

江南大学

全日制专业硕士学位论文

题 目:	基于图像处理和深度学习的		
	PCB 图像缺陷检测		
英文并列题目:	PCB Image Defect Detection Based on		
	Image Processing and Deep Learning		
研 究 生:	秦岭	专 业 领 域:	电气工程
研 究 方 向:	工业装备与节能控制		
导 师 ① 姓 名:	朱一昕	职 称:	副教授
导 师 ② 姓 名:	潘波	职 称:	高级工程师
学位授予日期:	2021 年 12 月		
	答辩委员会主席:		谢林柏

江 南 大 学

地址: 无锡市蠡湖大道 1800 号

二〇二一年 十二月

摘要

工业生产、制造水平的不断提高，加快了电子产品的更新换代，因此对于底层基础硬件设施提出了更高的要求。印刷电路板（Printed Circuit board）作为集成电路的重要基础元器件，其质量决定着电子产品的整体性能，但在 PCB 生产过程中难以避免因为各种因素导致制作产生的 PCB 带有缺陷。传统的 PCB 缺陷检测由人工标注其缺陷信息，效率较低，准确率也不高，近几年基于深度学习的缺陷检测算法逐渐得到研究，但是传统的卷积神经网络针对于 PCB 图像缺陷检测，难以兼顾全局与细节处的特征信息，检测的正确率以及标记的准确程度都不高。因此，针对于 PCB 图像的缺陷检测问题，本文提出了一种基于图像处理和深度学习的 PCB 图像缺陷检测算法。本文主要内容如下：

（1）首先，针对 PCB 缺陷图像中的特征信息不明显，并包含有很多与特征无关的噪声信息，提出了对图像进行图像灰度化、灰度直方图均衡化、高斯滤波等图像预处理方法，一方面除去噪声干扰突出特征信息，另一方面也为后续的神经网络训练提供训练和验证数据集。

（2）其次，在预处理过的 PCB 缺陷图像数据集上，通过机器视觉技术实现全局缺陷检测和定位。基于参照对比算法，使用 SURF 提取 PCB 标准图像和缺陷图像的特征点，并对特征点进行匹配实现图像配准，再用阈值分割将图像处理成二值图，随后进行图像差分，得到缺陷图像中包含的缺陷信息，通过形态学处理让缺陷更加清晰，以实现 PCB 图像的初步检测和定位。

（3）然后，基于深度学习方法对 PCB 图像进行特征提取。传统的残差神经网络 Resnet101 有效解决了网络层数过深带来的梯度弥散问题，但是低网络层高分辨率对图像语义特征提取不足，高网络层低分辨率对图像细节特征提取不足，本文提出多尺度特征融合的方式改进网络，以提升特征提取的质量，形成了 Resnet101-finetune 网络。通过迁移学习的方式在 PCB 图像数据集上训练该网络并进行验证，实验结果表明改进后的 Resnet101-finetune 神经网络的平均准确率达到了 94.2%，相比于其他神经网络得到明显提升。

（4）最后，基于 Faster-RCNN 算法实现 PCB 图像的缺陷目标检测。通过改进 Faster-RCNN 算法提升目标检测的效率和准确度。在相同的 PCB 图像数据集上采用 YOLO、R-CNN、SSD 和本文提出的算法进行对比，实验结果表明本文提出的算法在平均准确率上达到 95.5%，mAP 值也均高于其他算法，验证了本文算法的有效性，并通过实例检测结果验证了本文算法的通用性，可以适用于 PCB 图像缺陷检测，具有较高的实际使用价值。

关键词：图像处理；深度学习；残差神经网络；目标检测；缺陷检测

Abstract

The continuous improvement of industrial production and manufacturing levels has accelerated the upgrading of electronic products, and therefore put forward higher requirements for the underlying infrastructure hardware facilities. The printed circuit board is an important basic component of integrated circuits. Its quality determines the overall performance of electronic products. However, it is difficult to avoid PCB defects due to various factors during the PCB production process. Traditional PCB defect detection uses manual labeling of its defect information, which has low efficiency and low accuracy. In recent years, deep learning-based defect detection algorithms have gradually been studied, but traditional convolutional neural networks are aimed at PCB image defect detection. It is difficult to take into account the global and detailed feature information, and the accuracy of detection and the accuracy of marking are not high. Therefore, aiming at the problem of PCB image defect detection, this thesis proposes a PCB image defect detection algorithm based on image processing and deep learning. The main contents of this article are as follows:

(1) First of all, in view of the fact that the characteristic information in the PCB defect image is not obvious and contains a lot of noise information that is not related to the characteristic, it is proposed to perform image preprocessing such as image grayization, grayscale histogram equalization, and Gaussian filtering on the image. The method, on the one hand, removes noise interference to highlight feature information, on the other hand, it also provides training and verification data sets for subsequent neural network training.

(2) Secondly, on the preprocessed PCB defect image data set, global defect detection and positioning are realized through machine vision technology. Based on the reference comparison algorithm, SURF is used to extract the feature points of the PCB standard image and the defect image, and the feature points are matched to achieve image registration, and then the image is processed into a binary image by threshold segmentation, and then image difference is performed to obtain the defect image. The defect information contained is processed by morphology to make the defects clearer, so as to realize the preliminary detection and positioning of the PCB image.

(3) Then, feature extraction of PCB images based on deep learning methods. The traditional residual neural network Resnet101 effectively solves the problem of gradient dispersion caused by too deep network layers, but the low network layer high resolution is insufficient for image semantic feature extraction, and the high network layer low resolution is insufficient for image detail feature extraction. The method of multi-scale feature fusion is proposed to improve the network to improve the quality of feature extraction, forming the Resnet101-finetune network. The network is trained and verified on the PCB image data set

through transfer learning. The experimental results show that the average accuracy of the improved Resnet101-finetune neural network reaches 94.2%, which is significantly improved compared to other neural networks.

(4) Finally, the defect target detection of PCB images is realized based on Faster-RCNN algorithm. Improve the efficiency and accuracy of target detection by improving the Faster-RCNN algorithm. YOLO, R-CNN, SSD and the algorithm proposed in this thesis are compared on the same PCB image data set. The experimental results show that the algorithm proposed in this thesis achieves an average accuracy of 95.5%, and the mAP value is also higher than other algorithms. Verification The effectiveness of the algorithm in this thesis is verified, and the versatility of the algorithm in this thesis is verified through the example detection results. It can be applied to PCB image defect detection and has high practical value.

Keywords : image processing ; deep learning ; residual neural network ; target detection; defect detection

目 录

摘 要.....	I
Abstract	II
目 录.....	IV
第一章 绪论.....	1
1.1 课题研究背景及意义	1
1.2 国内外研究现状	3
1.2.1 PCB 缺陷检测方法的研究发展	3
1.2.2 图像处理和深度学习的研究发展.....	4
1.2.3 PCB 缺陷检测算法的研究现状	6
1.3 论文的主要研究内容与章节	8
第二章 基于图像处理的全局缺陷检测.....	10
2.1 PCB 缺陷图像数据集.....	10
2.2 图像增强	11
2.2.1 图像灰度化.....	12
2.2.2 灰度直方图均衡化.....	12
2.2.3 降噪滤波.....	15
2.3 缺陷全局检测和定位	19
2.3.1 图像配准.....	19
2.3.2 图像差分.....	23
2.3.3 形态学处理.....	25
2.4 本章小结	26
第三章 基于卷积神经网络的 PCB 缺陷检测	27
3.1 人工神经网络	27
3.1.1 感知器与多层感知器.....	27
3.1.2 反向传播算法.....	28
3.1.3 卷积神经网络.....	30
3.2 改进多尺度特征融合残差神经网络	33
3.2.1 残差神经网络模型结构.....	34
3.2.2 多尺度特征融合.....	35
3.2.3 损失函数与激活函数.....	38
3.2.4 训练集采集及模型数据迁移训练.....	39

3.2.5 实验对比.....	42
3.3 本章小结	44
第四章 基于 Faster-RCNN 的 PCB 缺陷目标检测	45
4.1 目标检测算法	45
4.1.1 two-stage 网络	45
4.1.2 one-stage 网络	47
4.2 基于 Faster-RCNN 的目标检测	48
4.2.1 Faster-RCNN 网络结构.....	48
4.2.2 模型训练.....	51
4.3 算法评价标准	52
4.4 实验结果与分析	53
4.4.1 实验环境与数据.....	53
4.4.2 实验结果分析.....	53
4.4.3 算法检测实例.....	56
4.4.4 不同算法检测结果对比.....	57
4.5 本章小结	59
第五章 主要结论与展望.....	60
5.1 主要结论	60
5.2 展望	61
参考文献.....	63
附录：作者在攻读学位期间发表的论文.....	67

第一章 绪论

1.1 课题研究背景及意义

伴随着科学技术的快速发展，越来越多的设备被研发制造出来，极大地改善了人们的生活。而近些年，随着工业时代 4.0 的到来，工业制造、生产的水平得到了大幅提升，也随之加快了电子产品的更新换代的速度。当前大数据、5G、人工智能等各式各样的全新技术正在火热发展，这些新兴技术的快速发展离不开高性能的工业产品，因此对于底层基础硬件设施的质量以及性能提出了更高的要求。

印刷电路板（Printed circuit board, PCB）是在 1936 年由奥地利人保罗·埃斯勒（Paul Eisler）发明出来的，又被称为印刷线路板或印制电路板。它是集成电路（Integrated Circuit, IC）中的重要基础部件，它作为载体进行安装和集成了各种电子元器件，并承载着连接各个电子元器件之间的导线，使其作为一个整体发挥作用，是微电子产品中至关重要的组成部分。PCB 技术最早出现在军用收音机当中使用，能够减小设备体积，加快传输速度，但是其技术相对不成熟且造价昂贵，因此在其他领域很难广泛推广。自 1950 年以来，集成电路技术不断得到发展，集成的复杂程度也由简入深，印刷电路板也随之由小规模生产发展到如今的特大规模。现如今，PCB 在各个领域都得到了极为广泛的运用，我们日常生活中使用的智能手机、便携平板电脑等都离不开 PCB 技术^[1]。

我国的微电子产业比起欧美发达国家来说起步较晚，但是近年来发展速度反而大幅超过其他国家，涌现出一大批优秀的集成电路企业，像华为海思、豪威科技、北京智芯微等等，为我国的 PCB 技术的发展做出了极大贡献，推动了电子产品行业快速发展。2021 年 7 月底发布的 NTI 全球 PCB 百强企业名单中的前十名如表 1-1 所示^[2]：

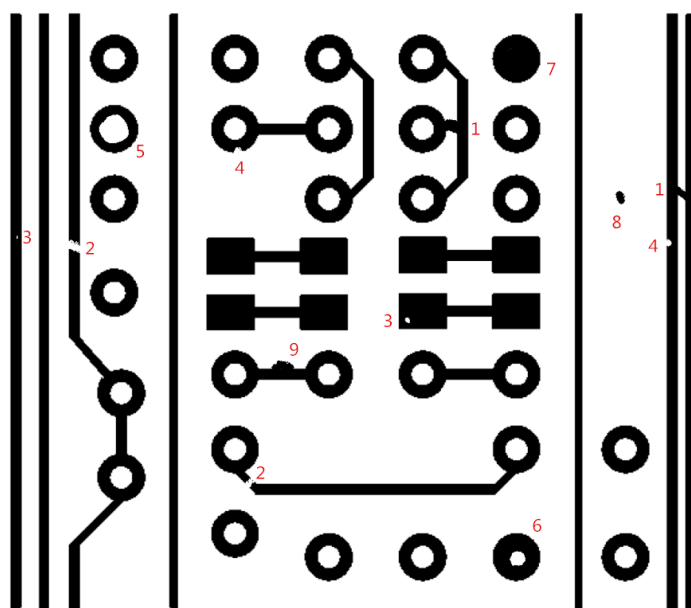
表 1-1 2021 年 NTI 全球 PCB 百强企业前十名

排名	企业		国家/地区	产值（亿美元）		增长率
	英文名	中文名		2019	2010	
1	ZDT	臻鼎	中国台湾	40.80	44.54	9.2%
2	Unimicron	欣兴	中国台湾	28.01	29.82	6.5%
3	DSBJ	东山精密	中国	21.50	27.19	26.5%
4	Nippon Mektron	旗胜	日本	26.52	26.39	-0.5%
5	TTM	迅达	美国	22.38	21.05	-5.9%
6	Compeg	华通	中国台湾	19.06	20.54	7.8%
7	Tripod	健鼎	中国台湾	18.51	18.85	1.8%
8	Shennan Circuits	深南电路	中国	15.25	16.80	10.2%
9	Ibiden	揖斐电	日本	12.38	15.56	25.7%
10	HannanStar	瀚宇博德	中国台湾	14.64	15.51	5.9%

2019 年中国内资 PCB 企业首次进入全球十强，并且近三年排名不断提升，2021 年中国大陆内资企业的东山精密和深南电路已经荣登第三和第七位，并且 2020 一年的产值增长率高达 26.5% 和 10.2%，相比于其他国家地区的产值增长较少甚至负增长，我国的 PCB 发展不可谓不快，并且随着国家对工业技术的不断深入，PCB 研发生产将会进一步提高，接下来在全球总产值的占比也将不断提升。

电子技术的不断发展使得人们不再仅仅满足于电子产品能够满足其使用的需求，而是对电子产品质量、性能、外观等等提出了更高的要求。随着技术不断地更新换代，硬件设施也不断有了更高的要求，PCB 正朝着微小型、精密型、轻便型、多层数、超薄型等新的目标发展，以满足不同领域的不同需求。PCB 作为电子产品中不可缺少的基础一环，PCB 的质量性能决定了后续电子产品的整体性能，然而在现代工业批量生产 PCB 的过程相当复杂，包含有多道工序，每一个步骤操作不慎都有可能产生制作的 PCB 产生各种各样的缺陷，进而使电子产品的性能和产品安全不达标。

最早开始研发 PCB 的美国逐步开始加大对 PCB 的质量控制技术的关注，对于 PCB 产出的精度控制有了巨大提升，并据此提出了 PCB 生产的标准。根据美国电子电路和电子互连行业协会所制定的 IPC 标准（Institute for Printed Circuits）以及电子产品实际工程生产过程的需求对 PCB 裸板的外观可接受条件、拒收条件进行了详细划分^[3]。PCB 图像中的缺陷有常见的短路、断路、毛刺、缺孔、多铜、针孔等等。缺陷的各种类型如图 1-1 所示。



1.短路；2.断路；3.针孔；4 缺损；5.过度蚀刻；6 刻蚀不足；7.缺孔；8.多铜；9.毛刺

图 1-1 PCB 缺陷示意图

因此便产生了对于 PCB 表面是否有缺陷而进行检测的需求^[4]，尤其是在精密领域，例如航空航天、微型芯片等对于 PCB 产品的质量和安全要求非常高，所以对于 PCB 表面进行缺陷检测显得尤为重要。我国现如今是全球第一大 PCB 生产大国，每年制造产生的 PCB 数以亿计，但同时因制造过程难以避免产生的残次品也进入到成品中，导致

我国制造 PCB 的精度始终难以赶上其他国家，因此如何进行 PCB 缺陷检测成为我国发展电子技术的关键难题。对 PCB 图像进行缺陷检测对于集成电路企业提高其 PCB 生产效率、提高产品质量、提升国际竞争力都有着至关重要的作用。对 PCB 缺陷检测进行的技术研究有以下几方面意义：

（1）对于生产包含 PCB 的电子产品的制造业有现实需求。企业需要对生产过程中制造出的 PCB 进行质量检测，以提高产品精度，但传统的 PCB 缺陷检测方法离不开人工，迫切需要采取新的计算机技术来解决这一难题。

（2）取代人工目测，避免不同工人对于产品的判断标准不一致或同一工人不同工作时间因劳累导致产品标准差别等人为主观因素的影响，采取新型技术进行 PCB 缺陷检测提高检测准确度、加快检测效率、减少因质量问题导致的安全问题。

（3）对检测出的 PCB 的缺陷类型进行分类分析研究，以改进 PCB 生产线，减少 PCB 产生缺陷的可能性，进而减少报废数量，提升 PCB 产量和精确度，对于节约资源、减少产能浪费、提高资源利用率也有着重要意义。

上世纪 90 年代，国外的大型 PCB 制造企业就已经意识到 PCB 缺陷检测技术在未来市场竞争的重要作用，并开始对其投入大量精力进行研究。而我国对于 PCB 领域的研究起步稍晚，对于缺陷检测更是没有完整成熟的技术，我国是 PCB 制造大国，因此研究出具有自主知识产权的 PCB 缺陷检测技术具有非常重要的意义和实用价值。

1.2 国内外研究现状

随着图像处理技术和深度学习技术的不断发展，为 PCB 缺陷检测这一问题提供了新的解决思路。传统的 PCB 缺陷检测常采用人工裸眼检测、电气检测或红外检测等^[5]，存在着检测准确率低，容易漏检、误检等问题。通过图像处理技术能够将 PCB 图像的复杂信息简单化，突出检测的重点部分。再通过深度学习技术替代传统的图像识别能够加快对于 PCB 图像的缺陷进行识别和分类的准确度和效率^[6]。此外，经过全世界不同研究所的努力，提出了很多能够有效实现 PCB 缺陷检测的算法，而我国在缺陷检测算法仍处于探索阶段，需要继续深入研究。

1.2.1 PCB 缺陷检测方法的研究发展

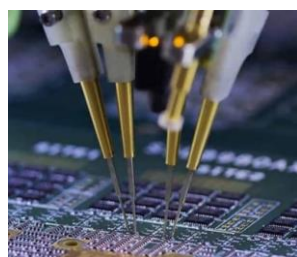
（1）人工检测。PCB 缺陷检测最早采用的是由生产线工人裸眼直接目检并手动挑出残次品来完成，如图 1-2（a）所示，生产效率极低并且很难满足企业实际的生产质检需求。人工检测主要通过制定统一标准，然后由工作人员按照自己对标准的理解进行检测。不同的人在不同的环境下对于检测标准的理解很难达到一致，PCB 大多缺陷细小难以观察，因此工人检测的准确度又受到人眼分辨率和灵敏度的制约，当 PCB 上电路导线之间的距离小于 0.2mm 时，人工检测的错误率高达 30% 以上，导致整个产品报废。并且工作人员在生产间长时间进行工作的状态下，也很容易因为疲劳或厌烦而导致检测结果波动。随着制造企业不断扩大、生产力水平不断提高、生产数量大量提升，人工目测的检测方式已经很难适应企业的生产需求，也被逐步淘汰。

(2) 电气检测。根据电路板的电路设计, 利用惠斯通电桥逐个测试电路板上网点之间的阻抗特性, 检测其是否与 PCB 设计目标一致, 如图 1-2 (b) 所示, 多用于检测电路是否短路或断路, 对于其他缺陷则无法检测, 具有一定的局限性。但由于需要仪器接触电路板表面进行检测, 用力不慎极易容易在电路板表面产生凹痕, 形成新的缺陷, 并且对每一个电路板的每一个网点进行测试耗时很长, 检测效率很低, 检测缺陷范围局限, 因此只适用于小规模电路板生产的缺陷检测, 不能满足企业大规模制造生产的需求。

(3) 红外检测。红外检测的原理是当电路板上存在缺陷时, 电流经过此处会因为受阻不均匀导致发热产生的红外线存在差异, 通过红外扫描仪检测电路板表面的温度差异进而确定电路板是否存在缺陷, 并能够准确定位到缺陷的位置, 如图 1-2 (c) 所示, 但是对于缺陷的类型并不能够明显区分。挪威的公司 Elkem 首次提出该方法并将其运用在 PCB 缺陷自动检测系统, 但是在实际操作过程中, 仅能检测出 PCB 的一些裂纹缺陷, 应用局限性较大。



(a) 人工目检



(b) 电气检测



(c) 红外检测

图 1-2 传统的 PCB 缺陷检测方法

1.2.2 图像处理和深度学习的研究发展

图像处理问题的研究由来已久, 最早在 20 世纪 20 年代产生对于图像数据进行处理分析的需求, 但是碍于当时硬件设施十分简陋, 相关技术发展并未成熟, 导致图像处理方法的研究始终进展缓慢。直至 20 世纪 60 年代第三代计算机面世, 计算机的运算速度达到了每秒几十万次到几百万次, 运算速度的大幅提升使得更多算法的实现成为可能, 数字图像处理技术得以快速发展^[7]。

图像处理指的是将图像数据进行信息化处理加工, 对图片进行数字化显示, 减少在对目标物进行图像采集时造成的细节损失, 以及将目标物的原始特征尽可能复原并强调, 以方便后续的研究。图像分析则是利用计算机技术对图像数据进行边缘检测^[8]、特征提取^[9]、图像分割^[10]等操作。而图像检测是研究在复杂背景下如何将所需目标检测出来的问题, 识别并标注目标。总体来说, 原始的图像数据信息冗杂, 而图像处理、分析和识别, 就是将图像中的信息熵进行提炼的过程, 图像中包含的信息越多, 与所需目标的无关特征信息就越多, 信息熵也越大, 因此将关键的特征信息进行提炼, 突出主要目标, 弱化无关信息, 就是图像处理的全过程。

早在 1994 年, Stricker 等人根据颜色信息首次提出了颜色直方图的概念^[11], 颜色直方图统计了一幅图像中每种不同的颜色占据整幅图像颜色的比例。针对于颜色的分布方式, Su 等人在 2011 年通过 HSV 空间上提出了颜色直方图的概念^[12]。Stricker 和 Orengo 等人提出了颜色矩^[13]的概念: 图像上的颜色信息不仅可以通过颜色直方图反映比例, 还可以使用颜色矩来表示空间分布, 包含有一阶矩、二阶矩、三阶矩等等, 并且颜色的主要信息主要集中在低阶矩上, 因此可以通过只用低阶矩来表示图像的色彩分布信息。

与颜色特征不同的是, 纹理特征同样作为图像的一种显著特征, 虽然容易被人眼识别, 但是很难以确切的定义去判断不同纹理, 因此通常采用图像像素点以及其相邻邻域内的灰度值分布来体现。1973 年, Haralick 等人首次提出了描述图像纹理特征的描述子, 灰度共生矩阵 (Gray Level Co-occurrence Matrix) ^[14]。随后 Ojala 等人提出了 LBP 算子^[15], LBP 算子主要表征了图像的局部纹理信息, 随着不断地研究深入, 渐渐完善成为旋转不变 LBP 算子。后来渐渐发展除了新的纹理特征提取方法, 有马尔可夫随机场 (MRF) 模型法、小波变换和 Tamura 纹理特征等等。

图像的内容千变万化, 想要通过单一的一种特征很难描述出图像完整的信息, 借助某一个特征进行的图像检索的结果也并不十分满意, 所以研究人员开始尝试将不同的特征融合, 采取多特征相结合的方式图像检测以提高准确率。不同的特征能够反映出不同部位的信息, 整体信息和细节信息也都不一样, 因此特征也可以被分为全局特征和局部特征。全局特征能够反映图像的整体信息, 往往对框架有很好的分类效果, 但是具体到某张确定的图像时, 局部特征才能更好的代表一张图片, 因此局部特征成为研究的关键。1999 年 David Lowe 等人提出了 SIFT 局部特征描述子^[16], 时至今日, 仍然是各种基于内容的图像检索技术的适用对象。此外, Huang 等人在前人的基础上改进了 SIFT, 形成了 SURF 描述子^[17], 还有 HOG 特征^[18]以及 GIST 特征^[19]等等都取得了良好进展。迄今为止, 图像处理技术经过几十年的发展已经能够在很多不同领域得到实现应用, 例如图像文本识别、人脸识别、医疗成像技术、指纹比对技术、机器人目标检测等等。

深度学习算法最早起源于人工神经网络技术, 该技术模拟人脑中的神经元结构, 尝试使用计算机进行学习过程。其中, 卷积神经网络 (Convolutional Neural Networks, CNN) 是一种典型的人工神经网络^[20], 这种神经网络包含有卷积计算, 并且有多层深度结构。卷积神经网络最早的研究, 是来自于日本的一位学者福岛邦彦 (Kunihiko Fukushima), 他在 1979 年提出了一个和传统神经网络不一样的, 具有多层深度的 neocognitron 模型, 因此被认为是最早的卷积神经网络。而第一个具有完整结构的卷积神经网络, 是由 Alexander Waibel 等人在 1987 提出的, 名为时间延迟网络 (Time Delay Neural Network, TDNN)。1988 年, Wei Zhang 提出了第一个二维卷积神经网络, 平移不变人工神经网络 (SIANN), 并将其应用于医学影像检测中。同样在 1988 年, Yann LeCun 同样构建了一个专用于机器视觉的卷积神经网络, LeNet 的最初版本^[21], 其包含两个卷积层, 两个全连接层, 并且在结构上与现代的卷积神经网络已经非常相

似了, Yann 在论述此神经网络时, 首次使用了卷积一词, 卷积神经网络的名字便由此而来。

直至 2012 年, Krizheysky 等人在图像识别大赛上首次提出了一种新的卷积神经网络, AlexNet^[22]。相比于 LeNet, 其具有更深的网络层, 同时采用了非线性的激活函数 Relu 方法大获成功, 夺得比赛冠军。在此之后, 2014 年牛津大学提出了 VggNet^[23], 相较于前两者, 他的网络层数更深, 并首次提出了卷积块的概念。虽然更深的网络层, 带来了更加高的精确度, 但是当网络层级增加到一定数目时, 模型变得难以训练, 测试卷积神经网络的速度以及准确度都会下降。为了解决这一网络层不断加深带来的梯度爆炸的问题, 何凯明等人研发出了残差神经网络 ResNet^[24]。ResNet 采用了独特的残差块结构, 使得输入层能够直接连接到输出层, 因此整个网络只需要学习残差, 大大简化了中间参数的数量。残差神经网络使得超级深的网络层成为可能, 避免了不断加深的网络层导致的准确率饱和的问题。因此, 目前在图像分类领域中, 残差神经网络运用最为广泛。

1.2.3 PCB 缺陷检测算法的研究现状

近几十年来, 世界各地研究学者在 PCB 图像缺陷检测领域提出了许多不同的缺陷检测算法, Moganti 学者首先在文献[25]中提出将缺陷检测算法分为以下三类: 参照对比算法 (Reference Image Comparison Method)、无参照对比算法(Non-reference Image Comparison Method)和两者结合的混合算法(Hybrid Method)。

