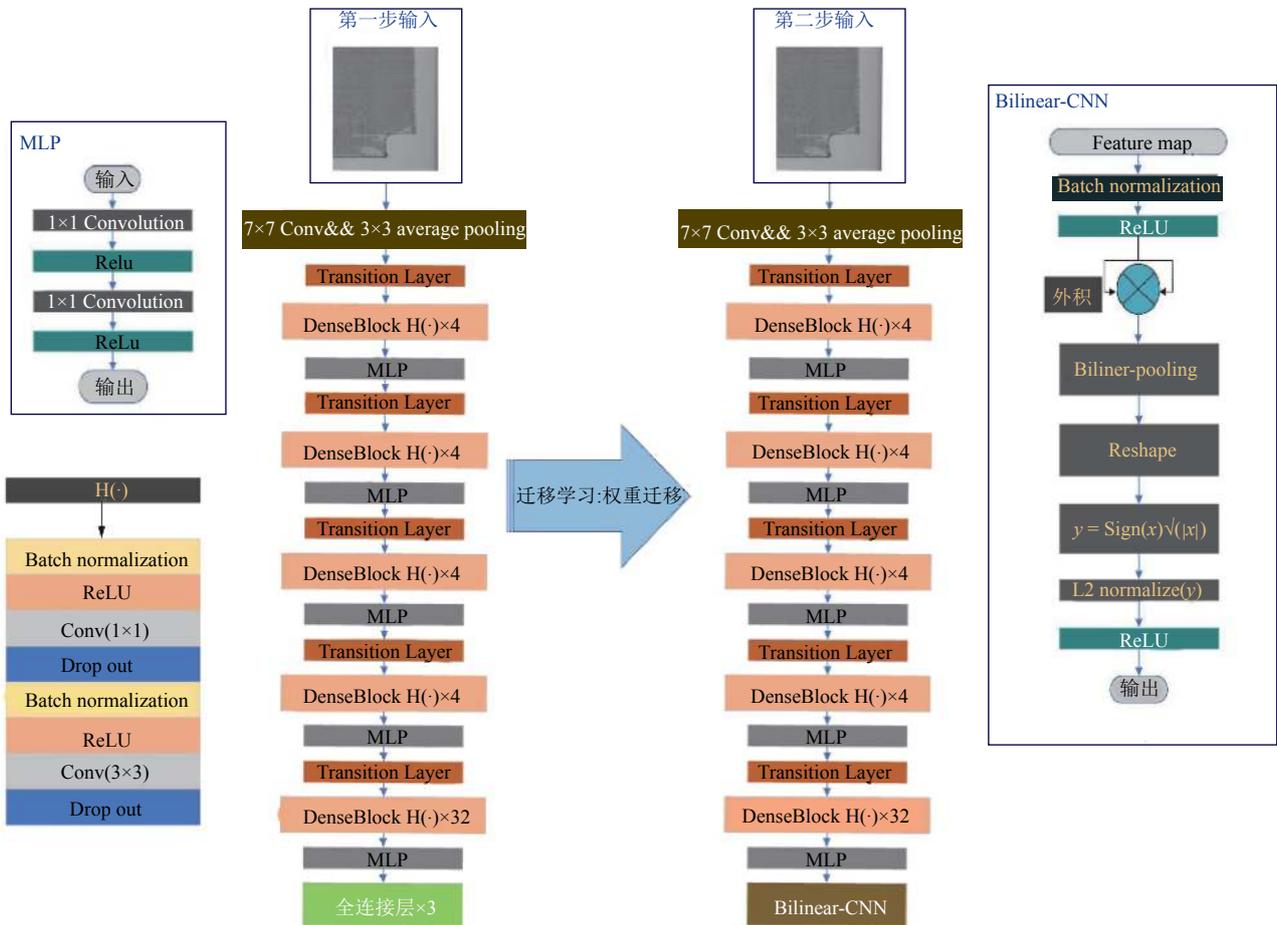


图4 本文算法网络结构

在特征计算单元中加入 DropOut 算法, 通过设置忽略特征检测器的概率, 使某些神经元的激活值以一定的概率停止工作. DropOut 的加入可以使当前模型泛化性更强, 减少对某些局部特征的依赖, 从而避免过拟合.