(一) 参照对比算法, 也称为差影法 (Reference Image Comparison Method)。文献[26]首次提出采用异或 (XOR) 逻辑计算的方法进行 PCB 缺陷检测, 该算法是利用标准 PCB 图像和待检测 PCB 图像之间的真实像素值或特征之间的比对, 这类方法可以准确检测出 PCB 中存在的明显缺陷, 如短路、断路、多铜、缺损等等。图像异或法 (XOR) 是最具有代表性的参照对比算法, 该算法运算简单直接, 将两者图像的像素值进行比较即可, 因此运行速度很快, 实现简单。图像异或是一种基于图像灰度值进行的处理方式, 分别对标准 PCB 图像和待检测 PCB 图像进行阈值分割, 得到其各自的标准二值图 $I_{template}$ 和待检测二值图 I_{sample} , 那么可以得到两者的异或结果:

$$I_{XOR} = XOR(I_{sample}, I_{template}) \quad (1-1)$$

其中当 I_{sample} 在点 (x, y) 处的灰度值与 $I_{template}$ 在点 (x, y) 处的灰度值相同时, I_{XOR} 的值取 0, 否则取 1。通过异或操作得到了两图之间灰度值差异非零区域即为缺陷, 随后进一步确定缺陷类型即可。但此种方法对于 PCB 的标准图像的精确度要求极高, 而在生产过程中包含的不确定因素影响制作结果太多, 因此往往很难有一个确定的值, 而是在一定范围内波动, 使得参照对象难以有明确标准。

(二) 无参照对比算法^[27], 也被称为设计规则校验法 (Non-reference Image Comparison Method)。这种方法按照 PCB 在制作过程中预设的功能对 PCB 图像中的图案进行设计规则, 假如 PCB 图像与设计的规则不相符, 则可以认定 PCB 裸板不符合生产要求, 以此来检验 PCB 是否存在缺陷。无参照对比算法可以检测出一些与 PCB 裸板

的图形形状尺寸等相关的缺陷类型，而不需要特意提供标准 PCB 图像，因此并不需要一套固定的标准去框定，但是制定的设计规则往往因为 PCB 的复杂程度而使得计算量巨大，并且在实际工程中，一套规则难以适应成千上万不同 PCB 的需求，因此也具有一定的局限性。

如今，随着深度学习技术的飞速发展，基于深度学习的缺陷检测算法在缺陷检测方面表现出明显的改进。文献[28]提出了一种基于深度学习神经网络的微小型缺陷检测的方法，该方法包含了基于经典神经网络的三个 DCNN（用于小型、中型和大型）的结构。D Soukup 等人在文献[29]中使用金属表面缺陷的光度立体图像数据集进行训练神经网络，得到了良好的检测结果。然而，这些基于传统的卷积神经网络的方法能够检测出缺陷的存在与否，但是对于缺陷的定位以及分类则很难完成。Cha Y J 等人在文献[30]提出了一种基于 Faster R-CNN 的缺陷检测算法来检测建筑结构中存在的一些缺陷：混凝土裂缝、钢筋腐蚀、螺栓腐蚀、钢筋分层等。尽管各行各业对于缺陷检测都有研究，但是针对 PCB 缺陷检测的深度学习神经网络还尚在研究中。

（三）两者结合的混合算法(Hybrid Method)^[31]。参照对比算法和无参照对比算法各有各的优势和劣势，混合法则是将这两者的优势共同结合形成的缺陷检测算法。由于 PCB 中缺陷的类型繁多，对于一个 PCB 来说，缺陷的数量也不少，有些缺陷适合采用参照对比算法，有些缺陷适合采用无参照对比算法，因此混合法就可以充分地发挥不同算法对于不同缺陷地检测效果，使检测结果更准确可靠。但是这种算法目前还不成熟，并且采取多种算法结合的方法计算量庞大、计算十分复杂，对于大规模 PCB 的缺陷检测适应性还不够。

将 PCB 缺陷自动视觉检测算法按照不同的检测原理进行分类，如图 1-3 所示：

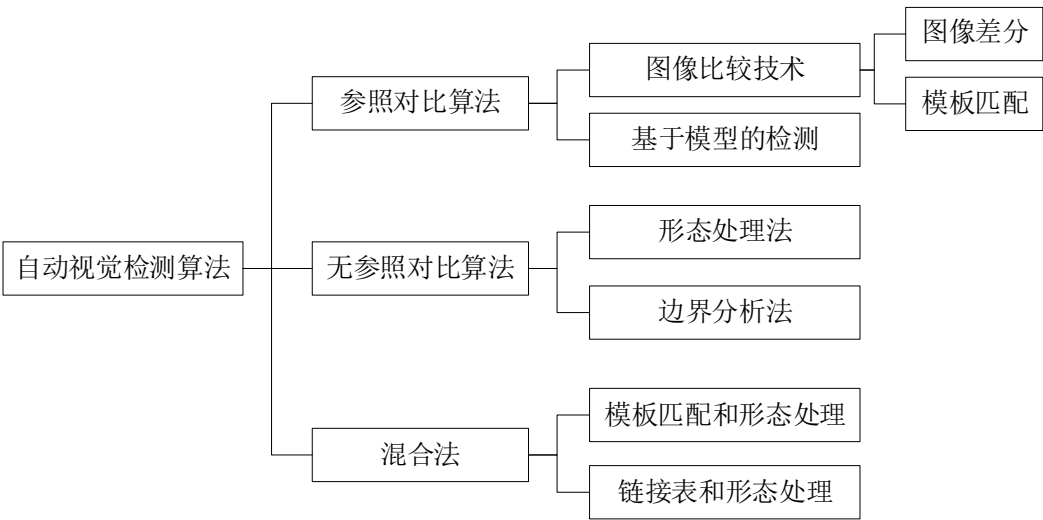


图 1-3 缺陷检测算法分类

1.3 论文的主要研究内容与章节

本文主要是针对 PCB 图像的缺陷检测问题，传统的 PCB 图像缺陷检测依靠人工进行视检或电气检测，但是检测效率、准确都很低，随后逐步研究出了基于不同特征提取方法的缺陷检测算法，但大都依赖图像处理技术进行特征提取，得到的结果很难令人满意。受近几年正飞速发展的深度学习在图像领域特征提取效果良好的启发，本文提出了基于图像处理和深度学习的 PCB 图像缺陷检测方法。

算法主要通过灰度处理、灰度直方图均衡化、高斯滤波等图像处理技术实现对 PCB 图像数据集的预处理，以消除噪声突出特征，并为后续神经网络训练提供数据集，再通过 SURF 图像配准、阈值分割、图像差分、形态学处理等机器视觉技术对 PCB 缺陷图像实现全局检测和初步定位。为了改善对图像的特征提取的质量，提出在 Resnet101 残差神经网络上采取多尺度特征融合和迁移训练的方式改进，得到了能够提取 PCB 图像良好特征的 Resnet101-finetune 网络结构。针对于目标检测问题，提出在 Faster-RCNN 目标检测算法上进行改进，与其他目标检测算法相比，在相同的 PCB 缺陷图像数据集上，表现出检测速度和准确率的明显提升。

针对于基于图像处理和深度学习的 PCB 图像缺陷检测的研究，本文的主要章节内容安排如下：

第一章：绪论。主要介绍了印刷电路板（PCB）生产厂商对于 PCB 进行缺陷检测这一需求由来已久，设计一套有效提高检测准确率的 PCB 图像缺陷检测算法具有十分重要的现实意义和研究价值，并分别介绍了 PCB 缺陷检测方法、图像处理和深度学习、PCB 缺陷检测算法各自的研究现状。

第二章：图像预处理。首先介绍了本文实验数据集的来源，并在数据集上进行预处理，通过灰度化、直方图均衡化、高斯滤波等操作实现了对图像进行平滑降噪去除干扰突出特征。并在预处理过的 PCB 图像数据集上通过 SURF 图像配准、阈值分割、图像差分、形态学处理等机器视觉方法实现对图像的全局缺陷检测和定位。

第三章：基于深度学习方法对 PCB 图像进行特征提取。传统的残差神经网络 Resnet101 有效解决了网络层数过深带来的梯度弥散问题，但是低网络层高分辨率对图像语义特征提取不足，高网络层低分辨率对图像细节特征提取不足，本文提出多尺度特征融合的方式改进网络，以提升特征提取的质量，形成了 Resnet101-finetune 网络。通过迁移学习的方式在 PCB 图像数据集上训练该网络并进行验证，实验结果表明改进后的 Resnet101-finetune 神经网络的平均准确率达到了 94.2%，相比于其他神经网络得到明显提升。

第四章：基于 Faster-RCNN 算法实现 PCB 图像的缺陷目标检测。通过改进 Faster-RCNN 算法提升目标检测的效率和准确度。在相同的 PCB 图像数据集上采用 YOLO、R-CNN、SSD 和本文提出的算法进行对比，实验结果表明本文提出的算法在平均准确率上达到 95.5%，mAP 值也均高于其他算法，验证了本文算法的有效性，并通过实例验证了本文算法的实用性，可以适用于 PCB 图像缺陷检测，具有较高的实际使用价值。

第五章：总结和展望。总结了本文主要的研究内容、成果以及实验的结论，并对本实验中包含的不足提出了后续改进的想法，以及对于本研究内容未来的一些发展方向提出新的期望。

第二章 基于图像处理的全局缺陷检测

PCB 缺陷检测的研究对象即为 PCB 裸板的图像数据，该图像数据是由工业相机拍摄所得，成像结果极易受到拍照时的光照条件、温度条件、空气条件的影响，此外传送带履带的抖动、拍摄相机的抖动都有可能导致 PCB 图像的质量，而 PCB 图像数据的质量直接影响了后续对于图像中的缺陷进行识别定位分类的准确程度，因此为了保证后续的深度学习算法的良好效果，需要对 PCB 图像数据进行图像预处理^[32]，这一步骤十分关键。

2.1 PCB 缺陷图像数据集

对于 PCB 的检测标准有很多，目前被业界广泛接受的是来自于美国电子电路和电子互连行业协会所制定的 IPC 标准（Institute for Printed Circuits），该标准对于 PCB 的生产过程各个阶段都建立了若干验收子标准，其中 IPC-A-600 标准^[33]则主要针对的是 PCB 裸板的表面外观的验收条件，即外部可观察特征，具体的缺陷分类如下表 2-1 所示：

表 2-1 IPC-A-600 缺陷类型

小尺寸缺陷	色差类缺陷	大尺寸缺陷	圆孔类缺陷	其他类型缺陷
针孔	字符缺陷	缺孔	偏孔	断路
毛发	铜板色差	大尺寸误腐蚀	缺孔	短路
导线间距过窄		多铜	针孔	鼠咬
		缺线	孔杂物覆盖	毛刺
		大尺寸遮蔽物		误腐蚀

本文将在 IPC-A-600 标准的基础上制定缺陷检测标准。PCB 图像数据集来自于北京大学智能机器人开放实验室免费公开提供的印刷电路板（PCB）瑕疵数据集 (<http://robotics.pkusz.edu.cn/resources/dataset/>)。该数据集共包含 1386 张具有多种不同缺陷类型的原始 PCB 图像，用于 PCB 图像的缺陷检测、分类和定位等问题的研究。本文从该数据集中选取了包含以下六种缺陷类型的原始 PCB 图像数据集作为实验数据集：缺孔（missing hole）、鼠咬（mouse bite）、断路（open circuit）、短路（short）、毛刺（spur）和多铜（spurious copper），如图 2-1 所示。

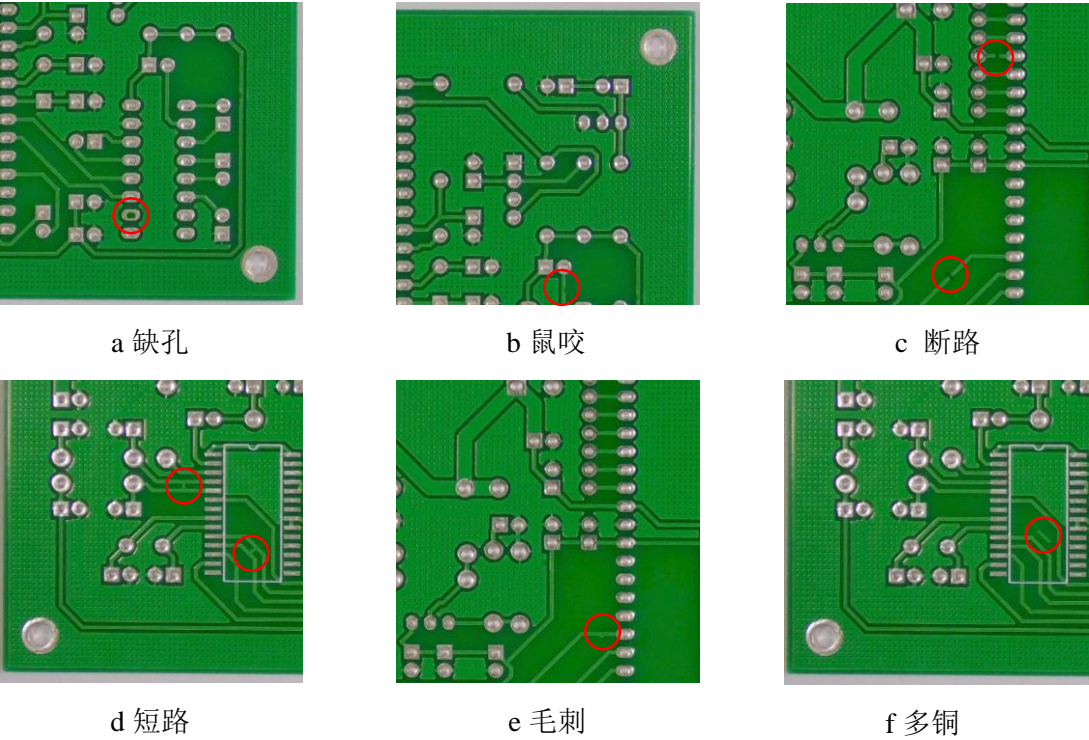


图 2-1 PCB 缺陷图像数据集

对原始 PCB 缺陷图像数据集进行数据增强^[34]，以增加模型训练时可用的图像数量，常用的数据增强的方式有旋转、镜像、亮度变换，随后将图像裁剪成合适训练使用的尺寸大小 600×600 ，最终得到 10668 张图像，本文数据集的具体情况如表 2-2 所示：

表 2-2 PCB 缺陷图像数据集

缺陷类型	图像数量	缺陷数量
缺孔	1832	3612
鼠咬	1852	3684
断路	1740	3548
短路	1732	3508
毛刺	1752	3636
多铜	1760	3676
共计	10668	21664

2.2 图像增强

图像增强指的是将图像中的与目标相关的关键信息突出出来，而对于与目标无关的信息尽可能去除的过程。对于在工业环境下采取工业相机拍摄得到的 PCB 缺陷图像，常常存在亮度过亮导致的某些细节曝光过度失真，或亮度过低导致细节信息丢失等问题，而噪声的存在也会影响到图像识别的过程。图像增强就是在不增加噪声的情况下，提高整体和局部之间的对比度，从而使图像中有用的信息更加凸显出来，以便于人眼

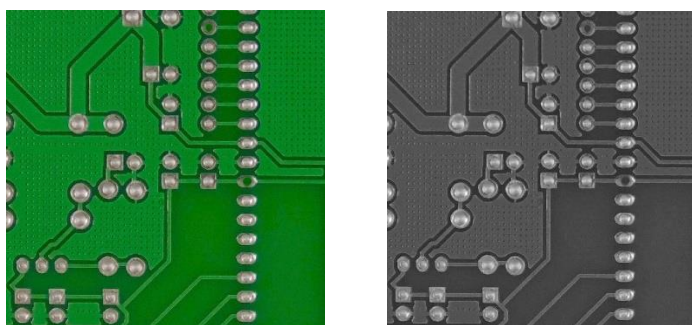
或机器进行观察识别。常用的图像增强的手段有图像灰度化、灰度直方图均衡化和降噪滤波。

2.2.1 图像灰度化

日常生活中对于彩色空间的描述常常采用 RGB 颜色空间，三维颜色 RGB 空间由三原色红绿蓝（RGB）构成，图像可以理解为由三原色按照不同的比例叠加而成。对于相机拍摄得到的图像进行数字离散化，则图像中的每一个像素的像素值都在[0, 225]的区间范围内，则图像就可以转变成为一个三维数组[H,W,C]，其中 H 代表图像的横向量，W 代表图像的列向量，C 代表图像的通道数量，在 RGB 颜色空间下即为 RGB 三个通道^[35]，当计算机对图像数据进行操作时，就需要计算三个通道的像素值，运算量太大，因此可以将三个通道的彩色图像转变成单通道的灰度图^[36]，参数计算量会大大减少，这便是图像灰度化，RGB 空间的灰度转换的公式如式 2-1 所示：

$$Gray(i, j) = 0.299 \times R(i, j) + 0.587 \times G(i, j) + 0.114 \times B(i, j) \quad (2-1)$$

其中 (i, j) 表示对应通道的像素矩阵。PCB 缺陷原图和灰度化处理之后的图像如图 2-2 所示：



(a) PCB 原始图像 (b) 灰度化图像

图 2-2 原始 PCB 图像与灰度化图像示例

经过灰度化处理之后，PCB 缺陷原始图像转换为单通道的灰色图像，失去了颜色信息，更加突出了前景与后景、整体与局部的亮度信息对比。从图 2-2 (b) 可以看出 PCB 灰度图像中的孔、线信息并没有损失，并且对于该图像中的缺陷信息“缺孔”更加清晰表现出来。采用灰度图像能够更加清晰地反映 PCB 图像中的缺陷信息。

2.2.2 灰度直方图均衡化

图像直方图^[37]反映了图像中不同亮度值的像素所占的比例，能够直观地看出图像的亮度分布，但是对于单通道的灰度图像来说，直方图则反映了该图像中不同灰度级出现的统计情况。图像的灰度直方图是一个一维的离散函数，如式 2-2 所示：

$$h(k) = n_k \quad k = 0, 1, \dots, 255 \quad (2-2)$$

其中， n_k 表示为图像的灰度值为 k 的像素点的个数，图像的灰度级有 256 种，因此 k 的值为 0~255。相应地，每个像素的灰度出现概率可以用式 2-3 表示为：

$$p(k) = \frac{n_k}{n} (0 \leq k \leq 255), \text{ 则有 } \sum_{k=0}^{255} p(k) = 1 \quad (2-3)$$

其中 $p(k)$ 表示图像像素中灰度值为 k 出现的概率。分别计算 PCB 原始图像和灰度图像的直方图，得到的结果如图 2-3 所示：

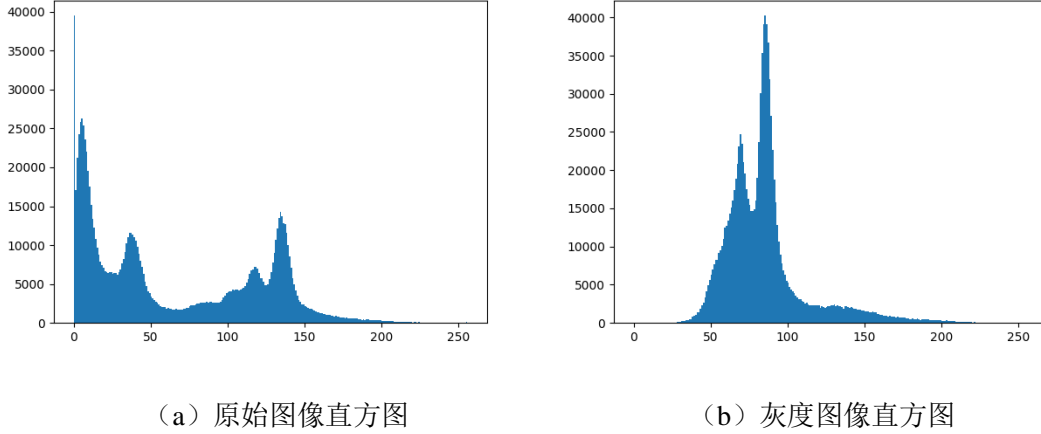


图 2-3 PCB 原始图像和灰度图像的直方图对比

其中，横坐标表示灰度值的范围是 0-255 灰度值，纵坐标表示该灰度值对应的像素的数量。对于原始 PCB 图像的直方图可以看出，图像的亮度分布很不均匀，而将原始图像处理过后的灰度图像的灰度值主要集中在[50, 100]之间，也就是灰度级的较低部分，图像的灰度值集中在较低部分时，图像可能会因为太过暗而使得部分细节信息难以观察，因此为了避免过暗或过亮，将图像中的目标缺陷信息细节清晰展示，需要对 PCB 的灰度图像进行直方图均衡化。

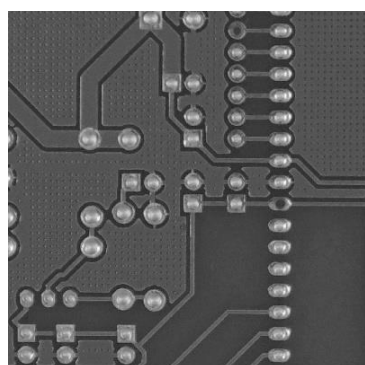
直方图均衡化^[38]是图像处理中一种常用于增强图像信息的技术，该技术的原理是对于图像中像素个数较多的灰度值进行展宽，而对于像素个数较少的灰度值进行归并，从而增强整幅图像的对比度，使图像清晰，达到增强的目的。所以该技术的关键就是建立一种灰度图像与灰度直方图均衡化图像之间的映射关系。假设输入图像为 T ，图像的宽度和高度的像素值分别设为 W 和 H ， $hist_{input}(k)$ 则表示该输入图像中像素的灰度值为 k 的像素点的数量。则计算灰度图直方图均衡化的公式如 2-4 所示：

$$\begin{aligned} hist_{input}(k) &\approx \frac{H \times W}{256} \\ \sum_k^m hist_{input}(k) &\approx (n+1) \frac{H \times W}{256} \end{aligned} \quad (2-4)$$

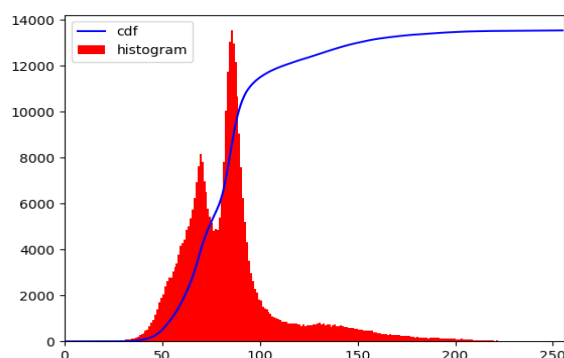
对于任意 $m \in [0, 255]$ ，总有 $n \in [0, 255]$ 使得上式成立，因此可以得到均衡化之后的灰度图像均衡化直方图 $Hist(T)$ ：

$$Hist(T) = \frac{\sum_k^l hist_{input}(k)}{H \times W} \times 256 - 1 \quad (2-5)$$

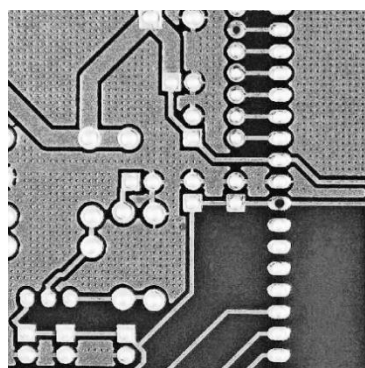
其中, I 表示像素的灰度值。根据该映射关系便可以得到灰度图像的均衡化结果。将 PCB 灰度图像进行灰度直方图均衡化处理, 均衡化前后的 PCB 灰度图像以及直方图如图 2-4 所示:



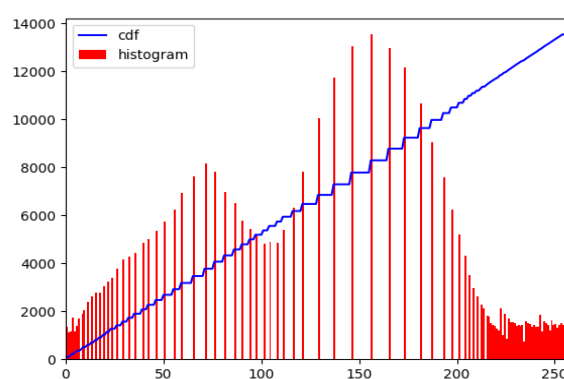
(a) 灰度均衡化之前



(b) 灰度均衡化之前直方图



(c) 灰度均衡化之后



(d) 灰度均衡化之后直方图

图 2-4 灰度均衡化前后的图像和直方图对比

如图 2-4 所示, 图 a 为灰度直方图均衡化之前的图像, 图像整体亮度较低, 像素的灰度值多集中在 $[50, 100]$ 之间, 图像内部的孔、线等边缘信息与图像背景的像素值差距较小。图 c 为灰度直方图均衡化之后的图像, 图像的整体亮度得到提升, 对于焊板上的孔、线等位置和形状信息更加突出, 且与背景的像素点区分明显, 比均衡化之前的灰度图像更能展现图像的细节信息。

图 b 和图 d 从直方图分布进一步表明了均衡化对图像的影响。横轴代表图像的灰度值, 纵轴代表该灰度值对应的像素点的数量。均衡化之前的灰度图像直方图如图 b 所示, 灰度值分布极端化, 大量像素点的灰度值都在 70 和 90 附近, 导致像素点之间的对比度很低, 表现在图像上就是轮廓等边缘信息模糊。图 d 为均衡化之后的灰度图像直方图, 将原先灰度值集中在 $[50, 100]$ 之间的像素点扩展到全部灰度值范围, 使不同像素点之间的对比度呈现明显差异, 方便了对于图像细节信息的观察。

2.2.3 降噪滤波

PCB 原始图像中不仅包含了我们需要的缺陷信息，还包含了外部原因导致在成像时产生的噪音，这些噪音的产生主要是在局部微小区域的像素点的像素值发生波动，这些噪音在后续对于图像的缺陷信息的判断有很大的负面影响，因此要想办法尽可能消除掉。为了消除或减少图像的噪音信息，同时不改变原有的特征信息，通常采用的方法是对图像进行降噪滤波，降噪滤波根据其原理可以分为空间滤波和频域滤波两大类，在空间上实现滤波的方式主要有均值滤波、中值滤波、高斯滤波和双边滤波，在频域上实现滤波的方式主要有小波变换和傅里叶变换等等。但是具体到针对 PCB 图像进行的降噪滤波，为了尽可能减少对图像内部特征信息的损耗，减少对图像预处理阶段的耗时，本文采用空间滤波的方式进行分析。

(1) 均值滤波

邻域均值滤波是在图像平滑处理算法中最常用的一种线性的、简单局部空间域降噪滤波算法^[39]。其滤波原理是设计一个可以滑动的窗口，根据滤波模板窗口在图像上滑动时覆盖的所有像素点的灰度平均值，作为中心像素的灰度值，实现图像的滤波。假设灰度图像的水平 and 垂直方向上的像素点的个数分别为 M 和 N ，则该灰度图像可以用下列二维矩阵来表示：

$$\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \cdots & \cdots & \cdots & \cdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}_{M \times N} \quad (2-6)$$

其中矩阵中的每一个元素代表该位置的像素点的灰度值，即点 (m,n) 的像素值即为 $f(m,n)$ 。对于图像中的点 $R_{m,n}$ ，该点处的噪音可以看作是外部影响造成的灰度值发生了小范围波动，则该点受到噪音干扰后的灰度值变为：

$$\hat{R}_{m,n} = R_{m,n} + N_{m,n} \quad (2-7)$$

其中， $N_{m,n}$ 是一个随机变量的噪音干扰，其针对图像中的每一个像素的均值为 0，方差为 σ^2 。那么均值滤波的基本思想就是采用一个像素值为 $(2x+1)(2y+1)$ 的窗口在灰度图像上遍历每一个像素点，并用窗口覆盖到的邻域内所有像素点的灰度平均值作为窗口中心位置的像素点的灰度值。那么受到噪音干扰的点 $\hat{R}_{m,n}$ 经过邻域平均滤波处理之后的像素灰度值变为了：

$$R_{m,n} = \frac{1}{(2x+1)(2y+1)} \sum_{j=-n}^n \sum_{i=-m}^m \hat{R}_{m-i,n-j} \quad (2-8)$$

由上式可以看出，受到噪音干扰的图像经过平均滤波处理之后的噪音的均值并不改变，而方差 σ^2 相比之前缩小了 $(2x+1)(2y+1)$ 倍。这种方法的本质是将某个像素点处的噪音干扰平均地分布在其周围邻域内地所有像素点，来削弱噪音对图像的影响，但同时也使图像本身的对比度降低，牺牲了图像的质量。采用 5×5 的窗口对 PCB 灰度图

像进行均值滤波得到的结果如图 2-6 (b) 所示, 由图中可以看出, 经过均值滤波处理之后图像整体变得模糊, 虽然对于噪声能够有效抑制, 但同时轮廓形状等关键边缘信息也受到损失。

(2) 中值滤波

针对于均值滤波只能解决线性平稳的图像的噪声干扰问题, 而获得的 PCB 缺陷图像是非平稳的, 并且均值滤波会损失图像的特征信息, 因此均值滤波方法并不适合, 而中值滤波的方法是非线性的, 其原理与均值滤波类似, 同样采用窗口在图像的像素点上逐一滑动, 但是对于窗口覆盖的邻域内的像素点的灰度值进行排序, 并将序列的中间值作为窗口中心的像素点的灰度值^[40]。由于中值滤波需要对像素灰度值进行排序并取中间值, 因此常采用奇数个像素点作为窗口, 例如常见的 3×3 、 5×5 的窗口形状, 如图 2-5 所示:

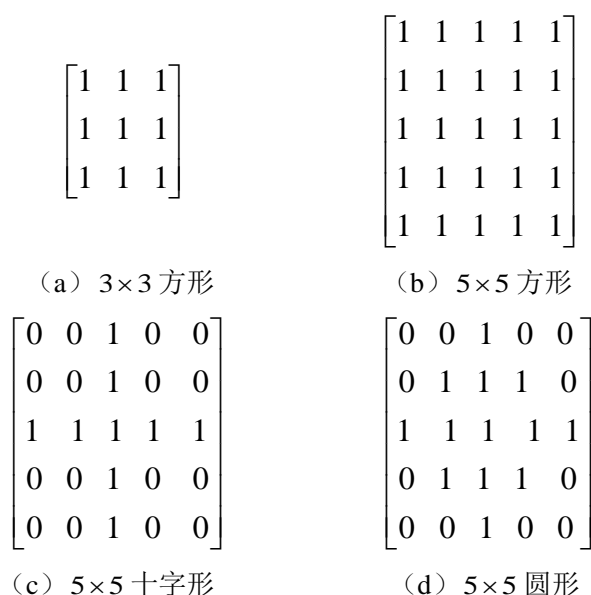


图 2-5 中值滤波窗口形状

在进行中值滤波时, 可采用上图的窗口对图像的像素点逐一计算, 窗口覆盖的像素点的权值为 1 时, 该像素点的灰度值保留, 否则舍弃, 随后对像素点的灰度值进行排序并将中间值作为结果赋予窗口中心的像素点。假设图像中位置为 (m, n) 的像素的灰度值为 $f(m, n)$, 则经过中心滤波处理之后该点的像素值为 $g(m, n)$, 则有:

$$g(m, n) = \text{Med}\{f(m, n)\} \quad (m, n) \in Z \quad (2-9)$$

其中, Z 表示以点 (m, n) 为中心的邻域模板窗口覆盖的像素点的集合, Med 表示对一个集合中的所有数值取中间值。对原始 PCB 灰度图像采取中间滤波处理之后的结果如图 2-6 (c) 所示, 由图中可以看出, 相比于均值滤波使得整个图像变得模糊, 中值滤波对于图像中的边缘信息并不改变, 图像的信号频谱基本保持不变, 同时, 对于图像内的孤立噪声点能够做到有效去除, 尤其是 PCB 图像中孔、线边缘在小范围内波动

产生的极大值或极小值噪声能够完全清除。但是中值滤波同样存在缺点，处理后对于边缘形状的改变仍具有局限性，并且不能消除高斯噪声。

(3) 高斯滤波

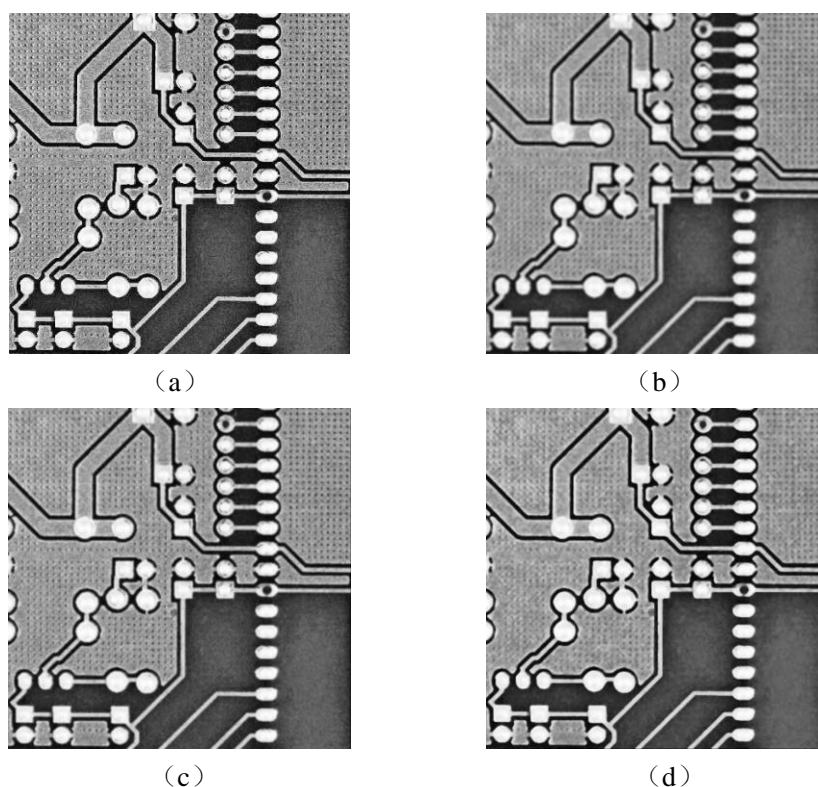
均值滤波认为图像的特征信息和噪声信息同等重要，因此在降噪的同时模糊了图像本身特征。中值滤波则对图像的特征信息和噪声信息一起排序，将中间值作为结果，同样没有考虑两者的权重。而在降噪滤波的过程中，我们希望尽可能保留图像的本身特征信息，而将噪声干扰尽可能消除，因此需要对图像和噪声采取不同的权值，这就是高斯滤波。高斯滤波^[41]是基于高斯函数，对于窗口内覆盖的邻域像素点的灰度值赋予不同的权值进行线性处理的平滑滤波器，其认为距离中心像素点越近则权重越大，反之则越小，在邻域内的变换范围呈现出离散高斯函数的形式。常用的二维高斯滤波器如式(2-10)所示：

$$\begin{aligned} R_{\sigma}(m,n) &= \frac{1}{2\pi\sigma^2} e^{-\frac{(c-\frac{W-1}{2})^2+(r-\frac{H-1}{2})^2}{2\sigma^2}} \\ &= \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(c-\frac{W-1}{2})^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(r-\frac{H-1}{2})^2}{2\sigma^2}} \\ &= R_{\sigma}(W)R_{\sigma}(H) \end{aligned} \quad (2-10)$$

其中， w 为高斯矩阵的宽， H 为高斯矩阵的高， c 的取值范围是 $[0, H-1]$ ， r 的取值范围是 $[0, W-1]$ 。因此高斯滤波的降噪效果主要与 σ 的取值相关， σ 的取值范围为 $[0,1]$ ， σ 的取值越小，则认为离窗口中心的像素点越远的像素点的权重越小，其对邻域内的影响作用越小， σ 的取值越大则认为离窗口中心的像素点越远的像素点的权重越大，其对邻域内的影响作用越大，当 σ 为 1 时，即为均值滤波。

高斯滤波器将中心像素的邻域内的像素灰度值的加权平均数代替中心像素的灰度值，而邻域内的每一个像素点的权值随着与中心点的距离增加而减小，大大降低了滤波对于图像的模糊效果，保留边缘锐利程度的同时，减小噪声干扰。此外二维的高斯函数具有旋转对称性，因此在对图像的像素点进行映射时，高斯函数认为该像素点在不同方向上的平滑程度都是均等的，因此平滑不会朝任何一个方向偏离，不会对原始图像带来新的干扰。

经过实验对比，当窗口采取 5×5 方形作为模板， σ 取 0.8 时，对于图像的降噪效果最为合适。采取 5×5 方形窗口， σ 取 0.8 对原始 PCB 灰度图像进行高斯滤波处理，得到的结果如图 2-6 (d) 所示。从图中可以看出，经过高斯滤波之后，图像仍保留较大程度的清晰度，图像中的孔、线的位置形状等信息并没有损失，PCB 表面多数特征信息均得到保留，并且对于背景板等无关信息进行了平滑处理，模糊了噪声信息，极大突出了图像特征。



(a) 原始灰度图；(b) 均值滤波处理后；(c) 中值滤波处理后；(d) 高斯滤波处理后

图 2-6 不同降噪滤波方法的结果对比

采取不同的降噪滤波方法对 PCB 灰度图像进行处理，得到的结果如图 2-6 所示。图中 a 表示原始的 PCB 灰度图像，其将图像中的所有信息均表现出来，既包括需要关注的孔、线的位置形状等特征信息，也包括了背景板或噪声干扰等无关信息。图 b 表示采取均值滤波处理后的灰度图像，将特征与噪声同时平滑，消去噪声的同时也降低了图像的特征，使图像整体失真，难以用于后续检测。图 c 表示采取中值滤波处理后的灰度图像，极大减少了边缘信息在小范围内波动产生的极大值或极小值的噪声干扰，但是对于边缘形状会产生变化，同时难以消去高斯噪声。图 d 表示采取高斯滤波处理后的灰度图像，其在最大程度范围内保留了边缘形状和锐利程度，对于图像的特征信息充分保留，并同时消去背景板和高斯噪声等干扰，达到了图像降噪滤波的目的。

综合上述分析，本实验采用高斯滤波的方法对图像进行平滑降噪，滤波器采用 5×5 方形窗口， σ 的值取为 0.8。

2.3 缺陷全局检测和定位

对 PCB 缺陷图像经过灰度化、灰度直方图均衡化、高斯滤波等处理之后，对于图像中的特征信息已经可以较为明显得表现出来了，因此接下来便可以通过数字图像处理技术先全局检测一下待测 PCB 图像是否包含缺陷，以及对缺陷进行定位。基于图像差分的缺陷检测方法属于参照对比算法，因此需要分别对 PCB 标准图像和 PCB 缺陷图像进行图像预处理，提取各自的特征并进行特征匹配，然后再进行图像差分，如果差分得到的结果在一定阈值内，则可认为待测 PCB 图像不存在缺陷，反之则存在，并对差分得到的结果进行形态学处理，使得缺陷信息更加清晰准确表现出来。具体的流程如图 2-7 所示：

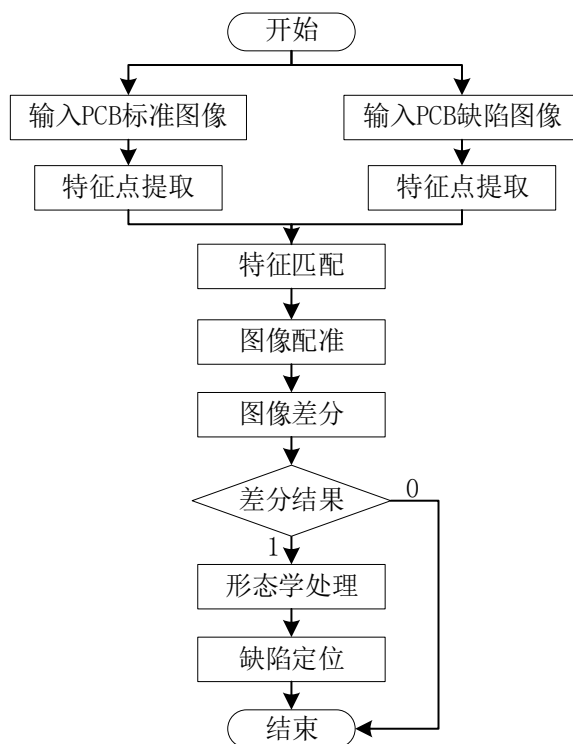


图 2-7 缺陷全局检测和定位流程图

2.3.1 图像配准

在进行图像差分来判断待测图像是否存在缺陷时，首先要做的便是将待测图像和标准图像进行配准。工业相机拍摄得到的 PCB 图像往往会因为拍照角度、传送速度等各种各样外界因素导致成像时发生偏转，如果直接拿角度偏转的待测图像与标准图像进行差分，得到的结果很难判断具体是否有缺陷信息。因此将待测图像与目标图像进行配准是必不可少的一步^[42]，常见的图像配准方法有基于模板匹配的图像配准、基于灰度的图像配准、基于模型的图像配准、基于特征的图像配准等等。因为上文已经对 PCB 图像数据进行了预处理，得到的灰度图像很好得表征了 PCB 图像的特征信息，因此这里采用基于灰度特征的图像配准方法，具体流程如图 2-8 所示。

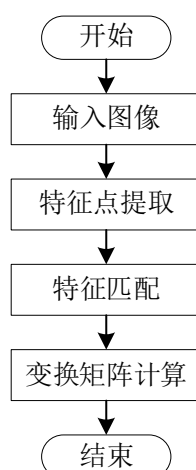


图 2-8 基于灰度特征的图像配准方法

如图 2-8 所示，首先输入经过高斯滤波之后的灰度直方图均衡化得到的 PCB 标准图像和 PCB 缺陷图像，特征点主要集中在灰度图像的角点边缘，分别对两者进行特征提取，并将特征点进行匹配，形成匹配特征点之后计算变换矩阵以实现图像配准。图像配准的结果主要取决于图像特征点的提取，提取出来的特征信息越全面越能表征图像的特征信息则配准的结果越好，因此采用合适的特征提取方法至关重要。目前常用的基于提取图像特征的配准算法主要有 SIFT、Harris、SURF。

(1) SIFT (Scale-invariant feature transform, 尺度不变性特征变换) 是由 Lowe 等人在 1999 年提出，至 2004 年得到完善的一种针对图像匹配的特征检测算法^[43]，它通过检测尺度空间的极值点来精确定位特征点并生成特征描述子。这种方法对于尺度、旋转以及一定视角和光照变化等都具有不变性。

(2) Harris^[44]是 1988 年由 Harris 和 Plessey 等人在 Moravec 方法的基础上改进提出的角点检测算法并命名。这种方法计算方式简单，提取出的特征点均匀稳定，具有很好的表征性，对于图像的旋转、亮度变换等都不敏感。

(3) SURF^[45] (Speeded Up Robust Features, 加速鲁棒性特征) 是由 Bay 等人在 2006 年对 SIFT 算法进行改进，提高了他的效率形成的特征检测算法。

分别使用上述三种方法对 PCB 缺陷图像进行特征点提取得到的结果如图 2-9 所示：

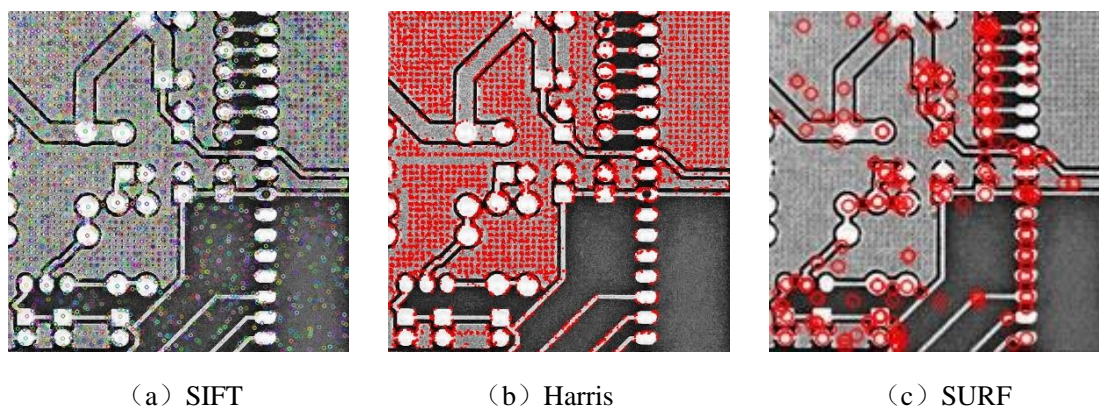


图 2-9 不同方法特征提取对比

由图 2-9 可以看出，图 a、b、c 分别为经过 SIFT、Harris、SURF 特征检测方法进行特征提取的结果。由图 a 可以看出，SIFT 能够精确定位全局所有的特征点并进行描述，但是针对于 PCB 图像配准，我们更关心的是孔、线等边缘位置形状特征，而对于背景板以及与关键特征无关的其他特征并不关心，全局所有的特征点都进行描述配对不仅不会提高准确率还会使效率降低，甚至因为引入过多特征点而导致配准效果误差增大。同样地对于图 b，Harris 角点检测得到的特征点信息非常均匀并且稳定，但是过多的无关信息特征引入是不必要的。此外对于图 a、b 来说，这张包含了“缺孔”缺陷的 PCB 图像，这两种方法在缺陷位置的特征点描述并不充分。

而对于图 c 可以看出，采取 SURF 特征检测方法提取出的特征点，非常贴合地表示了图像中所有的孔、线位置信息，并且对于我们特别需要关注的缺陷信息，在该处附近得到很多特征点，这对于后续的缺陷检测十分有利。因此我们采取 SURF 特征检测算法来实现 PCB 图像配准。

使用 SURF 特征检测算法分别对 PCB 标准图像和 PCB 缺陷图像进行特征点提取，并对特征提取结果进行匹配，得到的结果如图 2-10 和图 2-11 所示：

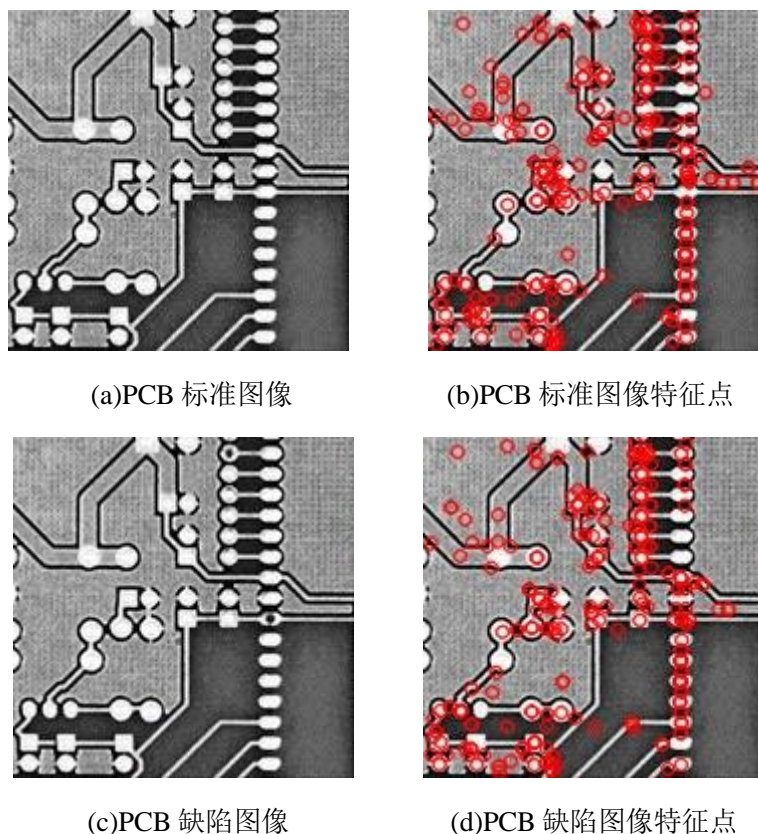


图 2-10 PCB 标准图像与缺陷图像特征提取对比

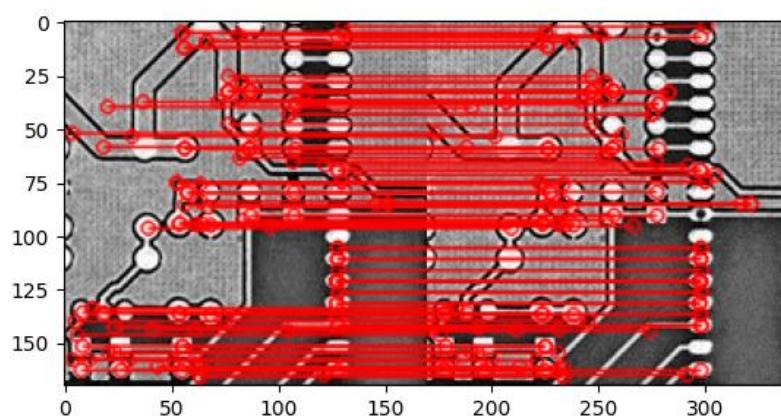


图 2-11 PCB 标准图像和缺陷图像采用 SURF 进行配准

由图 2-10 可以看出，图 a 和图 c 分别为 PCB 标准图像和待检测的 PCB 缺陷图像，图 b 和图 d 分别为两者的特征提取结果，可以看出相对于标准图像，待检测的 PCB 缺陷图像无论是在原图或是特征提取结果上都能明显看出具有两处“缺孔”。由图 2-11 可以看出，SURF 算法能够明确分辨出标准图像和待检测图像之间的不同之处，即缺陷信息，而对于相同的地方，也能够准确提取各自的特征点并进行特征匹配。

分别采用 SIFT、Harris、SURF 算法对 PCB 标准图像和缺陷图像进行配准，得到的相关数据如下表 2-2 所示：

表 2-2 不同图像配准算法对比

	标准图 特征点数	缺陷图 特征点数	正确匹 配对数	总匹配 对数	准确率	匹配率	匹配时间 (ms)
SIFT	7066	6295	1880	1880	1.00	0.26	6470
Harris	5849	5125	1130	1228	0.92	0.21	1580
SURF	4535	4532	1677	1677	1.00	0.37	6034

由表 2-2 可以看出，SIFT 特征检测算法能够提取出的图像特征点的数量是最多的，并且对于成功匹配的特征对都是正确的，说明基于 SIFT 的特征提取能够准确提取表征图像信息的特征点，但是他的匹配率不高，匹配时间最长，说明其提取特征不具备针对性，提取全局特征中关键特征信息占比较少，耗时更久。Harris 特征检测算法虽然耗时最短，但是其特征点匹配的准确率和匹配率都最低，说明其寻找特征点快速，但表征能力较差。而对于 SURF 特征检测算法，其耗时介于上述两者之间，准确率和 SIFT 算法一致，且匹配率得到大幅提升，说明其性能是这三者之中最为稳定且特征表现良好的。综合特征提取的质量以及算法的稳定性和效率，针对于 PCB 图像的配准，本文选择 SURF 特征检测算法进行图像配准。

2.3.2 图像差分

经过配准之后的 PCB 灰度图像如果直接进行异或操作，图像之间的像素很难一一对应，会产生很多差异，影响到后续的过程，因此在做差分处理之前，首先需要对配准过的图像进行阈值分割。常见的缺陷类型多存在于 PCB 上的孔、线等位置，通过阈值分割可以很好地将这些线路与焊板之间分割出来。其原理主要是基于灰度图像中的像素点的灰度值都是处于 $[0,255]$ 之间的离散值，按照设定的阈值对图像进行二值化处理，生成二值图。最常见的阈值分割方法是 OTSU 阈值分割^[46]，是一种自适应阈值确定方法。假设图像的灰度等级有 $[1,2,\dots,L]$ ，其中灰度值为 l 的像素的个数为 n_l ，则灰度值为 l 的归一化概率 p_l 可以通过式 2-11 计算：

$$p_l = \frac{n_l}{n_1 + n_2 + \dots + n_L} = \frac{n_l}{N} \quad p_l \geq 0 \text{ 且 } \sum_{l=1}^L p_l = 1 \quad (2-11)$$

其中 N 表示像素总数。假设图像的灰度值可以被 k 分为 C_0 和 C_1 两类，则 C_0 的灰度范围是 $[0,1,\dots,k]$ ， C_1 的灰度范围是 $[k+1,k+2,\dots,L]$ ，则这两类的概率分别为：

$$\begin{cases} \omega_0 = \omega(k) = \sum_{l=1}^k p_l \\ \omega_1 = 1 - \omega(k) = \sum_{l=k+1}^L p_l \end{cases} \quad (2-12)$$