在每个 DenseBlock 后面加入多层感知器 (MLP)^[20], 与 Caffe 中 MLP 的实现方法不同, 实验算法中的 MLP 算子中没有使用 3x3 大小的卷积核进行卷积, 仅使用两层 1x1 大小的卷积核进行卷积, 这样做有助于在加深网络卷积深度与增强网络容量和复杂度的同时, 减少网络参数数量^[21].

2.1.2 加速损失函数收敛

在实验中, 发现使用原 DenseNet 网络结构中的

GAP 结构, 无法有效降低损失函数的损失值. 在使用三层全连接层代替 Global Average Pooling (GAP) 之后, 损失函数的损失值快速下降. 对比前后结果, 三层全连接层作为一个小型网络, 相比于 GAP 算法其可以获取更多的特征. 使用 GAP 算法对高纹理、低对比的灰度图分类时, 会造成特征图上的特征丢失.

GAP 对所接收到的每一个特征图都进行了极端降维, 它将尺寸为 $h \times w \times d$ 的张量缩小为 $1 \times 1 \times d$ 大小的张量. GAP 激活区域与分类结果之间的关系证明如下所示^[22]:

设 $f_k(x, y)$ 为卷积层在输入图片 (x, y) 处得到的 K 个单元的映射. 对于 K 个单元执行全局平均池化的结果为:

$$F_k = \sum_{x,y} f_k(x,y) \quad (5)$$

假设当前所属类别为C, 令 w_k^c 为K个单元所属类别C对应的权重, 其中 w_k^c 代表了类别C中的 F_k 的重要性.

$$M_c(x,y) = \sum_k w_k^c f_k(x,y) \quad (6)$$

因此得出 $S_c = \sum_{x,y} M_c(x,y)$, 由此可以证明, $M_c(x,y)$ 代表着空间网格 (x,y) 处的激活, 并且导致图片分类到类别C.

使用三层全连接层可以有效地降低损失函数损失值, 加速了梯度下降, 促使网络可以更加快速的拟合权重参数, 找到当前网络的局部最优解. 但是随着三层全连接层的使用, 导致卷积神经网络的参数数量增加, 所以在下一步的迁移学习过程中通过修改全连接层结构的方法, 降低网络参数数量.

2.2 双线性 CNN 算法

双线性 CNN(如图 5) 由两组 CNN 特征提取器组成, 它使用两个卷积神经网络的输出在图像的每个位置进行外积相乘, 并在原图各个位置进行池化以获得最终的图像特征描述符. 通过外积操作捕捉特征通道之间的成对相关性, 对局部特征之间的交互进行建模. 由于双线性 CNN 的两个 CNN 的输出是线性的, 所以被叫做双线性 CNN^[23]. 在双线性 CNN 算法中, 有并行子网络 A、B, 网络 A 的作用是对图像中对象的特征部位进行定位, 而网络 B 则是用来对网络 A 检测到的特征区域进行特征提取, 由这两个网络共同完成了图像分类中最重要的两个任务: 区域检测和特征提取^[24].

双线性 CNN 的全局图像描述符, 表示如下:

$$B(X) = \sum_{s \in S} x_s x_s^T \quad (7)$$

其中 $X = (x_1, \dots, x_{|S|}, x_s \in R^c)$ 是一组局部描述符, S是空间位置的集合(行和列的组合). 其中局部描述符 x_s , 通常使用 Fisher vector、SIFT 等特征, 也可以使用卷积神经网络的正向传播提取特征. 描述符 $B(X)$ 是一个大小为 $c \times c$ 的矩阵, 因描述符本质上是一个向量, 也可以将其称为长度为 c^2 的向量.

$$bilinear(l, I, f_A, f_B) = f_A(l, I)^T f_B(l, I) \quad (8)$$

使用池化函数 P 用于聚合图片中所有区域的双线性特征, 其在实现的时候有两种方法, 一种聚合方法是将所有双线性特征累加, 另外一种是使用 max-pooling

进行双线性特征聚合.

累加聚合如下表示:

$$\Phi(I) = \sum_{l \in L} bilinear(l, I, f_a, f_b) \quad (9)$$

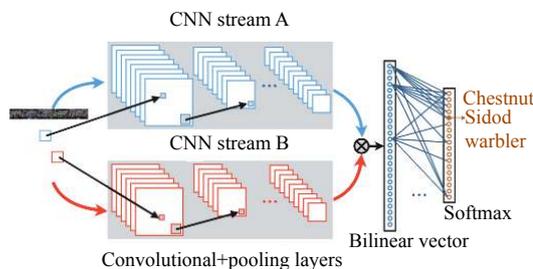


图 5 双线性 CNN 结构图

目前双线性 CNN 有两种实现方法, 第 1 种: 使用不同的两组卷积神经网络作为特征提取器, 第 2 种: 使用同一组网络获取的最终特征描述符进行运算. 通过对比研究, 因考虑到算法最终的应用环境, 最终选择第 2 种实现方式进行实验.