则每一类的均值可以通过式 2-13 计算：

$$\begin{cases} \mu_0 = \frac{\sum_{l=1}^k l p_l}{\omega_0} \\ \mu_1 = \frac{\sum_{l=k+1}^L l p_l}{\omega_1} \end{cases} \quad (2-13)$$

则两类的方差为：

$$\begin{cases} \sigma_0^2 = \frac{\sum_{l=1}^k (l - \mu_0)^2 p_l}{\omega_0} \\ \sigma_1^2 = \frac{\sum_{l=k+1}^L (l - \mu_1)^2 p_l}{\omega_1} \end{cases} \quad (2-14)$$

则类间方差 σ_B^2 计算方式如式 2-15 所示：

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_r)^2 + \omega_1(\mu_1 - \mu_r)^2 = \omega_0 \omega_1 (\mu_0 - \mu_1)^2 \quad (2-15)$$

其中, μ_r 为整体均值。图像类间方差 σ_b^2 表征了图像中这两个类之间的差异程度。阈值分割的目的是将线路信息与背板分隔开, 因此图像内这两类的差距尽可能大, 则方差 σ_b^2 也要尽量大。OTSU 算法的目标就是找到合适的阈值使得方差取得最大值。

依靠 OTSU 阈值分割算法设定的固定阈值很难适应不同的数据集, 尤其是 PCB 缺陷图像数据集的布线复杂, 孔线繁多, 因此也可以采取手动设定阈值的方法。假设 (m,n) 为图像宽、高分别为 M 、 N 的一个像素点, 手动设定阈值假设为 t (threshold), 则经过阈值分割后的像素点的灰度值 $f_{output}(m,n)$ 可以按照式 2-16 计算:

$$f_{output}(m,n) = \begin{cases} 255, & f_{input}(m,n) < t \\ 0, & f_{input}(m,n) \geq t \end{cases} \quad (2-16)$$

经过阈值分割得到的 PCB 标准图像和缺陷图像的二值图就可以进行图像差分了。假设两张图像分别为 F_1 和 F_2 , 对应的像素点分别为 $F_1(m,n)$ 和 $F_2(m,n)$, 则两者进行差分操作之后得到的结果便是:

$$F_{out}(m,n) = |F_1(m,n) - F_2(m,n)| \quad (2-17)$$

由图 2-4 (b) 可知, 对 PCB 缺陷图像进行灰度直方图统计之后, 其灰度值主要集中在 [50, 100] 之间, 因此这里手动阈值设定可以在该区间中选取, 此处选择了 55、65、75、85、95 阈值进行比较分析。分别对图像按照 55、65、75、85、95、OTSU 阈值进行阈值分割, 将分割后的图像进行差分^[47], 得到的结果如图 2-12 所示:

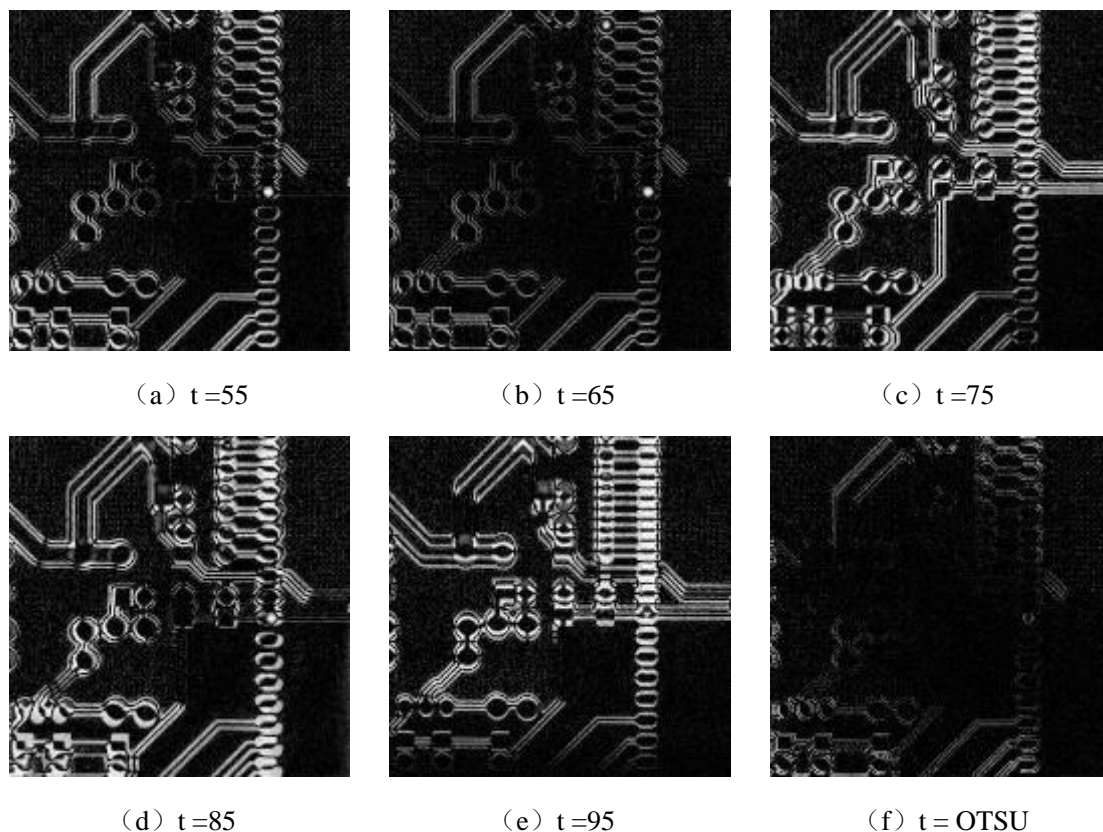


图 2-12 不同阈值下差分结果对比

从图 2-12 中可以看出, 手动设定阈值时, 阈值过大或过小都会使得差分的结果较差, 例如当阈值取 85 和 95 时的图 d 和图 t, 与缺陷无关的布线特征信息没有通过差分操作去除掉, 而缺陷的特征信息也没有被明显暴露出来。而其中当阈值选取为 65 时, 则可以明显观察到两处缺孔缺陷, 并且布线等正确信息相比于其他阈值也减弱很多, 对于无法完全消去的正确布线特征可以通过后续的形态学处理进一步操作。而对于阈值选取采用 OSTU 算法得到的差分结果如图 f 所示, 该图虽然实现了差分操作并将正确信息消去, 但从图中已经很难观察到缺陷特征了, 说明该阈值分割方法不能将缺陷特征与正确特征分开。因此, 阈值采取手动设定为 65 得到的结果最为合适。

2.3.3 形态学处理

经过阈值分割和差分处理之后的结果虽然也已经可以明显检测到缺陷的位置, 但是图像中仍然包含很多细小分散的噪声点以及没有消去的正确特征, 这些信息对于后续的缺陷检测和分类都会带来干扰, 甚至会被误认为是缺陷, 因此需要采取形态学运算^[48]的方法除去干扰信息。

常用的形态学处理方法有腐蚀、膨胀、开/闭运算、顶/底帽运算等等。针对于差分图像旨在削弱边缘线信息, 突出亮孔信息, 拟采用开、闭运算^[49]对图像进行处理。

腐蚀是指采用与滤波处理中的窗口类似的结构元, 将结构元在图像上滑动过每一个像素的点, 结构元覆盖的所有邻域内的像素点的灰度值的最小值作为结构元中心的像素点的灰度值^[50], 对于图像 T 来说, 假设结构元为 S , 则经过腐蚀操作后得到的图像 E 为:

$$E = T \odot S \quad (2-18)$$

显然经过腐蚀操作之后部分像素点的灰度值降低, 图像整体的亮度降低, 小面积区域的像素值的灰度值会降低至 0, 从而实现降噪的效果, 而对于亮孔部分灰度值则保持不变, 图像对比度增加。

相应地, 膨胀操作则是在结构元覆盖的所有邻域内的像素点的灰度值的最大值作为结构元中心的像素点的灰度值, 则经过膨胀操作后得到的图像 D 为:

$$D = T \oplus S \quad (2-19)$$

而开运算和闭运算就是在图像 T 上, 用结构元 S 先后进行腐蚀膨胀的操作:

$$\begin{aligned} T \circ S &= (T \odot S) \oplus S \\ T \bullet S &= (T \oplus S) \odot S \end{aligned} \quad (2-20)$$

即开运算就是用结构元在图像上先进行腐蚀后进行膨胀的操作, 腐蚀降低边缘线的灰度值, 膨胀提高亮孔缺陷的灰度值, 因此开运算可以有效过滤差分图中的无关信息, 突出缺陷信息。分别用开运算和闭运算对差分结果进行图像形态学处理, 得到的结果如图 2-13 所示:

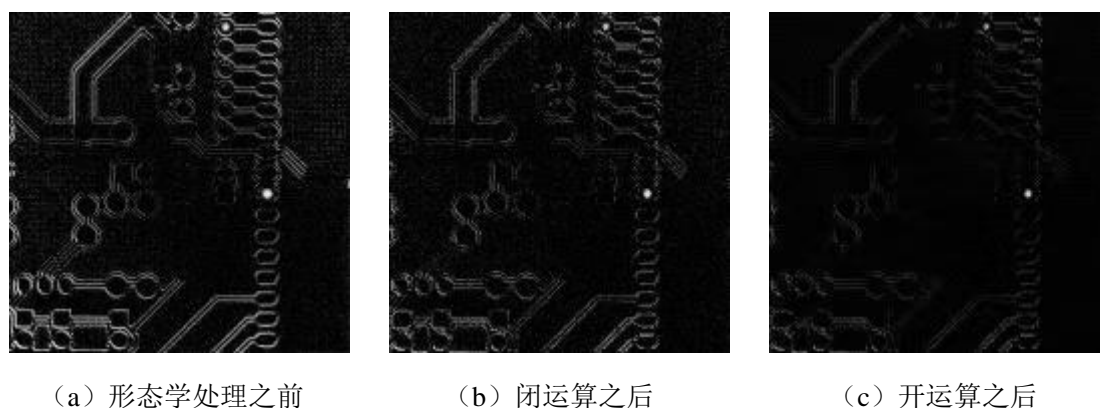


图 2-13 形态学处理前后结果对比

由图 2-13 可以看出，形态学处理之前的图 a 包含许多没有消去的边缘信息，经过闭运算之后得到的图像 b 已经消除了大部分但仍有残留，而对于开运算处理之后的图像 c 基本已经观察不到无关信息了，两个“缺孔”缺陷非常明显得表示了出来。

至此，已经通过图像处理方法实现了 PCB 缺陷图像的全局检测和定位。接下来便可以通过深度学习神经网络对图像中的缺陷信息进行识别分类，实现 PCB 图像的缺陷检测。

2.4 本章小结

本章主要介绍了本实验的数据集来源于北京大学智能机器人开放实验室免费公开提供的印刷电路板（PCB）瑕疵数据集，首先对 PCB 图像数据进行图像增强以加强图像的特征表达，图像增强的措施包括灰度化、灰度直方图均衡化和高斯滤波，从而使图像的缺陷信息更加清晰表现出来，并为后续的深度学习神经网络训练提供训练数据集。然后采用了 SURF 特征检测算法对缺陷图像和标准图像进行配准，随后手动设定阈值为 65 进行阈值分割并对两者进行差分，得到仅包含缺陷信息的差分图像，再利用形态学处理中的开运算过滤图像中的无关信息，最终实现了 PCB 缺陷图像全局检测和定位。将定位得到的缺陷信息映射到原始 PCB 图像中的对应位置，并分割出最小外接矩形区域，对该区域的缺陷类型进行分类标注，即可为后续的神经网络训练提供缺陷图像数据集。

本章主要基于参照对比算法对图像进行缺陷检测和定位，下一章则是基于无参照对比算法使用卷积神经网络对 PCB 图像的缺陷进行学习训练，实现缺陷检测和分类。

第三章 基于卷积神经网络的 PCB 缺陷检测

基于无参照对比算法的图像处理技术实现了对 PCB 缺陷图像的全局检测和初步定位,但是检测过程复杂,检测效率较低,检测准确度也受到图像处理技术的制约,并且对于缺陷的类别也无法确定。近几年正处于火热发展中的深度学习卷积神经网络技术在图像识别的领域得到大力发展,也为 PCB 缺陷图像检测这一问题带来了新的解决方案。通过将 PCB 缺陷图像形成训练集训练卷积神经网络,使训练好的模型能够准确检测 PCB 图像的缺陷并识别成为可能^[51]。

以往的人工神经网络和如今普遍使用的卷积神经网络之间最主要的差别在于深度以及学习能力,比如最开始的人工神经网络仅包含两到三层的网络层,实现简单的特征提取学习功能,但如今的卷积神经网络已经可以发展到成百上千的网络层,更深的网络层次实现了提取出的特征更加精确,算法检索的准确度和速度也得到了大大提升,例如残差神经网络已经能够达到了上千层,具备很强的计算机学习能力。本章在残差神经网络的基础上进行改进,实现了对 PCB 图像的缺陷检测,并提升了准确率。

3.1 人工神经网络

人工神经网络^[52]取名于人类的神经网络系统,生物的神经网络是以很多神经元为基本单元组合而成的,每个简单的神经元之间通过树突和轴突连接而成,也靠这些突触连接实现信息的传递。科学家以此为灵感设计一个类似于生物神经网络的拓扑结构,能够模拟出生物神经网络对外界刺激进行感知并且反馈,这个结构便被命名为人工神经网络,该网络有多个网络层之间的很多个神经元相同连接而成,通过这些神经元接收输入信息,每一层根据各自的功能以及需求进行操作,将输出作为下一层的输入渐渐传输下去,逐层处理之后并输出最终结果。

3.1.1 感知器与多层感知器

神经网络中第一步便是获取外界或是上一个神经元传递而来的信息,这依赖于感知器进行,每一个感知器都像人类的一个神经元,能够接收单个或多个的输入,并产生一个结果,采取合适的阈值函数作为激活函数,最终输出一个结果。如图 3-1 所示:

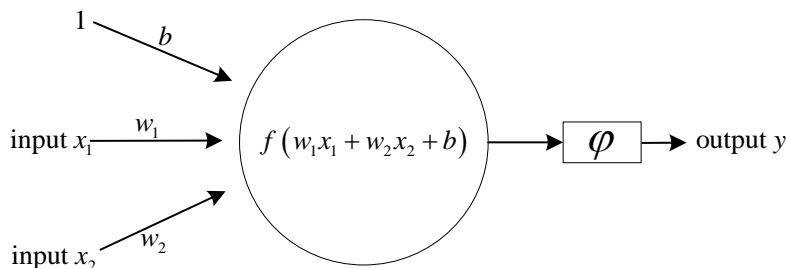


图 3-1 单层感知器

由图中可以看出，这个单层感知器模型有两个变量 x_1 和 x_2 输入，其权重分别为 w_1 和 w_2 ，还有一个常量也就是截距项输入，其偏置（bias）为 b ，感知器将每一个输入分量与其对应的权值进行相乘并求和，进行激活函数之后得到输出。

单层的感知器仅仅能够在线性可分的数据集上起到良好作用，快速达到收敛并且权重值的分类情况较好，但是面对线性不可分的数据集上，仅仅依靠单层的感知器很难得到与实际分类情况一致的权重值，权重值经过最大迭代次数也很难收敛，此时便需要将多个单层的感知器结合起来，形成多层感知器（Multi-Layer Perceptron, MLP），共同进行权值迭代。

多层感知器包含有输入层、隐层和输出层三个部分，其中输入层相比较于单层感知器来说，不仅仅可以接收连续型数据，还可以接受离散型数据。输出层与单层感知器一致，经过激活函数后得到结果输出。除了输入层和输出层，还有很多中间的网络层，这些网络层并不与感知器以外的信息结构连接，因此被称为隐层，具体的多层感知器结构如图 3-2 所示：

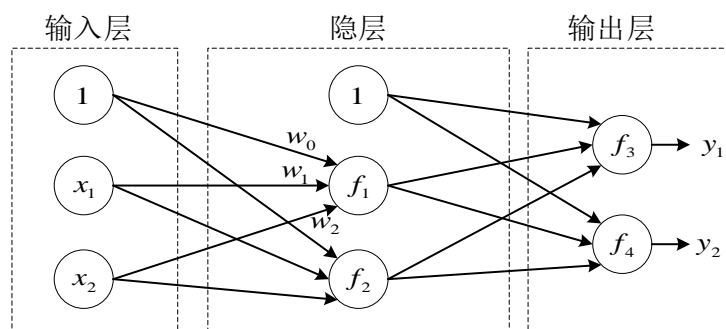


图 3-2 多层感知器结构

多层感知器的隐层根据需要可能有很多层，上图结构仅展示了包含一层隐层的多层感知器。并且上图仅标注了输入层 1、 x_1 、 x_2 到隐层中 f_1 的权重值 w_0 、 w_1 、 w_2 ，实际上每一个连接都有各自的权重值。输入层包含有两个外界变量和一个常量，按照其各自权重值相乘并求和，将计算得到的结果传输到隐层中，然后在隐层的多层网络中，分别按照输入值与对应附加的权重值进行相同的相乘并求和的计算，逐层传递下去，最终传递到输出层时，经过激活函数后得到的结果便作为输出传递出去。

3.1.2 反向传播算法

由多层感知器的训练过程可以得知，训练的本质就是对权值数据的更新，以求得最符合实际分类得权值。在感知器训练的一开始，所有的边权重（edge weight）都是在初始化阶段随机选取的，随着训练数据的不断输入，神经网络开始工作，按照初次的权值进行计算得出的结果，由人为监督的与实际分类进行比对，并将结果误差传播回前一层，然后适当的调整权值，循环往复多次之后，使得按照权值的预测结果与实际结果之间的误差趋于允许范围内，这个将误差由后层传向前层的过程便被称为反向传播算法^[53]（Backpropagation Algorithm, BP 算法），也是感知器训练过程实现“学习”能力的关键。

在神经网络的训练过程中，最重要的便是反向传播算法，他的核心是链式求导法则。反向传播算法 BP 算法主要包括两个部分：（1）前向传播。将输入的数据值按照预先设置的权值进行计算，逐层的由前向后传播，直至输出层，计算得出每个神经元的激励值和损失值。（2）反向传播。将感知器的预测结果与实际结果或期望结果进行对比，将预测值与真实值之间进行计算，计算得到的误差，逐层由后向前传播，根据误差的值调整各个神经元之间连接的权值。

反向传播算法计算过程如下：

假设神经网络总共有 L 层，损失函数为 J ，第 l 层的第 j 个神经元的输入值用 $z_j^{(l)}$ 表示，计算方式如下：

$$z_j^{(l)} = \sum_k w_{jk}^{(l)} a_k^{(l-1)} + b_j^{(l)} \quad (3-1)$$

其中， $w_{jk}^{(l)}$ 表示第 $(l-1)$ 层的第 k 个神经元与第 l 层的第 j 个神经元之间连接的权重值向量， $b_j^{(l)}$ 表示第 l 层的第 j 个神经元的偏置项。则第 l 层的第 j 个神经元的激励输出结果 $a_j^{(l)}$ 的计算方式如下：

$$a_j^{(l)} = f \left(\sum_k w_{jk}^{(l)} a_k^{(l-1)} + b_j^{(l)} \right) \quad (3-2)$$

其中， f 为激励函数。由此可以得出，最后一层的预测的输出值为 $a^{(L)}$ 。则第 l 层的第 j 个神经元的预测值与真实值之间的误差，也就是残差，用 $\delta_j^{(l)}$ 来表示，其计算方式如下：

$$\begin{aligned} \delta_j^{(l)} &= \frac{\partial J}{\partial z_j^{(l)}} = \sum_k \frac{\partial J}{\partial z_k^{(l+1)}} \cdot \frac{\partial z_k^{(l+1)}}{\partial a_j^{(l)}} \cdot \frac{\partial a_j^{(l)}}{\partial z_j^{(l)}} \\ &= \sum_k \delta_k^{(l+1)} \cdot \partial \left(w_{kj}^{(l+1)} a_j^{(l)} + b_k^{(l+1)} \right) \cdot f' \left(z_j^{(l)} \right) \\ &= \sum_k \delta_k^{(l+1)} \cdot w_{kj}^{(l+1)} \cdot f' \left(z_j^{(l)} \right) \end{aligned} \quad (3-3)$$

计算权重 $w_{kj}^{(l)}$ 的偏导数：

$$\frac{\partial J}{\partial w_{jk}^{(l)}} = \frac{\partial J}{\partial z_j^{(l)}} \cdot \frac{\partial z_j^{(l)}}{\partial w_{jk}^{(l)}} = \delta_j^{(l)} \cdot \frac{\partial \left(w_{jk}^{(l)} a_k^{(l-1)} + b_j^{(l)} \right)}{\partial w_{jk}^{(l)}} = a_k^{(l-1)} \delta_j^{(l)} \quad (3-4)$$

计算偏置项 $b_j^{(l)}$ 的偏导数：

$$\frac{\partial J}{\partial b_j^{(l)}} = \frac{\partial J}{\partial z_j^{(l)}} \cdot \frac{\partial z_j^{(l)}}{\partial b_j^{(l)}} = \delta_j^{(l)} \cdot \frac{\partial \left(w_{jk}^{(l)} a_k^{(l-1)} + b_j^{(l)} \right)}{\partial b_j^{(l)}} = \delta_j^{(l)} \quad (3-5)$$

根据以上两个偏导数的计算公式便可以计算得到经过误差修正后的权值以及偏置项的数值。

3.1.3 卷积神经网络

随着研究持续深入，神经网络开始逐渐加深用来拟合更加复杂的函数，渐渐便发展出了深度神经网络（Deep Neural Networks, DNN），深度神经网络 DNN 的结构里，每一层的每一个神经元和上层的所有神经元之间都能够形成连接，这就导致了训练过程中产生的参数的数量爆炸性增长，过多的权值需要训练直接使得整个网络十分容易过拟合，此外在图像处理领域，图像的信息并不是均匀分布于全局，而只是集中在局部，因此并不需要全局连接，通过卷积操作也使得学习的参数量过大的问题得以解决，因此将图像处理中的关键技术与神经网络技术相结合，就形成了卷积神经网络^[54]（Convolutional Neural Networks, CNN），也被称为前向神经网络（Feed-forward Neural Networks）。

卷积神经网络 CNN 是一种包含了卷积计算、具有深度网络结构的前馈神经网络，是多层感知器的一种变化形式，通过局部连接与权值共享，优化了传统的神经网络，缓解了模型过拟合问题，使其更加适合于处理二维图像的一些问题。LeNet-5 卷积神经网络^[55]是在上世纪 90 年代首次被提出，并广泛适用于手写数字的识别问题，其网络结构如图 3-3 所示：

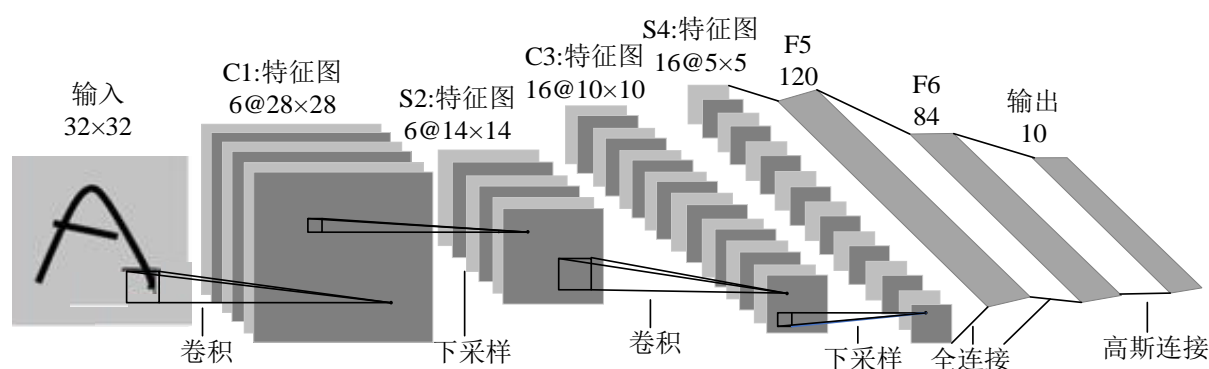


图 3-3 LeNet-5 卷积神经网络模型

卷积神经网络非常适合用于计算机视觉图像处理应用，在于其所具有的两个特性：局部连接和权值共享。

（1）局部连接

局部连接，也被称为稀疏连接。以往的深度神经网络由于每一层的每一个神经元都与上一层的所有神经元之间两两形成连接，因此需要训练更新的参数的数量非常巨大，这种完全连接的方式在使得全局中所有信息都被详细分析，但同时也带来了巨大的工作量，神经网络所需要训练的次数更多，训练的时间更长，训练所需的样本数据集更大。但是在图像处理领域中，图像的信息并不是均匀地分布在图像地每一个像素上，而是将主要信息集中在个别区域，因此对于全局地详细分析，并不能大幅提高图像检索的性能，反而使整个网络变得迟缓。然而，卷积神经网络通过不同网络层神经元之间得局部连接很好得解决了这个问题。

所谓局部连接，就是卷积层中感知器的节点仅仅与上一层感知器的部分节点进行连接，只用来学习局部的特征。假设第 n 层有 5 个神经元，上一层有 3 神经元，如果第 n 层的每一个神经元都与上一层的神经元进行相连，则应该有 15 个权值参数，但采用上图的局部连接方式，每个神经元仅与上层中相邻较近的几个神经元进行连接，共有 9 个权值参数，此时需要学习更新的权值参数大大减少。同样地，第 $n+1$ 层的神经元也与第 $n+2$ 层的神经元之间采取了局部连接的方式。采取这种局部连接的方式，可以极大地减少训练神经网络需要学习的权值参数的数量，提高了神经网络的训练效率，并且在一定程度上避免了过拟合现象的出现。