2.3 损失函数

在第一步的实验训练中尝试了不同的损失函数, 其中包括 focal loss 和交叉熵损失函数^[24], 在单目标分类实验中发现, 交叉熵损失函数的表现更为稳定, 最终选择交叉熵损失函数进行分类训练.

交叉熵损失函数公式如下所示:

$$L = -[y \log \hat{y} + (1 - y) \log (1 - \hat{y})] \quad (10)$$

3 实验分析

训练数据集共有 4059 张图片, 验证和测试数据集各有 896 张图片. 为应对不同工厂的不同产品, 旨在寻找简单易用的端到端的神经网络, 本文分类卷积神经网络算法在训练时并未对图片预处理.

实验电脑硬件配置如下:

- (1) CPU: Intel Core(TM) i5-4590; 四核四线程; 主频: 3.3 GHz.
 - (2) 实验训练用 GPU: NVIDIA Titan V; 显存: 12 GB; 内存: 250 GB.
 - (3) 实验测试用 GPU: NVIDIA GTX 1060; 显存: 6 GB.
 - (4) 本机内存: DDR3: 1600 MHz; 内存: 16 GB.
- 实验软件环境如下:
- (1) 操作系统: Windows 10, ubuntu 16.04.

(2) 编程语言: Python.

(3) 第三方库: Pillow, numpy, matplotlib, tensorflow, keras, scikit-image.

3.1 DenseNet 实验参数与结果

在第 1 步的训练中, 将训练集输入到网络中进行模型训练, 训练过程所涉及的参数有batch_size大小、学习率、激活函数、迭代次数和优化函数(训练参数设置如表 1 所示).

表 1 参数设置

训练参数	值
batch_size	4
学习率	0.0001
激活函数	ReLU
优化函数	Adam
迭代次数	约 60 次

使用实验网络模型对同一个数据集、在相同的训练参数下, 进行多次训练后发现, 每次训练后得到的准确率在 92% 左右震荡. 从侧面验证, 通过改进所得到的卷积神经网络具有较好的鲁棒性(loss 变化图如图 6 所示, 本文使用留一法数据进行展示, 用以凸显卷积神经网络拟合效果).

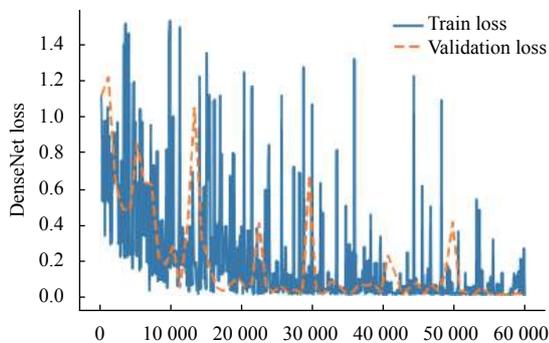


图 6 loss变化图

3.2 双线性 CNN 实验结果与分析

使用迁移学习技术对双线性CNN网络进行训练, 通过修改第一步中的DenseNet网络结构, 使用双线性 CNN 算法替换三层全连接层, 在减少网络参数的同时提高了分类准确率. 在第二步训练双线性 CNN 时依然采用交叉熵损失函数作为图像分类训练的损失函数(如图 7).

将训练集输入到网络中进行模型训练, 训练过程所涉及的参数有batch_size、学习率、激活函数、迭代次数和优化函数等(如表 2).

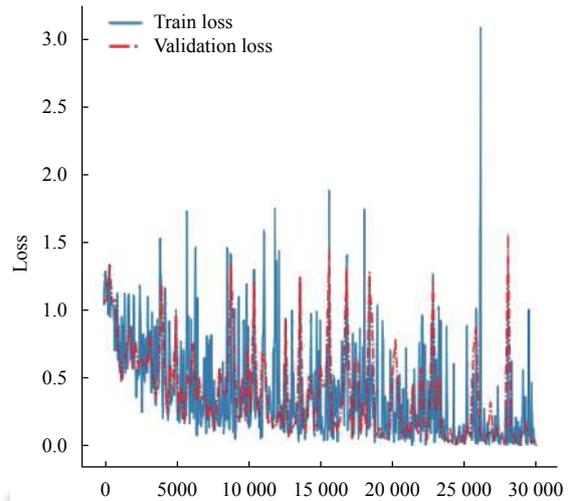


图 7 双线性 CNN loss 变化图

表 2 参数设置

训练参数	值
batch_size	4
学习率	0.1
激活函数	ReLU
优化函数	MomentumOptimizer
迭代次数	约 30 次
训练时间	约 3 小时

实验所用训练数据集(如图 8)由 3 类缺陷图片组成, 分别是无线、有线和气泡, 每类有 1353 张图片, 共计 4059 张训练图片, 图片大小为 $448 \times 448 \times 1$ 的灰度图片. 测试数据集和验证数据集各有 896 张图片, 在上述数据集中无线、有线和气泡的缺陷图片各有 300 张测试图片, 图片大小为 $448 \times 448 \times 1$ 的灰度图片. 使用实验算法在测试数据集上测试, 最终得到分类准确率为 95.53%.

经过对测试数据集图片内容逐张图片验证发现, 误判主要于有线板材被严重污染的情况下出现. 在对气泡和无线板材进行逐张图片验证时发现, 这两类图片的检测准确率为 100%. 根据工厂质量品控标准, 可以将合格的检测为不合格的, 但是不能存在不合格检测为合格的情况. 所以无线板材的准确检测是保障导光板质量的最低标准, 实验算法基本实现了导光板标记线的检测要求. 实验网络分类结果如图 9.

使用 VGG-16、V2-ResNet-101、AlexNet 与本文网络相对比, 分类结果如表 3 所示.

根据表 3 数据所示, DenseNet-BCNN 与其他主流网络对比结果如下.

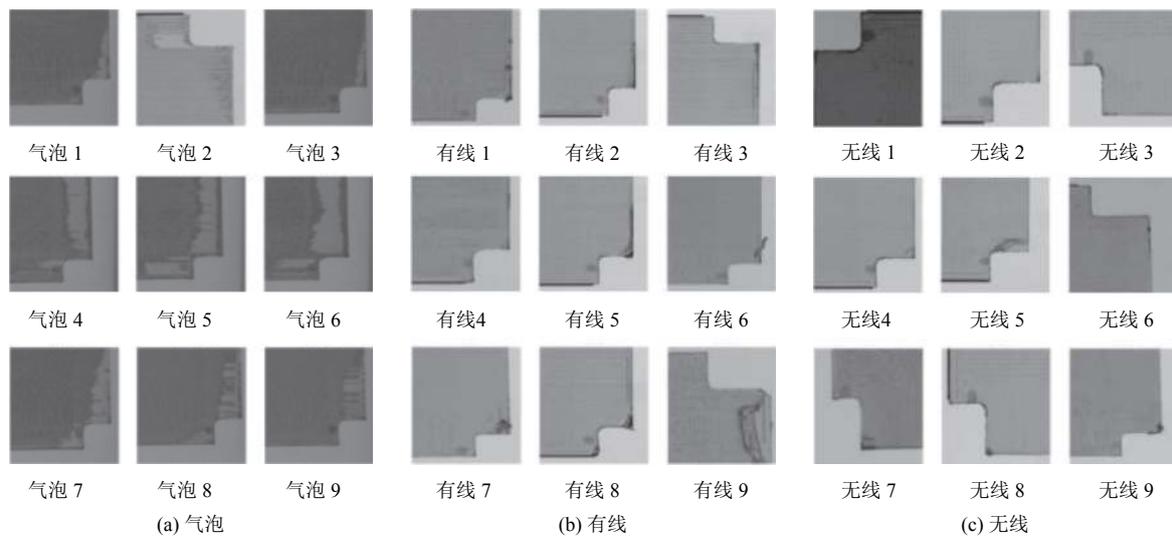


图8 数据集示意图

表3 分类网络结果对比

Architecture	Param. (MB)	GFLOPs	Accuracy	Recall	Precision	F1-score	Times (ms)
VGG-16	24.94	114.55	81.15	∅	∅	∅	∅
AlexNet	45.23	5.353	92.19	0.894	1.0	0.9440	10.6
V2-ResNet-101	40.58	58.411	95.09	0.953	1.0	0.9759	53.4
DenseNet-BCN	1.14	5.519	95.53	0.961	0.99	0.9752	40.1