(2) 权值共享

权值共享也成为参数共享。权值共享的效果和局部连接是一样的，都是为了减少在训练过程中产生的权值参数的数量。假设图像中一定区域内各个像素之间的距离大致相同，那么这些像素之间的相关性也大致相同，尽管具体参数可能稍有不同，但总体趋势都大致相同，那么我们可以将这些像素之间的相关性近似地认为是一致的。对应到神经网络中，则可以认为同一层地神经元与上一层的神经元之间的连接关系近似相同，也就是权值参数的数值可以近似认为相等，这就是权值共享，假设第 n 层有 5 个神经元，若采取全连接方式，则共有 15 的权值参数，采用局部链接之后则需要 9 个参数，而当我们采取权值共享的方式时，需要学习训练的权值参数则仅有 3 个，大大减少了训练权值参数的数量。

卷积神经网络一般都包含输入层、隐含层和输出层，而卷积神经网络的隐含层中又包括了卷积层、池化层和全连接层。因此一个完整的卷积神经网络的结构一般由以下几个部分组成：输入层、卷积层、激活层、池化层、全连接层和输出层。具体的结构图如图 3-4 所示：

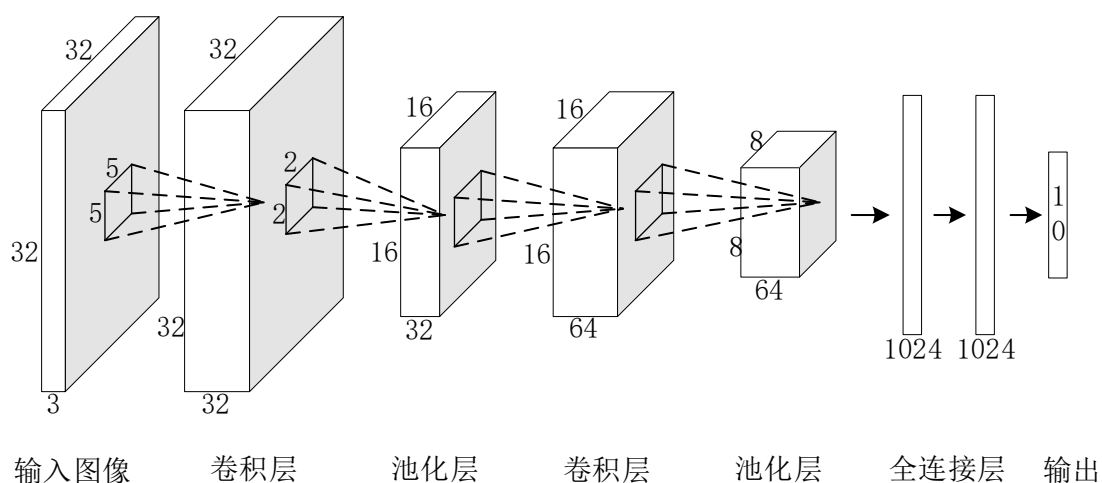


图 3-4 卷积神经网络结构图

(1) 卷积层

输入的图像在卷积层中，经过层层卷积层，逐步地将图像的特征提取了出来，卷积计算的过程是对输入图像的像素矩阵中的子区域逐个使用线性滤波器进行线性运算，

为了使得神经网络能够包含非线性因素的影响，因此通常还要再加上一个非线性运算变换，于是便可以通过计算得到输入图像的特征图。计算方式如下：

$$h_{ij} = f((W * x)_{ij} + b) \quad (3-6)$$

其中 h_{ij} 为该图像的特征图， w 为卷积核， $*$ 为卷积操作， b 为偏置量。

一个简单的卷积运算的例子如下图所示：

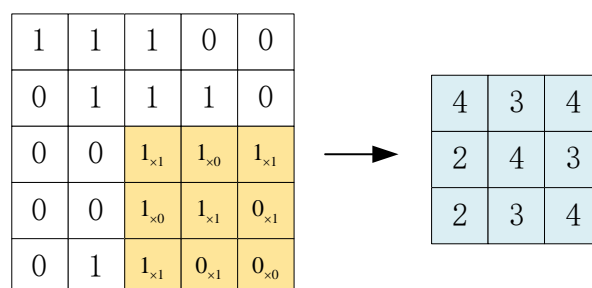


图 3-5 卷积运算示意图

(2) 池化层

池化层 (pooling)，也被成为欠采样或下采样，通常用作在卷积层之间加入池化层来进行特征降维、参数压缩等操作，来达到减小过拟合，以及提升模型性能的目的。常用的池化方法有两种：最大池化法 (Max Pooling) 和平均池化法 (Average Pooling)。最大池化法，就是在规定的空间区域内，从窗口的修正特征图中选取最大值作为该区域的最终值存入池化后窗口。例如下图中，整幅图像为 4×4 ，假设池化窗口为 2×2 ，则将原图像分割成 4 个 2×2 的子图，在子图中分别选取最大的数值 6、8、3、4 作为该子图的最终结果保留，便得到了原图像经过最大池化操作过后的池化图，具体如图 4-6 所示：

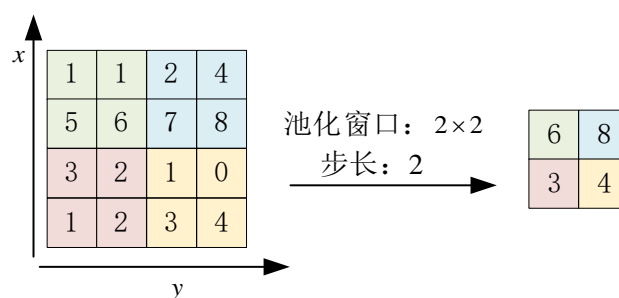


图 3-6 池化运算示意图

平均池化法，也是将图像分割成不同池化窗口，在每个区域中用平均值作为结果得到的池化图像。而在实际使用中，最大池化法的效果总是好一些。

(3) 全连接层

经过多次卷积、激活、池化操作之后，输入图像转变成具有较高质量的特征图像，在特征图像进入全连接层之前，有可能会出现过拟合现象，因此需要引入 dropout 操作，另外还有局部归一化 (LRN) 来增强系统的鲁棒性。卷积层之间的神经元总是采取的局部连接的方式，而全连接层也就是完全连接，他通常被分配在隐含层的最后一层，

连接方式则和传统的深度神经网络一致为完全连接。全连接层中将卷积层获得的特征图像进行加权和，转换成特征向量输出，因此得到的特征能够更加全面反映输入图像。

3.2 改进多尺度特征融合残差神经网络

残差神经网络是由微软研究院的 KaimingHe 等人提出的一种具有独特残差结构的卷积神经网络 ResNet^[56] (Residual Neural Network)。在 2015 年的 ImageNet 大规模视觉识别竞赛获得了冠军，top-5 的正确检索率提升到了 96.43%，同时大大减少了卷积神经网络模型训练过程中的参数的数量，缩短了模型训练需要的时间，整体性能相较于 VGGNet 获得了极大提升。其独特的残差结构如图 4-7 所示：

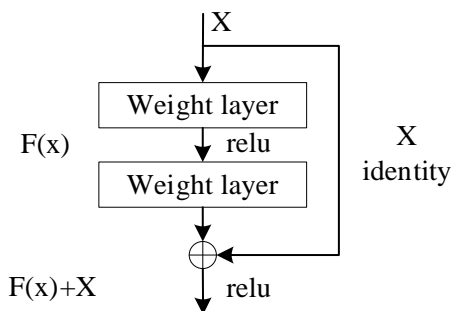
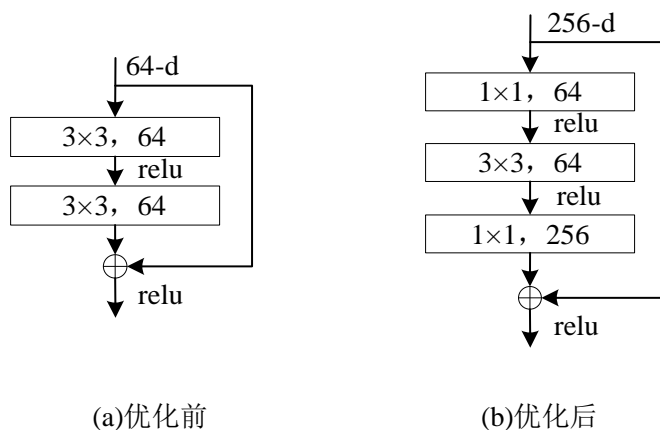


图 3-7 残差结构

图 3-7 为残差神经网络的独特结构，通过在多个卷积层之间设置一个直接进行连接的 shortcut 连接，这样的网络结构称为一个 block，即残差块。而在实际应用中，网络层数的大幅加深带来了数据计算成本大幅增加，因此使用大量的 1×1 卷积核来对残差块进行优化^[57]。优化结构如下图 3-8 所示



(a) 优化前

(b) 优化后

图 3-8 残差神经网络优化结构前后对比

由图 3-8 可以看出，将原来的 3×3 卷积核变成了首先采用 1×1 卷积核来减少计算量，经过 3×3 卷积核之后再通过一个 1×1 卷积核还原，在计算精度保持不变的基础上，大幅减少了计算量。因此相比于传统卷积神经网络，残差神经网络在解决因为网络层数过深带来的梯度弥散问题有独特优势。

3.2.1 残差神经网络模型结构

目前, 主流使用的残差神经网络结构有第一代残差神经网络 Resnet50, Resnet101 和第二代残差神经网络 Resnext50^[58]。这三种残差神经网络的结构如下表所示:

表 3-1 残差神经网络模型结构

卷积层	输出尺寸	Resnet50	Resnet101	Resnext50
Conv1	112×112		7×7, 64, stride 2 3×3 max_pool, stride 2	
Conv2_x	56×56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128, C = 32 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3_x	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256, C = 32 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Conv4_x	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512, C = 32 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Conv5_x	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 1024 \\ 3 \times 3, 1024, C = 32 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
Classifier	1×1	Averagr poor, 31 -d fc, softmax	Averagr poor, 31 -d fc, softmax	Averagr poor, 1000-d fc, softmax
FLOPs		3.8×10^9	7.6×10^9	4.2×10^9

残差神经网络结构包含有五个残差块: Conv1、Conv2_x、Conv3_x、Conv4_x 和 Conv5_x, 逐层计算图像的特征层次, 越浅的网络层提取出的图像特征分辨率越高, 其包含图像的信息更加全面, 但是对于图像的语义理解不够准确, 而深层网络层的分辨率依次下降, 表达图像语义能力加强, 每个残差块的最深层具有最强的特征。每一层的残差块的输入都是上一层残差块的输出, 则多层尺度计算以及级联方式过程可以用式表示:

$$F_n = \varphi_n(F_{n-1}) = \varphi_n(\varphi_{n-1}(\varphi_{n-2}(\cdots \varphi_1(I)))) \quad (3-7)$$

$$R = G(C_n(F_n), C_{n-1}(F_{n-1}), \cdots, C_{n-d}(F_{n-d})), n > d > 0 \quad (3-8)$$

式中各符号含义如表 3-2 所示:

表 3-2 符号含义

符号	含义
F_n	第 n 层卷积层的特征图
φ_n	第 $n-1$ 层到第 n 层特征图的非线性卷积运算
$\varphi_1(I)$	输入图像 I 得到第 1 层特征图的非线性卷积运算
$C_n(\bullet)$	在第 n 层特征图 F_n 上的检测结果
$G(\bullet)$	将所有不同尺度特征图的检测结果进行集合统计
R	算法最终的结果

由式 3-11、式 3-12 可知，第 n 层特征图上特征信息由第 $n-1$ 层特征图上信息决定，分辨率的大小决定了该层特征图上感受野大小。Resnet 用来预测的特征图 Conv1、Conv2_x、Conv3_x、Conv4_x 和 Conv5_x 中输出的特征最大尺度的分辨率为 56×56 ，如果输入的 PCB 图像中包含缺陷信息的特征尺寸比较小，经过 pooling 层池化后，有可能会丢失一部分低层网络所包含的全局细节特征，例如边缘信息，像素信息，这不利于对图像的特征进行提取，例如对一些小型缺陷信息在经过 4 个 pooling 层池化后，会使得该小型缺陷的特征图减少 $1/16$ ，甚至导致像素较小的缺陷数据的特征完全丢失。在低层特征图进行特征提取时，小尺寸的缺陷信息的特征没有明显表达，而经过层层卷积、池化后，高层低分辨率特征图保留的缺陷细节信息变少。

因此，低特征层次中高分辨率特征图像对于 PCB 缺陷检测这类微小物体很重要，高特征层次的特征图对于缺陷信息的表征也很重要。基于分析，提出在 Resnet101 网络上将深层神经网络提取到的抽象特征，配合相邻不同尺度特征图进行特征融合；低层网络结合更大分辨率特征层并与自上而下传递的高层语义融合，补充低层检测层上细节信息及语义信息。

3.2.2 多尺度特征融合

低层网络卷积层计算得到的特征图像拥有更大的分辨率，更好地概括了图像全局特征，深层网络对于特征的分类和识别更加有利。例如图 3-9 所示，将 PCB 图像输入到残差神经网络进行特征提取后，在 Conv2_3、Conv3_3 卷积层对特征图像进行可视化操作^[59]，得到的结果显示，Conv2_3 卷积层输出的特征图像的分辨率为 56×56 ，从特征图像上还可以观察到输入的 PCB 缺陷图像，但是细节部分特征并不明显。Conv3_3 卷积层输出的特征图像的分辨率为 28×28 ，相较于 Conv2_3 卷积层的高分辨率低特征细节，Conv3_3 卷积层则对细节特征表达更加丰富。

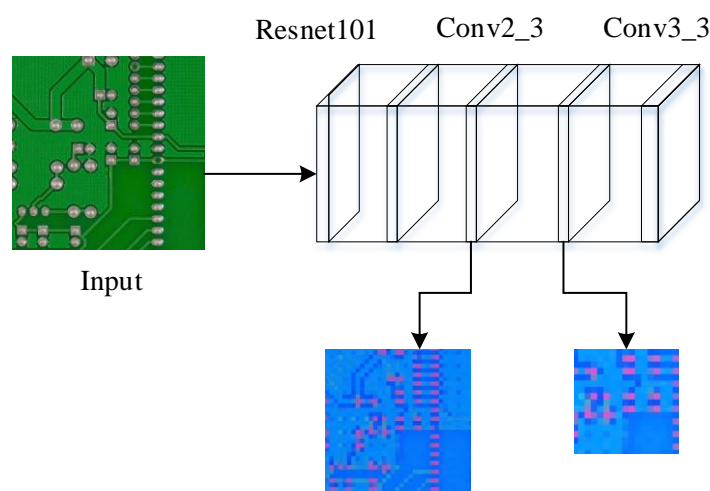


图 3-9 特征图像可视化

为了提高残差神经网络在低层网络对图像的微小缺陷特征的提取能力，改善神经网络对小型缺陷的检测效果，可以采取将相邻的低层高分辨率全局信息的特征图和高层低分辨率细节信息的特征图进行融合^[60]作为新的特征提取结果。网络特征融合结构如图 3-10 所示：

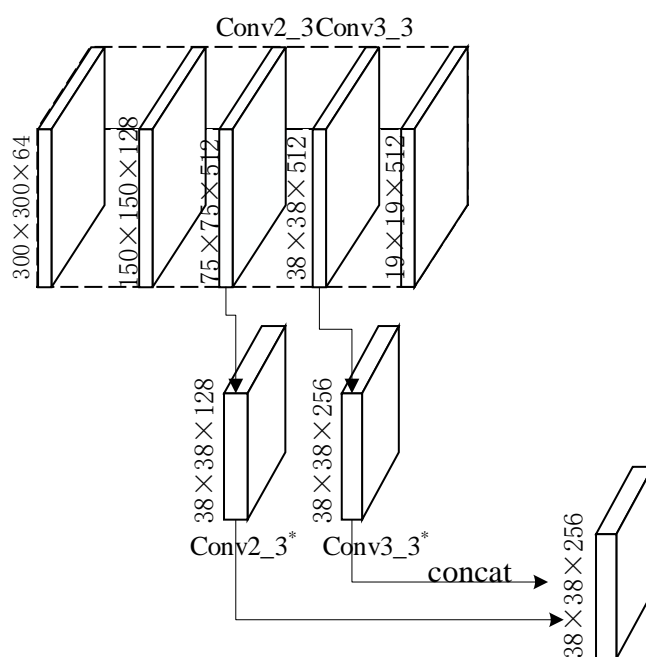


图 3-10 特征融合结构图

如图 3-10 所示，将 Conv2_3 层得到的特征图和 Conv3_3 层得到的特征图进行融合时，仍将更深层的 Conv3_3 层的输出尺寸作为融合结果的尺寸，采用 concat 方式融合两层的特征图，以补充 Conv3_3 层的卷积特征图的特征信息。

将相邻的低层网络特征图和深层网络特征图融合便可以补充深层网络丢失的细节特征，依次自上而下地将低分辨率强语义特征和高分辨率强结构特征横向连接起来，就形成多尺度特征融合金字塔结构，即 TDD-Net^[61]，结构如图 3-11 所示。

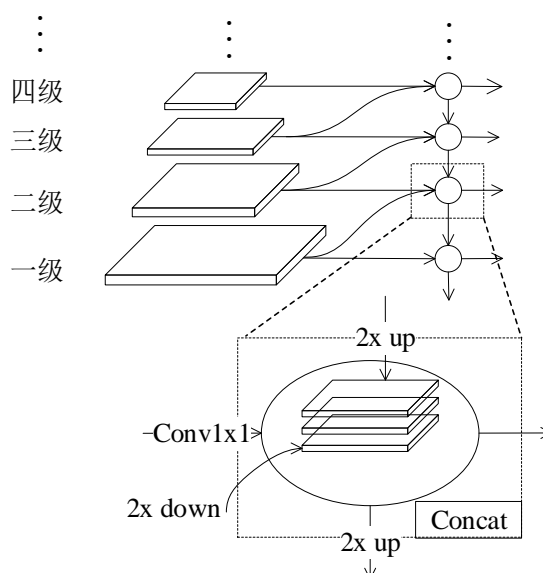
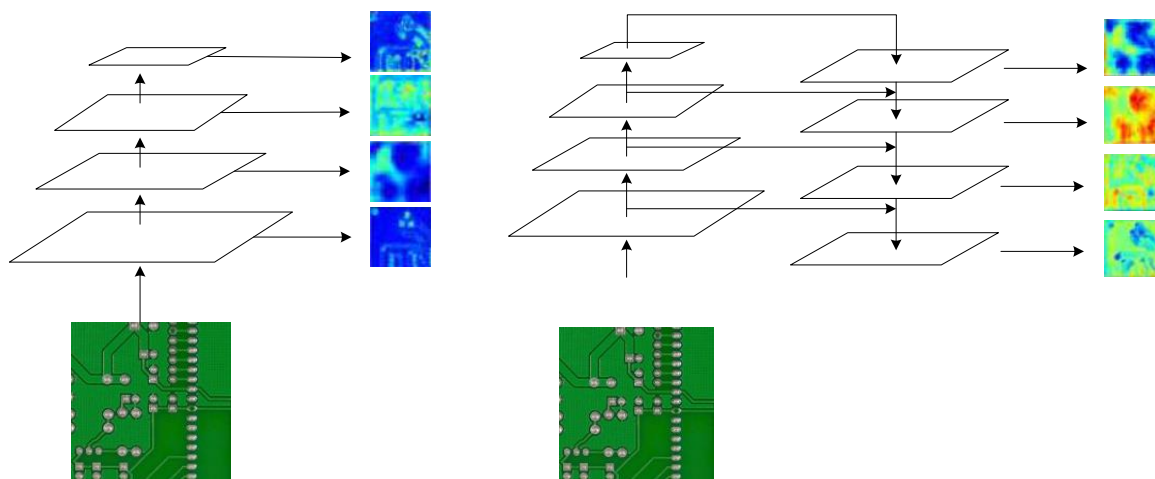


图 3-11 多尺度特征融合结构图

相邻两层特征图融合后不仅具有深层特征图的强特征并且具有低层特征图的强结构，良好的稳定性和准确性使得对于缺陷的检测和识别更加有利。如图 3-11 所示，在进行多尺度特征融合时，采用结合相邻两层不同分辨率的特征图，随后自上而下，将融合后的特征图依次传递，以逐步加深低层全局细节和高层强语义的特征融合。其中一级到四级在 Resnet101 神经网络中分别对应 Conv1 到 Conv4_x，则多尺度特征融合的具体流程如下：

- (1) 二级特征图经过 1×1 卷积核以减少通道维度简化计算，提取其特征并进行横向链接操作。
- (2) 一级特征图的分辨率是二级的两倍，对一级特征图采用最大池化下采样，以避免引入过量参数，并经过 1×1 卷积核计算。
- (3) 将 (1) (2) 步得到的特征图进行 Concat 融合，并将融合得到的结果进行线性插值两倍上采样操作。
- (4) 将 (1) - (3) 步得到的三个分辨率尺寸相同的特征图进行 Concat 融合以增强特征图像的语义信息。并且对融合结果经过 3×3 卷积核计算，以消除多次融合有可能产生的混叠效应。
- (5) 重复上述步骤，直至完成特征提取。

采用多尺度特征融合对 PCB 图像进行特征提取得到的结果如图 3-12 所示。至此完成了对残差神经网络 Resnet101 的改进，我们将改进后的残差神经网络称为 Resnet101-finetune，接下来便可以采用数据集对神经网络进行训练。



(a) Resnet101 最后一层特征图

(b) 多尺度特征融合结果

图 3-12 多尺度特征融合结果对比

3.2.3 损失函数与激活函数

残差神经网络对于缺陷的检测包括分类和框定，分别对应置信度损失（ $conf$ ）和定位损失（ loc ），损失函数即为这两者的加权和，计算方式如式 3-13：

$$L(x, c, l, g) = \frac{1}{N} (L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \quad (3-9)$$

其中， c 表示置信度， l 表示预测位置， g 表示真实位置， N 表示匹配对数， $L_{conf}(x, c)$ 表示置信度损失， α 为权值系数， $L_{loc}(x, l, g)$ 表示定位损失。

神经网络实现图像缺陷检测后得到正负样本以及每一个样本对应的位置，假设定位中心位置的横坐标为 cx ，纵坐标为 cy ，定位框图的宽和高分别为 w 和 h ，那么定位损失就可以通过下式计算：

$$L_{loc}(x, l, g) = \sum_{i \in Pos} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k smooth_{L1}(l_i^m - \hat{g}_j^m) \quad (3-10)$$

其中 $smooth_{L1}()$ 函数的输入并不是 (x_1, y_1, x_2, y_2) 或是 (cx, cy, w, h) ，而是经过编码计算之后的 $(\hat{g}_j^{cx}, \hat{g}_j^{cy}, \hat{g}_j^w, \hat{g}_j^h)$ ，具体计算方式如式 3-11、3-12

$$\hat{g}_j^{cx} = (g_j^{cx} - d_i^{cx}) / d_i^w \quad \hat{g}_j^{cy} = (g_j^{cy} - d_i^{cy}) / d_i^h \quad (3-11)$$

$$\hat{g}_j^w = \log\left(\frac{g_j^w}{d_i^w}\right) \quad \hat{g}_j^h = \log\left(\frac{g_j^h}{d_i^h}\right) \quad (3-12)$$

置信度损失的计算公式如 3-13 所示：

$$L_{conf}(x, c) = - \sum_{i \in Pos} x_{ij}^p \log(\hat{c}_j^p) - \sum_{i \in Neg} \log(\hat{c}_i^0) \quad (3-13)$$

$$\hat{c}_j^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \quad (3-14)$$

将定位损失和置信度损失进行加权求和便得到了损失函数。

激活函数中最常使用的是 Sigmoid 函数、Tanh 函数和 ReLU 函数，但是 Sigmoid 函数和 Tanh 函数均为饱和的非线性激活函数，卷积神经网络学习训练的前向和后向过程中，激活函数都参与其中，当神经网络输出的结果位于激活函数的饱和区域时，就会导致计算的权值结果急剧减小，因此采取适当的非饱和非线性激活函数 ReLU 函数最为合适，其数学公式如式 3-15 表示：

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (3-15)$$

由公式可以看出，ReLU 函数当 $x < 0$ 时，函数取值恒为 0，仍具有饱和性；当 $x \geq 0$ 时，函数为线性函数，导数值恒为 1，此时函数能够保持梯度不会衰减，从而解决了梯度弥散的问题。

3.2.4 训练集采集及模型数据迁移训练

第二章通过对 PCB 缺陷图像进行数字图像处理，实现了对 PCB 图像全局进行检测是否存在缺陷，并对缺陷定位。将定位得到的缺陷信息映射到原始 PCB 图像上，并将该区域的最小外接矩形进行分割，对分割得到的缺陷类型进行标注，即完成了 PCB 图像缺陷信息的定位和分类^[62]，来作为残差神经网络的训练数据集如图 3-13 所示：

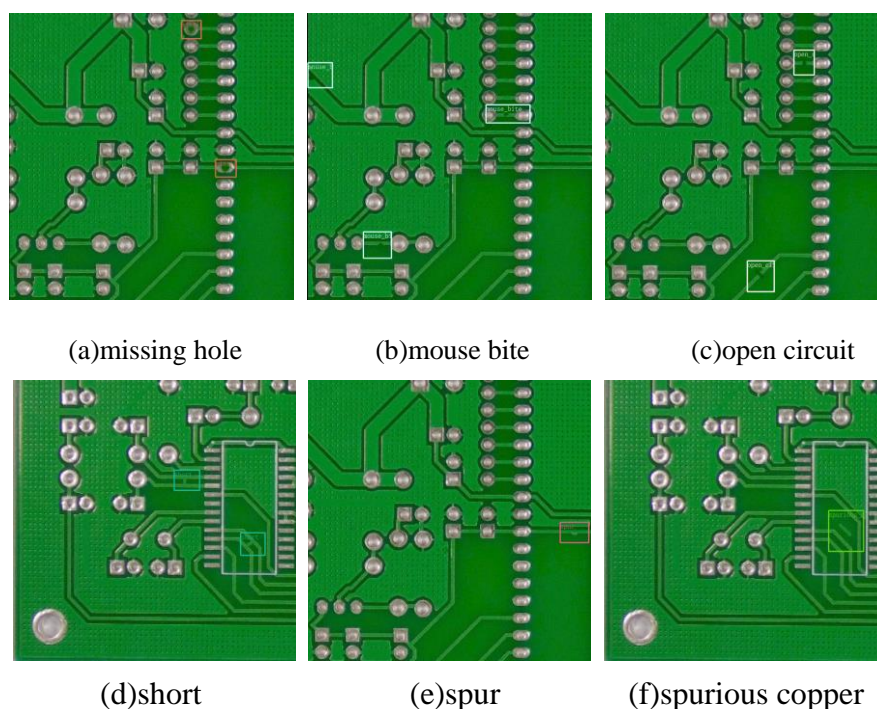


图 3-13 模型训练数据集

将精确标注得到的数据集按照接近 3: 1 分为训练数据集和验证数据集, 则训练数据集与测试数据集的具体情况如表 3-3 所示:

表 3-3 训练数据集与测试数据集

缺陷类型	训练数据集		测试数据集	
	图像数量	缺陷数量	图像数量	缺陷数量
缺孔	2748	5418	827	1630
鼠咬	2778	5526	835	1645
断路	2610	5322	786	1548
短路	2598	5262	725	1428
毛刺	2626	5454	734	1446
多铜	2640	5514	693	1452
共计	16000	32496	4600	9149

得到训练图像数据集的同时, 每张图片对生成一个与之对应的 XML 文件, 记录了该图像包含的缺陷数量、种类、位置等信息, 用于后续的训练以及模型验证。

每一个神经网络适用的领域都不十分相同, 在一些特定的领域, 我们很难收集大量的原始数据供神经网络训练学习, 如果要从 0 开始训练一个神经网络模型使其对待检测的图像数据能够达到良好结果所需要的原始图像以及训练的时间是非常庞大的, 这在现实中很难实现。例如, 在 PCB 缺陷检测方面, 能够用于模型训练的实际图像数据集仅有数千张, 仅依靠这些规模不是很大的图像数据集训练一个模型是很难的, 因此我们可以在已有针对图像检测的模型上加入我们需要识别的 PCB 缺陷图像数据集进行训练, 使得这个模型对于 PCB 图像也有很好的特征提取能力, 这就是数据迁移训练模型的方法。

数据迁移训练^[63]模型就是在大型数据集上预先训练一个深度卷积神经网络模型, 然后使用这个预训练的深度卷积神经网络模型的权值作为初始化设置, 继续训练学习我们需要的数据集, 大大缩减训练模型所需要的参数数量以及训练时间。数据迁移训练适用的场景一般有以下两种: (1) 新数据集规模较小, 但数据集内容与原始数据集类似。(2) 新数据集和原始数据集的规模内容均相似。在本文实验中研究对象为 PCB 缺陷图像数据, 而 ResNet 模型预训练的 Image 数据集中也包含有类似的图像数据, 因此, 我们可以采用数据迁移的方法在 ResNet 模型上继续训练我们的布料图像。

接下来便可以开始使用制作好的 PCB 缺陷图像数据集在多尺度特征融合改进过后的 Resnet101-finetune 神经网络上进行训练。分别将 PCB 缺陷图像训练数据集放在一个名为 train 的文件夹内, PCB 缺陷图像验证数据集放在一个名为 test 的文件夹内, 这两个文件夹分别对应 Resnet101 神经网络的训练和验证过程, 然后根据图像的名称、序号和类别号, 将这两个文件夹内的所有图像生成对应的 txt 文件, 按照两个文件夹名字对应, 分别为 train.txt 和 test.txt 文件。txt 文件内标签的格式为 “XXXX/X,X”, 前四个 X 对应于图像的序号, 分隔符之后的 X 代表图像的格式, 例如 jpg、png 等, 最后一个 X 代表图像所属的类别, 即图 3-13 中的缺陷分类。另外, 模型训练还需要 imagepairs.txt

文件，文件内包含了手工分类的两大部分，分别是属于同类的两张不同图像，和属于不同类的两张图像，用于后续验证迭代，文件格式即为“XXXX/X,X, XXXX/X,X, 0”和“XXXX/X,X, XXXX/X,X, 1”，其中前面为两个图像对应的标签格式，用中文逗号隔开，最后面的 0 或 1 代表这两张图像是否属于同一类别，若属于不同类别则为“0”，若属于同一类别则为“1”，该文件共包含 5000 条同类数据和 5000 条不同类数据。最后还需要一个空白文件 property。此外，由于 ResNet101 神经网络输入的图像尺寸为 256×256 ，因此，需要使用 convert_imageset.exe 程序将 PCB 缺陷图像数据集通过 resize 函数将图像分割成 256×256 的像素大小，并且将图像的格式转换成 Imdb 或 leveldb 格式。模型训练与验证过程所需的全部文件便准备好了。

首先在 PASCAL VOC2007（6）数据集上测试本实验改进的 Resnet101-finetune 神经网络，以验证网络的有效性，该神经网络的训练损失如图 3-14 所示：

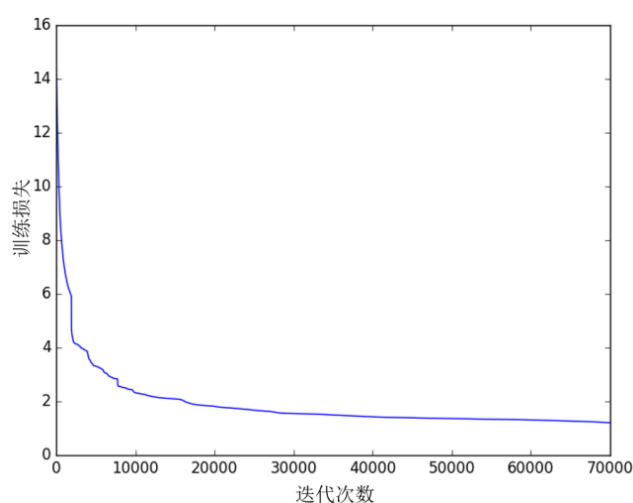


图 3-14 改进 Resnet101-finetune 训练损失

由图 3-14 可以看出，当迭代次数达到 10000 次以上时，训练损失趋于平稳，表明改进后的残差神经网络能够有效收敛，且在迭代次数达到 10000 次时已经基本达到检测效果。

残差神经网络的参数中 batch_size 为批样本数量大小，此数值若设置过大则单次训练压力过大，迭代次数较长，对计算机硬件设施要求较高，若设置过小，则单次训练迭代次数较少，但总体迭代训练增加，并且权值可能还没有收敛就已经达到最大迭代次数就停止迭代，随即停止更新权值，因此批样本数量大小应视情况选择合适的大小，此处设置为 48。

初始学习率 lr 的设置，基于 liner scaling rule 的原则，根据式 3-16 赋值：

$$lr = 0.00125 \times \text{Num_GPU} \times \text{IMG_PER_GPU} \quad (3-15)$$

其中，Num_GPU 为 GPU 的数量，本实验为 1，IMG_PER_GPU 为 batch_size 的图像数量，本实验为 48，根据式 3-16 计算得到本实验的初始学习率为 0.06

训练步数设置为 10000 步，权重每次更新时的衰减系数设置为 0.002，此时的参数能够帮助模型在训练过程中防止过拟合。则残差神经网络 Resnet101 的具体参数设置如表 3-4:

表 3-4 Resnet101 参数设置

参数	初始值	含义
Num_GPU	1	GPU 个数
Num_classes	8	图像类数
batch size	48	批大小
lr	0.06	学习率
train_steps	10000	训练步数
Weight_decay	0.002	衰减系数

具体的训练过程如下所示:

- (1) 准备 PCB 缺陷图像训练数据集以及验证数据集。
- (2) 准备好模型训练以及验证所需的文件: train.txt、test.txt、imagepairs.txt、property、convert_imageset.exe。
- (3) 将权值 W 和偏置项 b 设置为初始值，设定学习率为 0.06，设定批量样本数为 48，迭代次数为 10000 次。
- (4) 从 PCB 图像训练数据库中抽取第一次批量样本，传送到卷积神经网络中进行第一次训练。
- (5) 计算中间层的输出向量与神经网络最终得到的输出向量。
- (6) 根据得到的输出向量，以及图像本身标签所表示的向量进行计算，得到中间层向量之间的误差，以及最终的输出向量的误差。
- (7) 根据计算得到的误差，计算出对于权值和偏置项需要的调整量 ΔW 和 Δb 。
- (8) 根据上一步计算得到的调整量，更新神经网络的权值和偏置项。
- (9) 按照更新之后的权值的偏置项，重新训练训练数据集。
- (10) 迭代次数未达到设定值则返回第四步继续迭代，达到设定值之后结束。

3.2.5 实验对比

本文实验在 windows 系统上搭建了基于深度学习框架 MXNet 的编程环境，其详细的硬件和软件环境如表 3-5 所示:

表 3-5 硬件、软件环境配置

实验环境	配置说明
CPU	Intel (R) Core (TM) i3-2130 3.40GHz
硬件环境	GPU
	Nvidia GTX 750, 2GB 显存
	内存
	4GB
软件环境	操作系统
	Windows10, 64 位
	算法编程环境
	深度学习框架 MXNet

分别使用 Resnet50、Resnet101、Resnext50 和本文提出的 Resnet101-finetune 残差神经网络基于相同的表 3-4 给出的参数进行初始化设置，在表 3-3 给出的 PCB 缺陷图像训练数据集上进行训练，并在测试数据集上进行测试，各神经网络得到的测试结果在不同缺陷类型上的准确率如下表 3-5 所示：

表 3-5 不同神经网络准确率对比

缺陷类型	Resnet50	Resnet101	Resnext50	Resnet101-finetune
缺孔	0.918	0.965	0.904	0.973
鼠咬	0.884	0.932	0.868	0.940
断路	0.903	0.903	0.879	0.911
短路	0.915	0.938	0.909	0.943
毛刺	0.827	0.897	0.795	0.909
多铜	0.961	0.929	0.850	0.979
平均准确率	0.901	0.927	0.867	0.942

由表 3-5 中的数据可以看出，对于相同的缺陷类型：缺孔、鼠咬、短路和毛刺，不同神经网络的准确率对比，总有 Resnet101-finetune > Resnet101 > Resnet50 > Resnext50，当缺陷类型为断路时，Resnet50 和 Resnet101 的准确率相等。仅当缺陷类型为多铜时，有 Resnet101-finetune > Resnet50 > Resnet101 > Resnext50。由此可以看出，在这六种缺陷中，无论哪种缺陷类型，本文提出的改进后的残差神经网络 Resnet101-finetune 总能够对 PCB 缺陷图像进行准确检测并且识别缺陷类型。此外，对于平均准确率，Resnet-finetune 也是这几种神经网络中最高的，进一步说明了经过改进之后的残差神经网络对于 PCB 图像缺陷检测是有效的，并且在已有的方法的准确率上得到了进一步提升。

分别对这四种神经网络在测试数据集上测试的过程中的平均准确率、总时间、平均每张检测时间、平均每秒检测数统计结果如表 3-6 所示：

表 3-6 不同神经网络训练过程对比

	平均准确率	总时间	平均每张检测时间	平均每秒检测数
Resnet50	0.901	2822s	0.6135s	1.63
Resnet101	0.927	2857s	0.6211s	1.61
Resnext50	0.867	2721s	0.5917s	1.69
Resnet101-finetune	0.942	2805s	0.6098s	1.64

由表 3-6 可以看出，采用不同神经网络进行检测的速度相差不多，其中检测速度最快的是 Resnext50，但是其检测的平均准确率相对最低，根据表 3-5 中给出的数据，对于毛刺类型缺陷，Resnext50 的准确率相对于其他神经网络平均降低接近 0.1，因此其在 PCB 缺陷检测的应用上很难达到较好效果。而本文提出的基于多尺度特征融合的 Resnet-finetune 残差神经网络在速度上仅次于 Resnext50，但准确率却是最高。综合检测的效率与准确率，Resnet101-finetune 都是能达到理想效果的。

3.3 本章小结

本章主要提出了基于残差神经网络 Resnet101-finetune 的 PCB 图像缺陷检测方法。结合上章的图像预处理，针对传统 PCB 图像特征提取的局限性，提出使用卷积神经网络来对 PCB 缺陷图像进行特征提取。常用的 Resnet101 残差神经网络有效改善了卷积神经网络因为网络层数过深带来的梯度弥散参数爆炸问题，但是低网络层的特征图检测不到细节特征，高网络层的特征图检测不到全局信息，因此提出采用多尺度特征融合的方法进行改进，形成 Resnet101-finetune 残差神经网络，通过在相同的 PCB 缺陷图像数据集上进行对比，本文提出的方法平均检测准确率达到了 94.2%，在速度和检测准确度上均得到提升。该方法对于 PCB 图像进行缺陷检测提供了新的解决方案，具有研究意义和实用价值。

第四章 基于 Faster-RCNN 的 PCB 缺陷目标检测

通过用 PCB 缺陷图像训练残差神经网络，能够实现对输入的 PCB 图像进行检测其包含的缺陷类型，准确率也符合需求，但如何将缺陷信息定位并标记出来则需要依赖目标检测算法来实现。目标检测是机器视觉领域中一个非常重要的研究方向，也是如今深度学习的一个非常热门的问题，他的目的主要是将事先提取到的图像的特征进行精确定位，用方形边框标记出其缺陷所在的位置，并对特征的类别进行记录的过程^[64]。这一研究需求在各个领域都有着广泛的应用，例如交通行人检测、工业视觉检测、医疗病理检测等等。

4.1 目标检测算法

目标检测算法最早发源于 20 世纪末 21 世纪初，经过近二十年的不断发展，目标检测算法也得到不断改进，由早期的基于人工手动标注的传统方法，例如 Hog 特征结合 SVM 算法、Hear 特征结合 Adaboost 算法和 DPM 算法等等，传统算法无论是在特征的提取还是定位框标记等问题上都难以满足需求，已经逐步淘汰。随着深度学习的大力发展，如今基于深度学习的目标检测算法^[65]成为主流，例如 R-CNN 算法、YOLO 算法、SSD 算法等，相比于传统算法，基于深度学习的目标检测算法具备特征提取质量高、检测准确度高、检测速率高等优势，因此得到广泛应用。根据检测模式的不同，基于深度学习的目标检测算法可以分为以 YOLO^[66]为代表的 one-stage 网络^[67]和以 R-CNN 系列算法为代表的 two-stage 网络。

4.1.1 two-stage 网络

R-CNN 算法是由 Ross 等人在 2014 年最早提出的用于目标检测的算法^[68]，该算法的结构如图 4-1 所示：

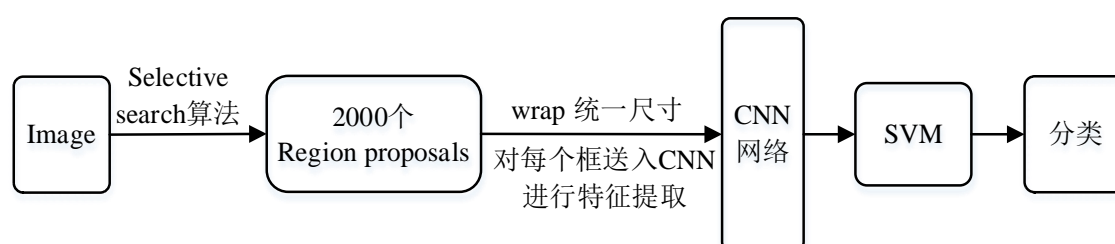


图 4-1 R-CNN 网络结构

由图 4-1 可以看出，R-CNN 算法首先通过 Selective Search 方法挑选出将近 2000 个 Region proposal 候选框，随后对这 2000 个候选框进行 crop 或 warp 归一化以统一所有候选框的尺寸大小，并将归一化处理后的候选框送入卷积神经网络进行特征提取，对特征提取的结果进一步使用多个 SVM 分类器进行分类。R-CNN 是最早能够实现目标

检测的算法，其对于目标的标记基本达到需求，准确率相比于之前也得到很大提升，但是由于 Selective Search 方法得到的 Region proposal 候选框数量过多，并且需要对每一个候选框都进行区域特征提取和计算，因此耗时很长，而且采用 Selective Search 方法没有考虑到候选框之间的重叠问题，因此虽然得到了尽可能完整的候选框，但是其中包含大量重复交叉的候选框内容，这些重复的大量计算极大降低了算法的运算效率。因此 R-CNN 算法在检测效率上较差，在存储资源上有大量浪费。

针对于 R-CNN 存在的候选框重复导致运算效率低的问题，何凯明等人在其基础上提出了 SPP-Net 算法^[69]，算法结构如图 4-2 所示：

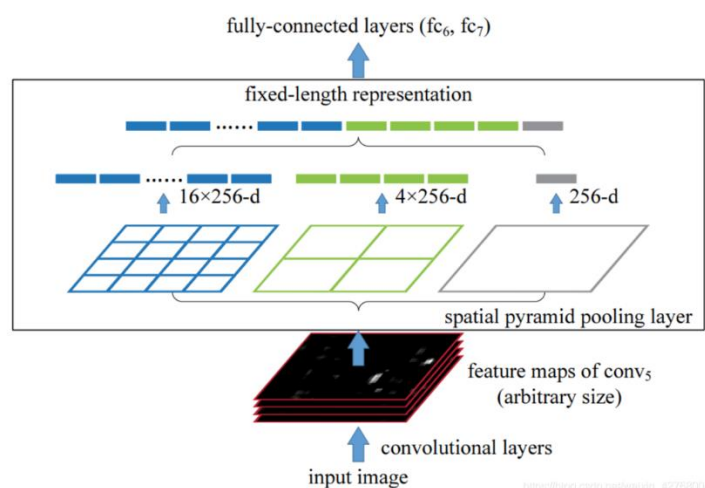


图 4-2 SPP-Net 网络结构

相对于 R-CNN 先提取候选框后对图像特征提取，SPP-Net 直接将一整张图像一次性送入卷积神经网络中进行图像特征提取，这就避免了 Selective Search 方法的大量重复无效计算，同时为了防止 CNN 图像中 2000 个候选框 crop/wrap 归一化造成的复杂操作，SPP-Net 算法在 CNN 末端添加空间金字塔池化层（SPP），提取固定大小的特征向量，大大减少了工作量。SPP-Net 算法在 R-CNN 算法上的两点改进，检测速度远远高于 R-CNN，但仍然没有解决存储空间消耗的问题。

2015 年，针对于 R-CNN 和 SPP-Net 算法在计算量大、存储资源浪费、检测速率低等问题，提出了 Fast-RCNN 算法^[70]，该算法的结构如图 4-3 所示：

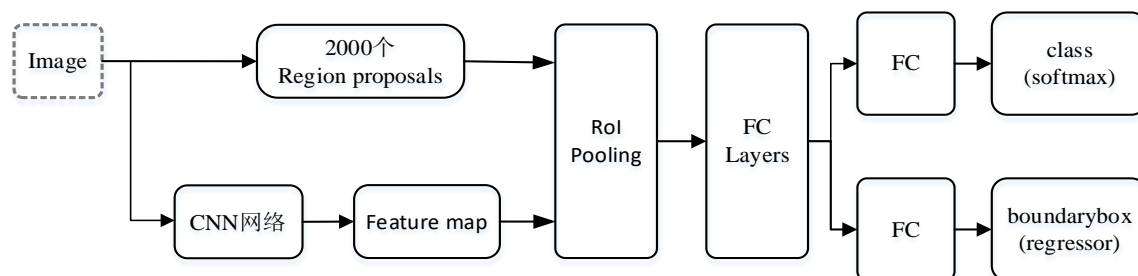


图 4-3 Fast-RCNN 网络结构

对比图 4-1 和图 4-3 可以看出，Fast-RCNN 算法使用了 ROI Pooling 来代替 SPP-Net 算法中的 SPP 层，并且不再使用 R-CNN 算法中的 SVM 分类器，取而代之的是将全连接层进行 SVD 分解，进而得到了两个输出向量，分别用于 Softmax 分类和用于边框回

归。这种改进的结果便是不再需要 SVM 分类器，而是将图像特征提取到的结果直接存储在显存中，大大节约了计算的数据量，并且对于存储空间和资源也极大减少消耗，使得整个算法的检测效率得到提升，并且使用 SVD 分解向量的方法也并不改变原有的检测准确度。

虽然在特征向量的分类上，Fast-RCNN 算法简化了计算，提高了检测效率，但是算法真正耗时的部分 Selective Search 方法仍然没有改变，检测的速度仍然难以达到实际需求。因此，对于候选框的生成选取方法需要进一步优化，据此提出了 Faster-RCNN 算法^[71]，该算法的结构如图 4-4 所示：

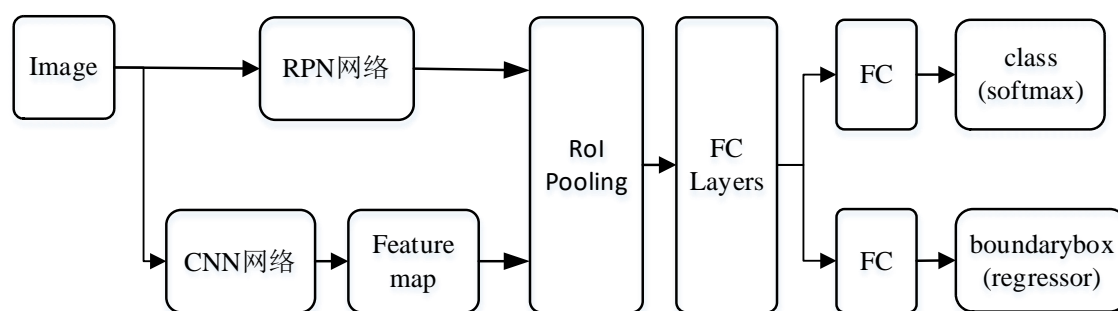


图 4-4 Faster-RCNN 网络结构

Faster-RCNN 算法不再采用计算量大、耗时长 Selective Search 方法，而是采用区域生成 RPN 网络，其产生的候选框是长宽具有一定比例的锚点框，很好地解决了使用 Selective Search 方法产生的大量时间，进一步提高了检测速度，使目标识别真正实现端到端过程。

从 R-CNN 算法能够初步实现目标检测的需求，到 Faster-RCNN 算法在候选框的生成方法和图像特征向量的分类方法不断优化改进，基于 Faster-RCNN 的目标检测算法在准确率和检测速度都能达到实际需求，因此在不同领域中都得到了广泛运用。

4.1.2 one-stage 网络

不同于 two-stage 算法的检测原理是将候选框生成和特征向量分类分为两步进行，one-stage 算法提出将“region proposal 问题和分类问题”转化成“regression”，一步到位直接得到目标的位置以及分类。one-stage 算法作为基于深度学习的目标检测算法的另一大分类，在检测的速度上远远过 two-stage 算法，但是在检测的准确程度上则稍显逊色，因此在实际情况中，应该综合分析检测的实时性与准确性的具体需求，再选择不同的目标检测算法。

2016 年，Joseph redom 等人首次提出 YOLO 算法，该算法作为 one-stage 算法的代表，抛弃了对于候选框的生成的计算过程，而是将整个目标检测的任务作为回归任务来处理，只需要将图像处理后直接送入卷积神经网络中进行特征提取，根据特征提取的结果便可以快速预测目标的类别、置信度分数以及 Boundingbox 的坐标。其检测流程如图 4-5 所示：

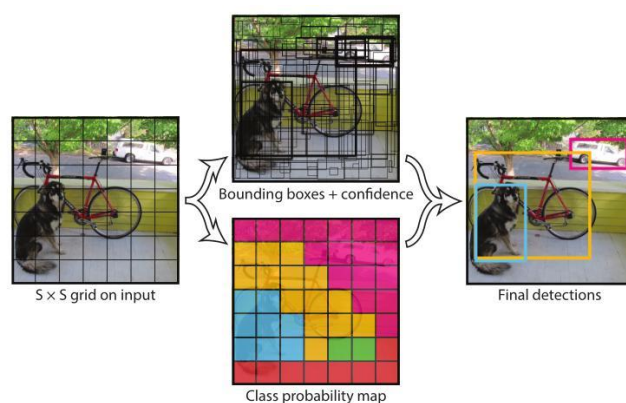


图 4-5 YOLO 检测流程示意图

由图 4-5 可以看出，YOLO 算法将输入的图像缩放到一定尺寸，并将图像划分成 $n \times n$ 个 cell（网格），每个网格负责检测物体中心落在该网格内的物体。YOLO 目标检测算法采用了 cell 这一网格形式为锚点框中心，抛弃了候选框的生成过程，每个 cell 同时预测两个尺度的 bounding box 以及类别信息，一步操作同时实现了位置定位和目标分类。YOLO 目标检测在检测速度上可达 45/s，基本能够满足不同场合的目标检测的实时性要求，但在精度上相较于 Faster-RCNN 差距较大，因此，YOLO 算法是在提升检测速度的同时牺牲了检测精度。此外，YOLO 算法使用每个 cell 预测两个 bounding box，当有多个物体中心落在同一个 cell 内时，或对于微小目标物体的检测，YOLO 目标检测算法十分容易出现漏检现象。并且，由于 YOLO 定位框的选择基于 cell 十分粗糙，因此对于目标物体的定位精确度较 Faster-RCNN 差。

4.2 基于 Faster-RCNN 的目标检测

近些年，在不同领域中都逐渐产生了目标检测的需求，不同的目标检测算法也被使用在各自领域中不同的数据集上。自 2005 年至 2012 年每年举行的图像识别挑战赛为图像识别和分类提供了一整套标准化的数据集 PASSCAL VOC，该数据集专用于图像分类、目标检测、图像分割等问题的研究。根据不同目标检测算法在 VOC2007 和 VOC2012 数据集上的检测效果，YOLO、SSD 等 one-stage 模型在检测的速度上更胜一筹，但是整体的检测准确度很难达到需求，并且漏检、误检的比例更大，而以 Faster R-CNN 为代表的 two-stage 模型在检测速度上虽然较慢，但是准确率有明显提升。考虑到对于 PCB 图像进行缺陷检测的需求，对于速度的要求并不是十分苛刻，而更加看重于检测的准确程度，因此本文在 Faster R-CNN 目标检测模型的基础上进行优化调整^[72]。

4.2.1 Faster-RCNN 网络结构

Faster R-CNN 的整体结构如图 4-6 所示，主要由卷积层、RPN（region proposal network）、RoI pooling、分类回归层四个部分组成。