(注: ∅ 为无意义项)

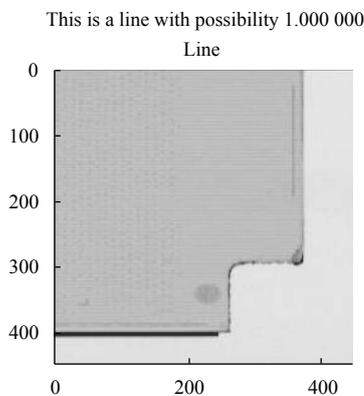


图9 实验网络分类结果示意图

在网络参数方面, 较 V2-ResNet-101 卷积神经网络参数减少了 97.2%, 较 AlexNet 卷积神经网络参数减少了 97.5%。

在准确率方面, 较 V2-ResNet-101 卷积神经网络准确率提升了 0.44%, 较 AlexNet 卷积神经网络准确率提升了 3.34%。

在平均单张图片检测时间方面, 较 V2-ResNet-101 卷积神经网络快了 13.3 ms, 较 AlexNet 卷积神经网络慢了 29.5 ms。

4 结语

本文卷积神经网络主要分两步完成, 第 1 步: 通过修改 DenseNet 网络结构降低参数数量、提高检测速度、减少图片检测所用时间以及避免训练过拟合。第 2 步: 使用迁移学习方法, 将第 1 步训练所的权重迁移到双线性 CNN 进行进一步训练, 通过外积操作捕捉特征通道之间的成对相关性, 提高分类准确率。使用本文算法在低对比、高纹理、高相似的灰度图片小数据集上进行分类, 最终取得了 95.53% 的准确率。相比于 V2-ResNet-101 网络结构, 准确率提升了 0.44%, 参数减少了 97.2%, 平均单张图片检测时间大约减少 25%。因为使用机器视觉的方法对导光板缺陷进行检测的相关文献较少, 使用深度学习方法的文献更加稀少。因此现有的导光板检测方面的文献, 无法提供有价值的参考和对比。由于工场计件生产规则限制, 无法在工厂生产线上进行长时间的、大负荷的、连续的算法鲁棒性和稳定性评测。但是在与传统算法搭配进行缺陷检测的过程中, 就目前的算法检测结果来看, 已基本满足工业生产的缺陷检测需要。下一步, 在面对减少缺陷检测时间、提高缺陷检测准确率、保证算法鲁棒性和稳定性等方面, 需要更深入的研究和分析。

参考文献

- 1 Lu CJ, Tsai DM. Automatic defect inspection for LCDs using singular value decomposition. *The International Journal of Advanced Manufacturing Technology*, 2005, 25(1-2): 53-61.
- 2 Wang JW, Chen WY, Lee JS. Singular value decomposition combined with wavelet transform for LCD defect detection. *Electronics Letters*, 2012, 48(5): 266-267. [doi: [10.1049/el.2011.3746](https://doi.org/10.1049/el.2011.3746)]
- 3 Li WC, Tsai DM. Defect inspection in low-contrast LCD images using Hough transform-based nonstationary line detection. *IEEE Transactions on Industrial Informatics*, 2011, 7(1): 136-147. [doi: [10.1109/TII.2009.2034844](https://doi.org/10.1109/TII.2009.2034844)]
- 4 Zhang Y, Zhang J. A fuzzy neural network approach for quantitative evaluation of Mura in TFT-LCD. *Proceedings of 2005 International Conference on Neural Networks and Brain*. Beijing, China. 2005. 424-427.
- 5 Chen LF, Su CT, Chen MH. A neural-network approach for defect recognition in TFT-LCD photolithography process. *IEEE Transactions on Electronics Packaging Manufacturing*, 2009, 32(1): 1-8. [doi: [10.1109/TEPM.2008.926117](https://doi.org/10.1109/TEPM.2008.926117)]
- 6 Li TY, Tsai JZ, Chang RS, *et al.* Pretest gap Mura on TFT LCDs using the optical interference pattern sensing method and neural network classification. *IEEE Transactions on Industrial Electronics*, 2013, 60(9): 3976-3982. [doi: [10.1109/TIE.2012.2207658](https://doi.org/10.1109/TIE.2012.2207658)]
- 7 Li CJ, Fang YC, Cheng MC. Study of optimization of an LCD light guide plate with neural network and genetic algorithm. *Optics Express*, 2009, 17(12): 10177-10188. [doi: [10.1364/OE.17.010177](https://doi.org/10.1364/OE.17.010177)]
- 8 张锡英, 宋宇鹏, 陈曦, 等. 融合 STN 和 DenseNet 的深度学习网络及其应用. *计算机工程与应用*, 2019, 55(17): 131-136. [doi: [10.3778/j.issn.1002-8331.1901-0094](https://doi.org/10.3778/j.issn.1002-8331.1901-0094)]
- 9 穆龙涛, 高宗斌, 崔永杰, 等. 基于改进 AlexNet 的广域复杂环境下遮挡猕猴桃目标识别. *农业机械学报*, 2019, 50(10): 24-34. [doi: [10.6041/j.issn.1000-1298.2019.10.003](https://doi.org/10.6041/j.issn.1000-1298.2019.10.003)]
- 10 张海涛, 刘景鑫, 赵晓晴, 等. 基于 VGG16 的急性淋巴细胞白血病血液细胞显微图像人工智能辅助诊断分类研究. *中国医疗设备*, 2019, 34(7): 1-4, 9. [doi: [10.3969/j.issn.1674-1633.2019.07.001](https://doi.org/10.3969/j.issn.1674-1633.2019.07.001)]
- 11 Schmidhuber J, Huber RM. Learning to generate artificial fovea trajectories for target detection. *International Journal of Neural Systems*, 1991, 2(1-2): 125-134.
- 12 Larochelle H, Hinton GE. Learning to combine foveal glimpses with a third-order Boltzmann machine. *Proceedings of the 23rd International Conference on Neural Information Processing Systems*. Vancouver, BC, Canada. 2010. 1243-1251.
- 13 Bazzani L, de Freitas N, Larochelle H, *et al.* Learning attentional policies for tracking and recognition in video with deep networks. *Proceedings of the 28th International Conference on International Conference on Machine Learning*. Madison, WI, USA. 2010. 937-944.
- 14 Denil M, Bazzani L, Larochelle H, *et al.* Learning where to attend with deep architectures for image tracking. *Neural Computation*, 2012, 24(8): 2151-2184. [doi: [10.1162/NECO_a_00312](https://doi.org/10.1162/NECO_a_00312)]
- 15 Ranzato MA. On learning where to look. *arXiv: 1405.5488*, 2014.
- 16 Mnih V, Heess N, Graves A, *et al.* Recurrent models of visual attention. *Proceedings of the 27th International Conference on Neural Information Processing Systems*. New York, NY, USA. 2014. 2204-2212.
- 17 Tang YC, Srivastava N, Salakhutdinov R. Learning generative models with visual attention. *Proceedings of the 27th International Conference on Neural Information Processing Systems*. New York, NY, USA. 2014. 1808-1816.
- 18 Huang G, Liu Z, Van Der Maaten L, *et al.* Densely connected convolutional networks. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI, USA. 2017. 2261-2269.
- 19 吴仁彪, 赵婷, 屈景怡. 基于深度 SE-DenseNet 的航班延误预测模型. *电子与信息学报*, 2019, 41(6): 1510-1517. [doi: [10.11999/JEIT180644](https://doi.org/10.11999/JEIT180644)]
- 20 Singh G, Sachan M. Multi-layer perceptron (MLP) neural network technique for offline handwritten Gurmukhi character recognition. *Proceedings of 2014 IEEE International Conference on Computational Intelligence and Computing Research*. Coimbatore, India. 2014. 1-5.
- 21 Zhang T, Qi GJ, Xiao B, *et al.* Interleaved group convolutions. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy. 2017. 4383-4392.
- 22 Zhou BL, Khosla A, Lapedriza A, *et al.* Learning deep features for discriminative localization. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA. 2016. 2921-2929.
- 23 Lin TY, RoyChowdhury A, Maji S. Bilinear CNN models for fine-grained visual recognition. *Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile. 2015. 1449-1457.
- 24 Lin TY, Goyal P, Girshick R, *et al.* Focal loss for dense object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318-327. [doi: [10.1109/TPAMI.2018.2858826](https://doi.org/10.1109/TPAMI.2018.2858826)]