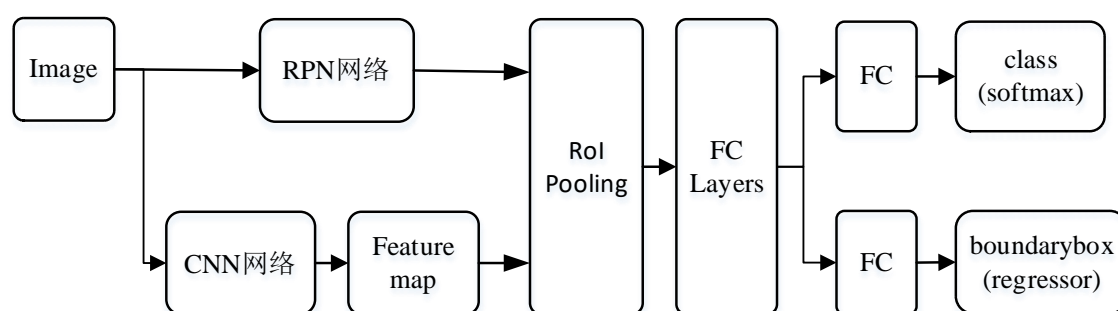


图 4-6 Faster-RCNN 网络结构

其中，卷积神经网络层由 3.2 给出的 Resnet101-finetune 神经网络实现，将输入的图像经过图像预处理调整为合适的格式，再通过该神经网络提取出图像的特征图像，传送至后续的 RPN 网络 and 全连接层。

相比于基础 R-CNN 目标检测算法，Faster-RCNN 目标检测算法的巨大提升在于其使用了 RPN 网络^[73]，该网络提出不再采用 Selective Search 方法，而是引入 anchor 机制来生成 region proposal。RPN 网络结构如图 4-7 所示：

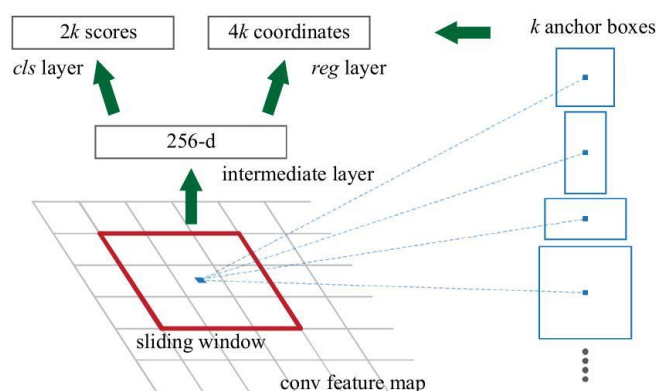


图 4-7 RPN 网络结构图

anchor 机制使用预先设定固定尺寸的窗口在特征图像上进行滑动的方式遍历图像的每一个像素，进而避免了对于图像中每一个候选框都进行计算的过程，减少了计算量，降低了学习的难度。RPN 网络生成 region proposal 的方式便是根据窗口滑动过程中的中心点作为采样结果生成 k 个 anchor，每一个 anchor 分别对应四个回归值 x 、 y 、 w 、 h 代表 anchor 的中心坐标及其宽度和高度，以及两个分类值 p_i 、 t_i 代表 anchor 为前景和背景的概率。

RPN 网络生成 region proposal 的过程显然是一个监督学习过程，即在假设空间中选取模型作为决策函数，对于给定的输入 X ，由 $f(X)$ 给出相应的输出 Y ，这个输出的预测值 $f(X)$ 与真实值 Y 可能一致也可能不一致，损失函数则是用于评估模型预测错误的程度，其首先是一个非负实值函数，通常用 $L(Y, f(X))$ 来表示。则 RPN 网络的损失函数定义为：

$$L(\{p_i\}\{p_i^*\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (4-1)$$

其中 i 代表一个 mini-batch 中 anchor 的索引, p_i 代表第 i 个 anchor 为前景的概率, 第 i 个 anchor 为前景时 p_i^* 为 1 否则为 0, t_i 是网络预测的 bounding box 的坐标, t_i^* 是 ground truth 的坐标。分类损失则为前景和背景两类的对数损失:

$$L_{cls}(p_i, p_i^*) = -p_i^* \log p_i - (1 - p_i^*) \log(1 - p_i) \quad (4-2)$$

回归损失为:

$$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*) \quad (4-3)$$

其中, R 为 $Smooth_{L1}(x)$ 损失函数:

$$Smooth_{L1}(x) = \begin{cases} 0.5x^2 & , |x| \leq 1 \\ |x| - 0.5 & , |x| > 1 \end{cases} \quad (4-4)$$

对于 bounding box 的回归采用以下 4 个坐标的参数化:

$$\begin{aligned} t_x &= (x - x_a) / w_a, & t_y &= (y - y_a) / h_a \\ t_w &= \log(w / w_a), & t_h &= \log(h / h_a) \\ t_x^* &= (x^* - x_a) / w_a, & t_y^* &= (y^* - y_a) / h_a \\ t_w^* &= \log(w^* / w_a), & t_h^* &= \log(h^* / h_a) \end{aligned} \quad (4-5)$$

其中, x 、 y 、 w 、 h 分别表示框的中心坐标及其宽度和高度, 变量 x , x_a , x^* 分别表示 bounding box, anchor 和 ground truth 的 x 坐标 (y , w , h 同理)。

Faster-RCNN 的算法完整结构如图 4-8 所示:

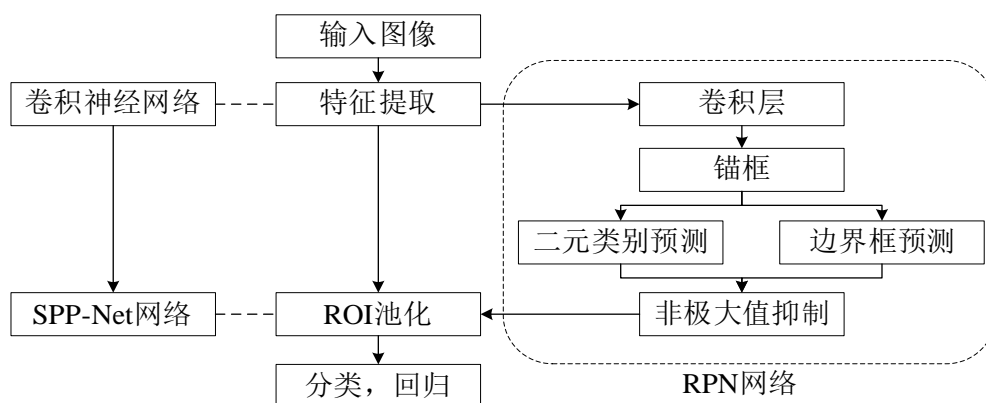


图 4-8 Faster-RCNN 算法结构

Faster-RCNN 算法在算法结构上与 Fast-RCNN 基本类似, 在其基础上进行了少量修改, 使用 RPN 网络生成候选框代替了 Selective Search 方法, 这样的改动结果使得目标检测的整个流程更加简洁, 检测速度得到了进一步提升。

4.2.2 模型训练

基于 Faster-RCNN 算法的目标检测最基础也是最重要的是 anchor 的设计, 在目标检测的 bounding box 的回归过程中, 模型会根据预先设计好的 anchor 进行反复迭代学习, 因此 anchor 的设计决定了训练过程的效率以及准确程度。假如 anchor 的设计接近 ground truth 的尺寸, 那么对于训练过程的学习难度将会大幅降低, 可以有效节省训练时间, 提高检测的准确程度。常用的 Faster-RCNN 算法中的 anchor 常被设计为三种大小 $\{128^2, 256^2, 512^2\}$ 以及三种比例 $\{1:1, 1:2, 2:1\}$, 也就是窗口在滑动过程中每个位置可以产生 9 个 anchor, 这种常用的设计方法对于常规物体的目标检测十分奏效, 但是对于 PCB 缺陷这种微小目标的检测并不十分有利, 因此需要对训练集的 ground truth 进行 K-means 聚类, 以生成合理的 anchor 尺度。

K-means 聚类是一种典型的无监督学习算法, 将原始样本集合选择其中 k 个类中心, 将样本与中心最接近的类别放在同一类中, 划分成 k 类, 然后计算每一个类的均值并作为新的类中心, 再次计算样本与类中心的距离并再次划分成 k 类, 如此循环往复直至收敛。其数学过程即为给定类中心 (m_1, m_2, \dots, m_k) , 计算一个划分 C , 使得目标函数极小化:

$$\min_C \sum_{l=1}^k \sum_{C(i)=l} \|x_i - m_l\|^2 \quad (4-6)$$

即对于确定 k 个类中心, 将样本集合分为 k 类, 使得所有样本和其类中心的距离总和最小。然后, 对于按照划分 C 得到的 k 类, 重新计算各个类中心 $(m_1^*, m_2^*, \dots, m_k^*)$, 再次计算使得目标函数极小化:

$$\min_{m_1^*, m_2^*, \dots, m_k^*} \sum_{l=1}^k \sum_{C(i)=l} \|x_i - m_l^*\|^2 \quad (4-7)$$

其中, 样本之间的距离的度量方式采用 IoU 函数, 如式 4-8 所示:

$$d(box, centroid) = 1 - \text{IoU}(box, centroid) \quad (4-8)$$

于是便通过 K-means 聚类方法得到了 anchor 尺寸, 结合本文的 PCB 缺陷图像数据集中的缺陷数据的分布, 最终采取的 anchor 尺寸为四种大小 $\{25^2, 30^2, 40^2, 50^2\}$ 和三种比例 $\{1:1, 1:2, 2:1\}$ 。

设定好 anchor 尺度之后, 便可以训练 RPN 网络, RPN 网络的输入训练数据为 3.2 节中 Resnet101-finetune 神经网络对 PCB 缺陷图像进行特征提取得到的特征图, 输出的便是良好的 region proposal。分别计算样本的 ground truth 和预先设计的 anchor 之间的 IoU 值, 将样本分为正样本和负样本两类, 其中 $\text{IoU} > 0.7$ 的为正样本, $\text{IoU} < 0.7$ 的为负样本。则训练过程, 对于每一个 batch 任意选取 256 个 anchor, 其中正负样本的比例为 1:1, 数量不足时可用部分相反样本补齐。

4.3 算法评价标准

常用于算法的评价标准有：准确率（precision），召回率（recall），平均检索精度（mAP）和 F1-Score。

准确率和召回率是目前在图像检测领域中应用最为广泛的评价指标。对于二分类问题，可以根据测试样本的真实类别和模型的预测类别划分为真正例（True Positive, TP），假正例（False Positive, FP），真反例（True Negative, TN），假反例（False Negative, FN），则混淆矩阵如表 4-1 所示：

表 4-1 混淆矩阵

		预测类别	
		正例	反例
真实类别	正例	TP	FN
	反例	FP	TN

准确率（precision）又叫查准率，指分类过程中模型所预测的所有正例内预测正确的比例，其计算公式如下：

$$P = \frac{TP}{TP + FP} \quad (4-9)$$

召回率（recall）又叫查全率，指分类过程中模型预测正确的正例占真实正例的比例，其计算公式如下

$$R = \frac{TP}{TP + FN} \quad (4-10)$$

不同的数据集对应的检索准确率与召回率也不相同，这两者通常与数据集本身的种类大小非常相关，因此为了使不同算法在不同数据集的检索结果之间相互对比，通常需要将准确率和召回率进行调和，即 F-Score。根据不同的权重值会得到不同的 F-Score，在实际使用中，通常认为准确率和召回率是同等重要的，因此二者的权重值相同，此时的 F-Score 值即为 F1-Score，计算方式如下：

$$\text{F1-Score} = \frac{2 \times P \times R}{P + R} \quad (4-11)$$

以准确率为纵轴，召回率为横轴作图，即可得到模型的 P-R 曲线。

平均检索精度（mean Average Precision, mAP）不同于上述几个指标，能够反映算法的整体性能，在多次预测结果中均正确得将相关的类别预测到了更前面的位置，则 mAP 的值也就越大，代表算法的整体性能越好。

mAP 的计算方式分为两步：首先需要计算出算法预测结果的平均准确率 AP，假设根据待检测图像输入后，算法根据输入图像进行预测并最终返回了 K 个算法认为最相关的类，并将类按照相似度排序，其中正确的相关类图像的排序位置分别为 x_1 、 x_2 、...、 x_k ，则该类别的平均准确率的计算公式如下：

$$AP_i = \frac{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_k}}{K} \quad (4-12)$$

第二步，对平均准确率 AP 进行算数平均数计算，计算公式如下：

$$mAP = \frac{1}{M} \sum_{i=1}^M AP_i \quad (4-13)$$

其中， M 为总类别数。

4.4 实验结果与分析

4.4.1 实验环境与数据

实验硬件配置：CPU 为 Intel (R) Core (TM) i3-2130 3.40GHz，GPU 为 Nvidia GTX 750，2GB 显存，8GB 内存。软件环境：操作系统为 Ubuntu16.04，深度学习框架为 MXNet，编程语言为 Python3.4，第三方库为 OpenCV，以及 CUDA。

实验采用的数据集为表 3-3 给出的印刷电路板 (PCB) 瑕疵数据集，包含 20600 张图像，6 个类别。数据集分为 16000 张图像的训练数据集和 4600 张图像的测试数据集，如图 3-13 所示。输入图像的尺寸为 256×256 。网络卷积层权重采用 xavier 初始化方法，整个网络采用批归一化 BN(BatchNormal)处理，lr 设为 $1e-4$ ，每进行 2 万次迭代 (iter)，学习率下降 10 倍，momentum 设为 0.9，weight_decay 设为 $5e-4$ ，batch_size 设为 4，进行 7 万次迭代。

4.4.2 实验结果分析

综合本实验改进算法，目标检测算法采用 4.2 节设计的改进 Faster-RCNN 目标检测，其主干网络采用 3.3 节设计的基于多尺度特征融合的残差神经网络 Resnet101-finetune，基于 4.4.1 给出的参数以及 3.2.4 给出的数据集进行训练和验证，得到趋于收敛的算法模型。

在该模型上对 2.1 节给出的 PCB 缺陷图像进行缺陷检测，分别计算在不同缺陷类型的情况下，模型检测的准确率与召回率，并绘制成各自的 P-R 曲线，该曲线反映了算法对于不同缺陷类型能否准确检测出来并对其缺陷类型进行分类的准确程度，P-R 曲线越靠近上方，则表示算法对于该种类型缺陷的检测性能越好，反之则越差。得到的不同缺陷类型的 P-R 曲线结果如图 4-9 所示：

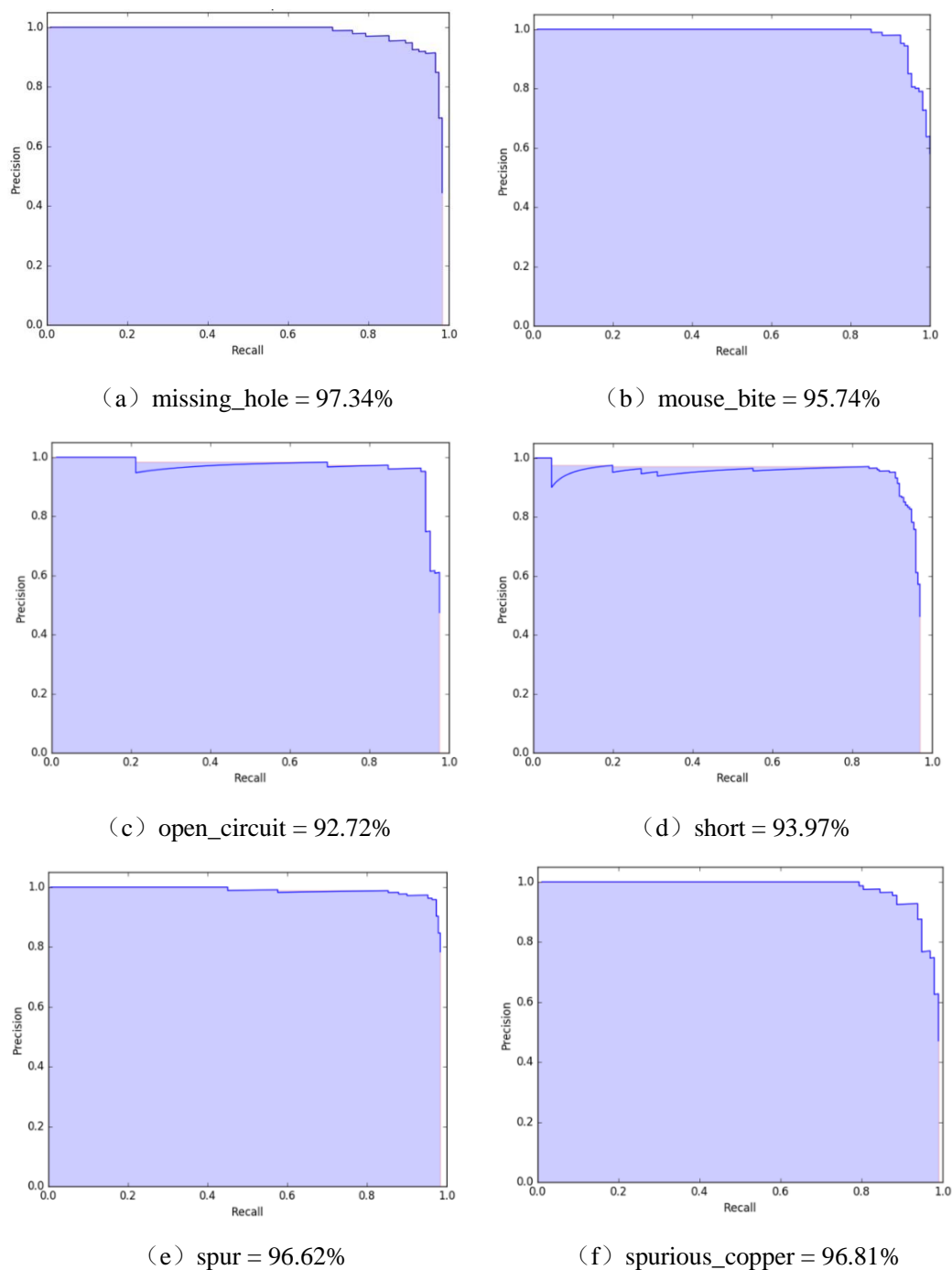


图 4-9 不同缺陷类型的 P-R 曲线

由图 4-8 中可以看出本文提出的基于图像处理和深度学习的 PCB 缺陷检测算法在数据集上的六种不同类型的缺陷的检测准确率 (AP) 均达到了 92% 以上。其中, 对于缺孔类型的缺陷的检测准确率最高, 达到了 97.34%, 而整体的平均准确率也高达 95.5%, 检测速度 FPS 为 3.7, 检测精度取得了良好的效果之外, 检测的速度也基本满足了工业生产线上对于 PCB 缺陷检测得实际需求, 由此可以看出本文提出的算法能够有效实现 PCB 图像的缺陷检测。

为测试改进后算法的检测性能, 采用 mAP (Mean Average Precision) 为评价指标, 即所有类标签的平均精确率。选取印刷电路板 (PCB) 瑕疵数据集中包含 6 类缺陷的图

像 4600 张，分别对比 YOLO、R-CNN、SSD 和本文改进的算法的 mAP，得到的实验结果如表 4-2 和图 4-10 所示：

表 4-2 不同算法的 mAP 对比

缺陷类型	YOLO	R-CNN	SSD	Ours
missing_hole	0.937	0.938	0.905	0.973
mousee_bite	0.857	0.942	0.861	0.957
open_circuit	0.895	0.918	0.877	0.927
short	0.902	0.899	0.896	0.939
spur	0.912	0.885	0.781	0.966
spurious_copper	0.905	0.915	0.812	0.968
mAP	0.901	0.916	0.855	0.955

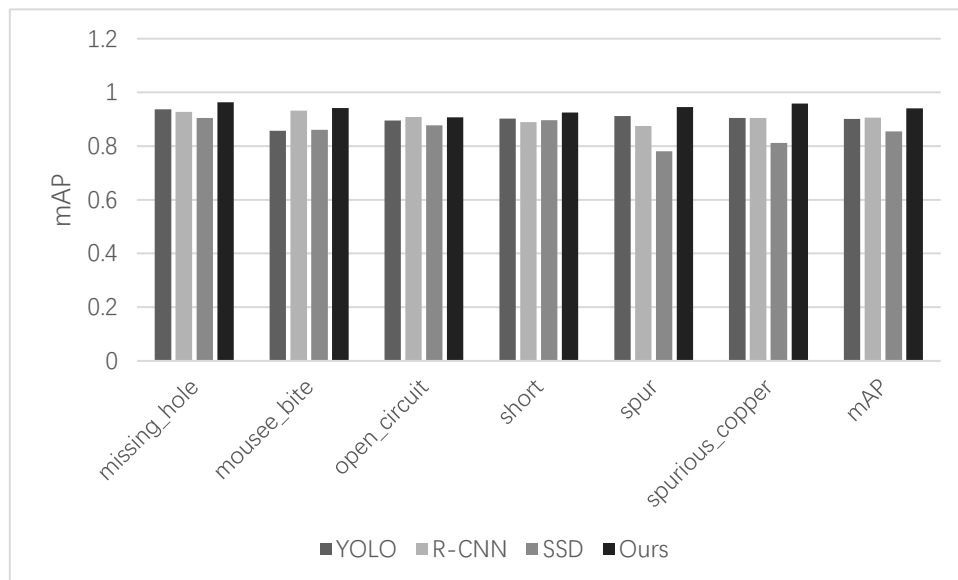


图 4-10 不同算法的 mAP 对比

由表 4-2 和图 4-10 可以看出，YOLO 算法和 R-CNN 算法的 mAP 分别为 90.1% 和 91.6%，检测准确度相近，表明针对于 PCB 图像进行缺陷检测采用 one-stage 和 two-stage 目标检测算法的准确率相近，而 YOLO 算法在检测的时间上相对较短。同为 one-stage 的 SSD 算法的检测结果是这四种方法中最差的，为 85.5%。而本文提出的算法的平均检测精度达到了 95.5%，相比于 YOLO 算法和 R-CNN 算法有明显的提升，其中对于缺孔、多铜等缺陷类型的检测准确率最高，达到了 97% 左右，而对于鼠咬、断路的检测结果与其他算法接近。对于同一种缺陷类型检测结果来说，本文提出的算法的平均检测精确度相比于其他算法总有所提升，实验结果表明了本文提出的算法在实现 PCB 图像缺陷检测上有良好的效果，且相比于已有的算法，在平均检测精确度上有了明显提升，具有较高的实际应用价值。

4.4.3 算法检测实例

为了验证本文算法的通用性，拟对包含多种缺陷的 PCB 板进行图像采集，并对采集到的图像进行预处理，随后送入预先训练好的模型，对 PCB 图像进行缺陷检测，检测结果的准确率验证本文算法的通用性。

使用海康威视的 MV-CA060-10GC 千兆以太网工业面阵相机对 PCB 板进行图像数据采集，该相机的感光芯片类型为 CMOS，即面阵数字相机，具有集成性高、速度快和成本低等优势，相机的最大分辨率为 3072×2048 ，最大帧率为 17fps，数据接口和像素格式分别为 Gigabit Ethernet 和 Mono，相机图像如图 4-11 (a) 所示。光源采取两条线性 LED 光源，如图 4-11 (b) 所示。



(a) MV-CA060-10GC 工业相机



(b) 线性 LED 光源

图 4-11 工业相机与光源选型

为了避免环境光照条件对拍摄成像的影响，将整个图像采集装置放置在黑箱环境中，在黑箱环境中，用两条线性 LED 光源分别位于 PCB 板前后侧上方进行光照补偿，以避免光照直射 PCB 板导致局部光照过强而四周光照较弱产生的局部曝光。将 PCB 板水平放置在桌面上，并使用 MV-CA060-10GC 工业相机以垂直俯视角度进行拍摄，并对 PCB 板进行图像采集，图像采集的装置以及实时界面如图 4-12 所示：

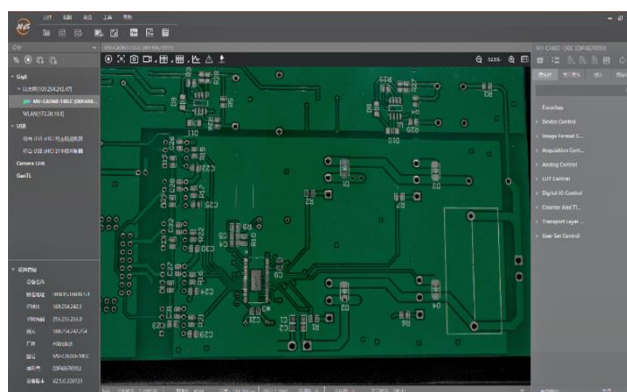
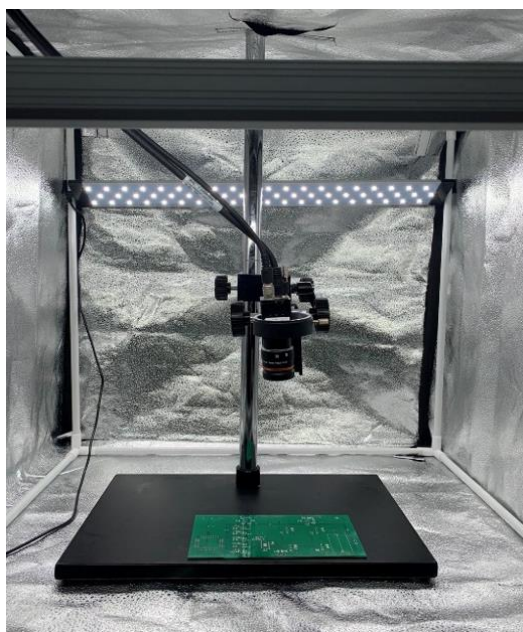


图 4-12 PCB 图像采集装置及界面

将拍摄得到的 PCB 图像进行预处理成合适的送入算法的格式,得到的 PCB 图像如图 4-13 所 (a) 示,由图中可以看出该 PCB 包含两处“缺孔”、一处“短路”、一处“断路”、两处“多铜”、一处“毛刺”等缺陷信息。将预处理过后的 PCB 图像送入算法中进行缺陷检测,检测后得到的结果如图 4-13 (b) 所示。

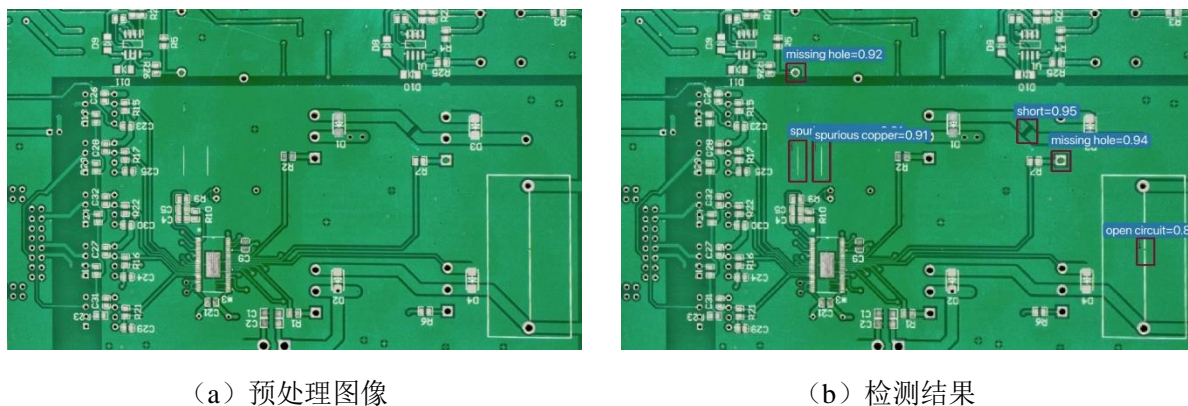


图 4-13 PCB 图像预处理及检测结果

由检测结果可以看出,算法检测出了该 PCB 的六处缺陷信息,框定出各自的具体位置并对缺陷类型进行预测。其中,对于缺孔、短路和多铜等缺陷信息的位置框定十分准确,缺陷分类准确,预测的概率也较高,表明算法对于这几种缺陷类型能够准确检测出来,并能准确框定其位置以及分类其缺陷类型。对于断路缺陷信息的位置和分类较为准确,但预测概率相对较低,推测原因是该处断路缺陷并不十分明显,断路的地方特征相对较弱,导致算法对于该处的检测不确定性增大。

通过对 PCB 进行实物图像采集,并将采集到的图像送入本文提出的算法,经过预处理和缺陷检测之后得到了准确的检测结果,表明本文算法对于 PCB 图像能够进行准确的缺陷检测,对于一般的 PCB 均具有通用性。

4.4.4 不同算法检测结果对比

基于北京大学智能机器人开放实验室免费公开提供的 PCB 微小缺陷图像数据集,在包含相同的缺陷信息的 PCB 图像上,如图 4-14 所示,分别采取检测效果相对较好的 YOLO 算法、R-CNN 算法以及本文提出的算法,对其进行缺陷检测,比较分析不同算法之间的检测结果,以验证本文提出的算法的改进效果。

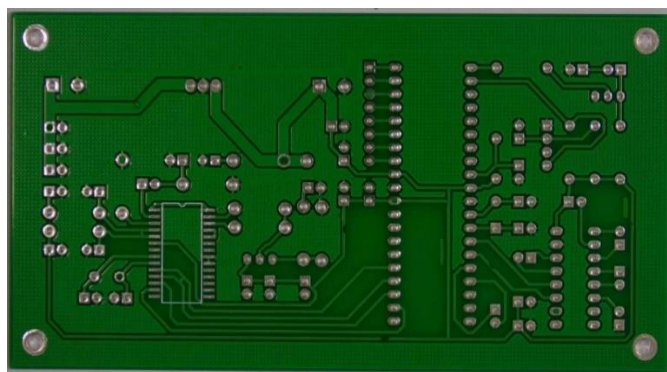
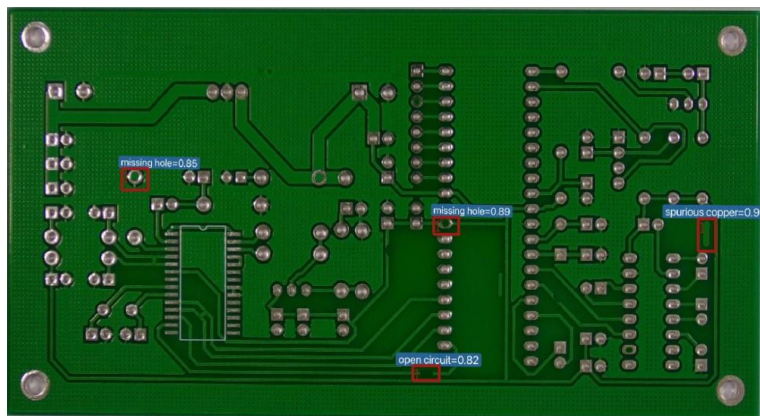
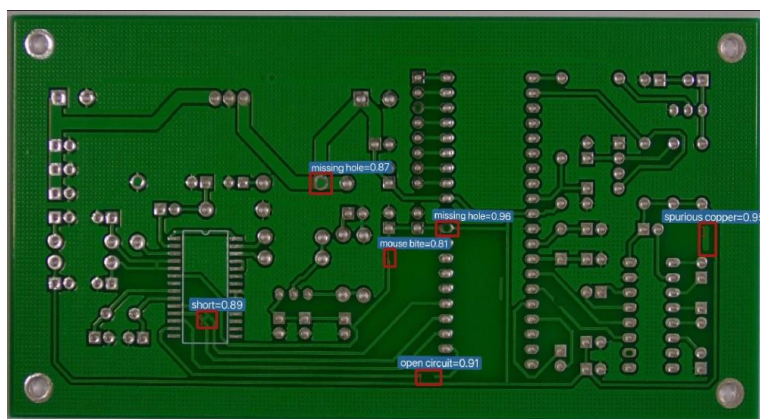


图 4-14 待检测 PCB 图像

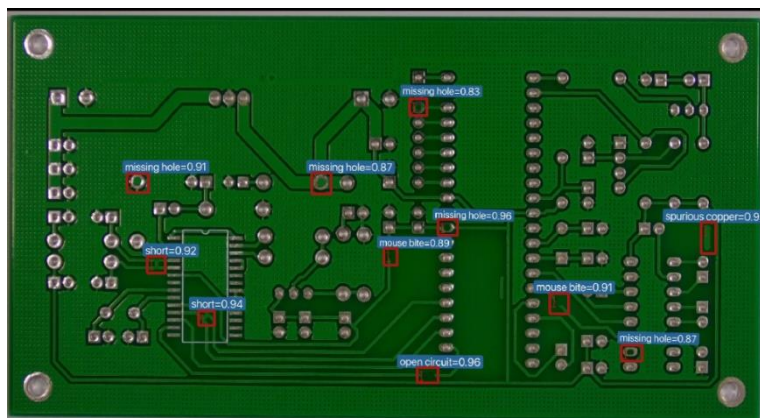
由图中可以看出,该 PCB 板孔、线等特征的形状和位置信息十分清晰,并且包含有五处“缺孔”缺陷、两处“鼠咬”缺陷、两处“短路”缺陷、一处“断路”缺陷和两处“多铜”缺陷。经过不同算法进行缺陷检测得到的检测结果如图 4-15 所示。



(a) YOLO 算法检测结果



(b) R-CNN 算法检测结果



(c) 本文算法检测结果

图 4-15 不同算法的检测效果对比

由图 4-15 可以看出,对于同一张 PCB 图像的预测结果,三种算法检测出的缺陷类型、数量以及概率都不一样。检测结果的数据统计如表 4-3 所示。YOLO 算法仅能够检测出了缺孔、断路、多铜三种类型的五个缺陷,且对于每一个缺陷预测概率并不高,算法对于 PCB 中存在的缺陷信息不敏感,虽然检测速度相对较快,但是漏检、误检现象经常发生。R-CNN 算法相对于 YOLO 算法在检测数量上有所提升,共计七个缺陷,

表 4-3 不同算法检测结果数据对比

缺陷类型	YOLO	R-CNN	Ours
missing_hole1	0.85	未检测出	0.91
missing_hole2	未检测出	0.87	0.87
missing_hole3	0.89	未检测出	0.83
missing_hole4	未检测出	0.96	0.96
missing_hole5	未检测出	未检测出	0.87
mousee_bite1	未检测出	0.81	0.89
mousee_bite2	未检测出	未检测出	0.91
open_circuit1	0.82	0.91	0.96
short1	0.75	0.89	0.92
short2	未检测出	0.87	0.94
spurious_copper1	未检测出	未检测出	未检测出
spurious_copper2	0.91	0.95	0.95
共计	5	7	11

并且检测出的缺陷标记的位置与类别基本正确，但对于一些微小昏暗的缺陷信息仍未能检测出来，该算法只适合用于检测较大尺寸的缺陷类型，对于微小型特征的提取识别能力较差。而本文提出的算法在该 PCB 图像上共检测出了 11 个缺陷，仅有一处多铜未能检测出，并且对于检测出来的缺陷分类准确，定位清晰，表明本文算法对于 PCB 图像中的缺陷能够精确定位和分类，采取多尺度特征融合的卷积神经网络对于微小型缺陷特征也有很好的提取和识别效果。总之，本文提出基于 Resnet101-finetune 特征提取和改进 Faster-RCNN 目标检测算法针对 PCB 图像进行缺陷检测是十分有效的，并且相对于已有的方法，在检测的准确度、全面性上均有所提升。

4.5 本章小结

根据残差神经网络 Resnet101-finetune 提取的 PCB 缺陷图像的特征图，设计了基于改进 Faster-RCNN 的目标检测算法，实现了在 PCB 图像上检测缺陷信息，并对缺陷分类、定位和预测。通过在相同数据集上，分别对比了 YOLO、R-CNN、SSD 和本文提出的算法的检测结果，根据 mAP 值反映出本文提出的算法的准确率和平均检测精度均高于其他方法。并通过实例采集 PCB 板图像进行缺陷检测验证了本文算法的通用性。表明本文提出的 PCB 图像缺陷检测算法是有效且具有广泛使用价值。

第五章 主要结论与展望

5.1 主要结论

本文以印刷电路板（PCB）瑕疵图像为研究对象，针对大规模的 PCB 图像数据集缺陷检测困难的问题，提出了一种基于图像处理、Resnet101-finetune 特征提取和改进 Faster-RCNN 的 PCB 图像缺陷检测算法。相比于传统的基于图像处理的特征提取或 onestage 目标检测算法，本文提出的缺陷检测算法在检测的准确率、检测时间、学习能力、鲁棒性等方面都得到了较大提升，不仅对电子产业的缺陷检测具有现实意义，而且对于图像数据处理领域具有研究意义。本文主要的工作内容以及研究成果如下：

（1）对于原始的 PCB 图像，往往因成像条件不同导致图像特征不明显，并伴随产生很多噪声，不能直接用于特征提取。本文采用将图像灰度化，以摒弃色彩信息干扰，再对灰度图进行直方图均衡化以增强前景与后景的对比，最后通过高斯滤波等方式降低图像的噪声干扰，实现了 PCB 图像突出特征，过滤噪声的目的。

（2）预处理之后得到的 PCB 图像可以首先通过机器视觉图像处理技术进行缺陷的全局检测和初步定位。基于参照对比算法，使用 SURF 算法提取 PCB 标准图像和缺陷图像各自的特征点，并对特征点进行匹配实现图像配准，再用阈值分割将图像处理成二值图，随后进行图像差分，便可以得到缺陷图像中包含的缺陷信息，通过形态学处理让缺陷更加清晰，以实现 PCB 图像的初步检测和定位。

（3）传统的基于图像处理提取的 PCB 图像特征不明显，基于深度学习方法对 PCB 图像进行特征提取有效提高了特征提取的质量。传统的残差神经网络 Resnet101 有效解决了网络层数过深带来的梯度弥散问题，但是低网络层高分辨率对图像语义特征提取不足，高网络层低分辨率对图像细节特征提取不足，本文提出多尺度特征融合的方式改进网络，以提升特征提取的质量，形成了 Resnet101-finetune 网络。通过迁移学习的方式在 PCB 图像数据集上训练该网络并进行验证，实验结果表明改进后的 Resnet101-finetune 神经网络的平均准确率达到了 94.2%，相比于其他图像处理方法或神经网络均得到明显提升。

（4）基于 Faster-RCNN 目标检测算法实现了对 PCB 图像的缺陷目标检测。通过改进 Faster-RCNN 算法提升目标检测的效率和准确度。在相同的 PCB 图像数据集上采用 YOLO、R-CNN、SSD 和本文提出的算法进行对比，实验结果表明本文提出的算法在平均准确率上达到 95.5%，mAP 值也均高于其他算法，验证了本文算法的有效性。并通过实例验证采集 PCB 板图像，处理后送入算法中进行检测，检测结果表明了本文算法的实用性，可以适用于 PCB 图像缺陷检测，具有较高的实际使用价值。

5.2 展望

限于客观条件，本文也有一些不足之处，可以通过以下几个方面继续改进：

（1）基于机器视觉图像处理技术可以实现对于 PCB 图像的缺陷进行全局检测和初步定位，但是对于图像的缺陷定位并不精确，对于缺陷的类型也无法判断，因此可以尝试继续通过更多的形态学处理方法，将图像差分得到的结果进行图像形状的处理，根据不同的缺陷类型对应的差分结果的不同形状，来对缺陷进行分类。

（2）基于残差神经网络 Resnet101 进行特征融合后对于 PCB 图像的特征提取大大增强，但是这仍然是在原有网络上进行改进，受制于残差神经网络的识别准确率，很难有大幅提升，目前卷积神经网络正处于飞速发展的时期，各种不同的网络模型层出不穷，因此可以尝试使用更多不同的卷积神经网络进行学习，以找到最适合应用于 PCB 图像特征提取的神经网络。

（3）用于训练神经网络的 PCB 图像数据集规模较小，且采用迁移学习的方法进行训练，而深度学习模型对于数据集的规模需求较大，因此训练的效果有限，后续可以对数据集进行全方位补充，包括更多种类型的缺陷、不同成像条件下的图像、不同型号的 PCB 板等，以增强神经网络对于复杂情况的适应能力。

致 谢

时光荏苒，转眼间在江南大学七年的时间已接近尾声。我在这里度过了最美好的大学时光，也收获了最美好的回忆。一路上有欢笑有成长，也有泪水有磨砺，感谢帮助过我的人，也感谢一直陪伴的人，更重要的是，感谢自己一路走来的艰苦奋斗、砥砺前行勇气。

首先我要感谢的是，我的原导师于振中老师，于老师给我提供了十分宝贵的平台：哈工大机器人合肥国际创新研究院，在人工智能所机器视觉组我度过了为期一年半的实习生涯。感谢我的组长瞿顶君，在实习过程中不断指导我的研究方向，并给出建设性意见，让我在机器学习方向收获良多。

我要感谢我的现研究生导师朱一听，在我毕业论文撰写过程中一直帮助我，为我出谋划策解决困难。我要感谢我的研究生同学们，三年的同学情谊难以磨灭，还望日后不忘初心。最后我要感谢我的室友王佳玮，是你一直陪伴着我，度过了所有开心或不开心的时光，让我不至于太孤单，并且一直包容我关心我，万分感谢。

研究生生涯即将结束，我也找到了满意的工作，希望在接下来的社会生涯中我能够继续保持在学校时的刻苦学习精神，不断提升超越自己。

参考文献

- [1] 陈茹. 印制电路板产业的发展对我国经济的影响[J]. 经济研究导刊, 2017(28):38-39.
- [2] 杨宏强. 中国内资 PCB 产业现状研究[J]. 印制电路信息, 2021, 310(05): 6-10.
- [3] 谭安菊, 顾炳林. 基于 IPC-7351B 的表贴器件 PCB 封装设计[J]. 兵工自动化, 2019 (7): 11-11.
- [4] Vishwakarma S, SahilTikke, Manurkar C, et al. PCB DETECTION AND CLASSIFICATION USING DIGITAL IMAGEPROCESSING[J]. Image Detection, 2014, 52(6): 1327-1350.
- [5] 马灿. PCB 缺陷智能视觉检测系统研究与设计[D]: [硕士学位论文]. 长沙: 湖南大学, 2012.
- [6] Liu Z, Qu B. Machine vision based Online detection of PCB Defect[J]. Microprocessors and Microsystems, 2021, 82(9):103807.
- [7] Meenu Maninder, et al. A concise review on food quality assessment using digital image processing[J]. Trends in Food Science & Technology, 2021, 118(PA): 106-124.
- [8] Wang Y, Peng W X, Qiu C H, et al. Fractional-order Darwinian PSO-Based Feature Selection for Media-Adventitia Border Detection in Intravascular Ultrasound Images[J]. Ultrasonics, 2018, 04(16): 24-35.
- [9] 白明柱, 杨立森, 阿木古楞. 图像特征提取的研究进展[J]. 科技创新与应用, 2013 (04): 46-46.
- [10] 刘沛津, 王曦, 贺宁. 改进 GSO 与二维 OTSU 融合的红外图像多阈值分割方法[J]. 应用光学, 2021, 42(4): 1-7.
- [11] 姜兰池, 沈国强, 张国焯. 基于 HSV 分块颜色直方图的图像检索算法[J]. 机电工程, 2009, 26(11): 54-57.
- [12] 张少博, 全书海, 石英, 杨阳, 李云路, 程姝. 基于颜色矩的图像检索算法研究[J]. 计算机工程, 2014, 40(06): 252-255.
- [13] Patil J K, Kumar R. Classification of Maize Leaf Diseases using Color Moments[J]. Digital Image Processing, 2011, 37(12): 68-72.
- [14] 薄华, 马缚龙, 焦李成. 图像纹理的灰度共生矩阵计算问题的分析[J]. 电子学报, 2016 (01): 155-158+134.
- [15] 赵玉丹. 基于 LBP 的图像纹理特征的提取及应用[D]: [硕士学位论文]. 西安: 西安邮电大学, 2015.
- [16] Parashivamurthy R, Naveena C, Kumar Y. SIFT and HOG features for the retrieval of ancient Kannada epigraphs[J]. IET Image Processing, 2021, 40(1): 20-23.
- [17] Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features[J]. Proceedings of the 9th European conference on Computer Vision - Volume Part I, 2006, 22(09): 74-83.
- [18] Terasawa K, Tanaka Y. Slit Style HOG Feature for Document Image Word Spotting[C] International Conference on Document Analysis & Recognition. IEEE Computer Society, 2009, 26(10): 110-118.
- [19] 薛超, 牟云平, 张尧. 基于多核学习的 GIST 全局和 SIFT 局部特征融合遥感图像检索方法[J]. 青岛大学学报(自然科学版), 2020 (1): 25-29.

- [20] Gul M J, Urfa G M, Paul A, et al. Mid-term electricity load prediction using CNN and Bi-LSTM[J]. The Journal of Supercomputing, 2021, 58(02): 74-86.
- [21] 李丹, 沈夏炯, 张海香, 朱永强. 基于 Lenet-5 的卷积神经网络改进算法[J]. 计算机时代, 2016 (08): 4-6+12.
- [22] Guo L, Song J, Li X R, et al. Haze Image Classification Method Based on Alexnet Network Transfer Model[J]. Journal of Physics: Conference Series, 2019, 1176:032011.
- [23] 王婷, 李航, 胡智. 一种 VGGNet 的图像风格迁移算法设计与实现[J]. 计算机应用与软件, 2019, 36(11): 224-228.
- [24] 王文成, 蒋慧, 乔倩, 祝捍皓, 郑红. 基于 ResNet50 网络的十种鱼类图像分类识别研究[J]. 农村经济与科技, 2019, 30(19): 60-62.
- [25] Moganti M, Ercal F. Automatic PCB inspection systems[J]. IEEE Potentials, 1995, 14(3): 6-10.
- [26] Sven K. Computer-implemented multi-dimensional medical image comparison method for comparison of e.g. tumor image, with reference image of patient, during diagnosis of tumor, involves detecting signal to use context for image on other image: DE102009011540[P]. 2010.
- [27] Suzuki T, Sagara N, Sugiyama K. Non-reference estimation method of image quality summarizing four kinds of detection for the coded degradation[J]. Ieice Technical Report, 2011, 111: 23-26.
- [28] Faghih-Roohi S, Siamak H, Alfredo Núñez, et al. Deep Convolutional Neural Networks for Detection of Rail Surface Defects[C] International Joint Conference on Neural Networks. IEEE, 2016. 23-147.
- [29] D Soukup, Huber-Mrk R. Convolutional Neural Networks for Steel Surface Defect Detection from Photometric Stereo Images[C] International Symposium on Visual Computing. Springer, Cham, 2014. 284-472.
- [30] Cha Y J, Choi W, Suh G, et al. Autonomous Structural Visual Inspection Using Region - Based Deep Learning for Detecting Multiple Damage Types[J]. Computer ided Civil and Infrastructure Engineering, 2018, 33(4): 1-17.
- [31] Suliman B, Featherston C A, Kennedy D. A hybrid method for modelling damage in composites and its effect on natural frequency[J]. Computers & Structures, 2019, 213(MAR.): 40-50.
- [32] 蒋东晖. 浅析计算机图像处理和识别技术的应用[J]. 信息通信, 2020 (07): 126-127.
- [33] 佚名. 生益电子国内首家通过 IPC-6012/IPC-A-600QML 认证[J]. 印制电路资讯, 2016, 000(004): 66-66.
- [34] 冯晓硕, 沈樾, 王冬琦. 基于图像的数据增强方法发展现状综述[J]. 计算机科学与应用, 2021, 11(2): 13-13.
- [35] Chernov V, Alander J, Bochko V. Integer-based accurate conversion between RGB and HSV color spaces[J]. Computers & Electrical Engineering, 2015, 46: 328-337.
- [36] 袁誉乐, 曹建民, 赵勇, 等. 一种图像灰度化方法及存储介质[P]. 中国专利, 108765501B. 2020-01-04.

- [37] 周果, 司秀丽. 基于颜色分量与直方图的玉米苗期背景分割方法研究[J]. 农业网络信, 2012 (08): 22-24.
- [38] 徐建东. 一种直方图均衡插值的图像细节增强方法[J]. 江苏理工学院学报, 2020 (2): 23-29.
- [39] 张威. 空间域中均值滤波与中值滤波去噪的应用研究[J]. 产业科技创新, 2020, 2(12): 65-66.
- [40] 彭俊, 魏琼, 王君, 童明浩. 改进中值滤波算法及其 FPGA 硬件实现[J]. 湖北工业大学学报, 2016, 31(02): 42-44+72.
- [41] 赵洁, 贾春梅, 虞凌宏. 高斯滤波算法在缺陷视觉检测中的应用研究[J]. 宁波工程学院学报, 2014, 26(04): 7-11+22.
- [42] Greene W H, Chelikani S, Purushothaman K, et al. Constrained non-rigid registration for use in image-guided adaptive radiotherapy[J]. Medical Image Analysis, 2009, 13(5): 809-817.
- [43] 刘敏, 周聪, 汤靖博. 基于 SIFT 图像配准算法优化研究[J]. 湖北工业大学学报, 2020, 35(02): 32-36.
- [44] 吴一全, 谢芬. 基于对比度 Harris 的快速鲁棒图像配准算法[J]. 北京理工大学学报, 2020, 30(03): 90-98.
- [45] 张锐娟, 张建奇, 杨翠. 基于 SURF 的图像配准方法研究[J]. 红外与激光工程, 2009, 38(001): 160-165.
- [46] Rajinikanth V, Dey N, Raj A, et al. Harmony-Search and Otsu based System for Coronavirus Disease (COVID-19) Detection using Lung CT Scan Images[J]. arXiv, 2020 (11): 23-29.
- [47] 吴闯, 于大泳. 基于轮廓修补和图像差分的手持玻璃盖板表面缺陷检测[J]. 软件工程, 2020 (7): 41-45.
- [48] 李绍坚, 于立军, 王国立, 李海振, 张福华. 形态学图像处理在缺支检测中的应用[J]. 科技资讯, 2011 (36): 251-251.
- [49] 陈则西. 基于灰度形态学开运算的遥感图像处理[J]. 科技创新与应用, 2012 (10): 18-18.
- [50] 任柳清. 基于图像处理的 PCB 焊接缺陷检测方法研究及实现[D]: [硕士学位论文]. 绵阳: 西南科技大学, 2021.
- [51] 何国忠, 梁宇. 基于卷积神经网络的 PCB 缺陷检测[J]. 图学学报, 2021 (07): 1-8.
- [52] 袁曾任. 人工神经网络及其应用[M]. 北京: 清华大学出版社, 1999, 7-13.
- [53] 李晓峰, 刘光中. 人工神经网络 BP 算法的改进及其应用[J]. 四川大学学报(工程科学版), 2000, 32(2): 105-109.
- [54] Cai Pei, Quan Hui min. Face anti-spoofing algorithm combined with CNN and brightness equalization[J]. Journal of Central South University, 2021, 28(1): 194-204.
- [55] Maraoui A, Messaoud S, Bouaafia S, et al. PYNQ FPGA Hardware implementation of LeNet-5-Based Traffic Sign Recognition Application[C]. 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD). 2021, 112-118.
- [56] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning[J]. Computers and Science, 2016 (23): 126-132.

- [57] 薛昕惟, 刘日升, 王祎, 等. 一种基于优化算法结合残差网络的单幅图像除雨方法[P]. 中国专利, 110111267. 2019-07-26.
- [58] Miret S, Chua V S, Marder M, et al. Neuroevolution-Enhanced Multi-Objective Optimization for Mixed-Precision Quantization[J]. Medical Image Analysis, 2021 (17): 112-119.
- [59] Shang J, Yang L, He K. 基于特征可视化分析深度神经网络的内部表征[C]. CCFAI, 2019.
- [60] 何婧媛, 谢生龙, 田原, 田琴琴. 多尺度特征融合的目标检测算法[J]. 河南科学, 2021, 39(07): 1045-1051.
- [61] R Ding, Dai L, Li G, et al. TDD-Net: A Tiny Defect Detection Network for Printed Circuit Boards[J]. CAAI Transactions on Intelligence Technology, 2019 (31): 27-35.
- [62] Li Xiaofeng, Wang Yanwei, Cai Yingjie. Automatic Annotation Algorithm of Medical Radiological Images using Convolutional Neural Network[J]. Pattern Recognition Letters, 2021, 152: 158-165.
- [63] 石祥滨, 房雪键, 张德园, 等. 基于深度学习混合模型迁移学习的图像分类[J]. 系统仿真学报, 2016, 028(001): 167-173.
- [64] 葛雯, 宫婷, 王媛, 等. 基于深度学习的目标检测算法[J]. 微处理机, 2019 (3): 29-33.
- [65] 张婷婷, 章坚武, 郭春生, 陈华华, 周迪, 王延松, 徐爱华. 基于深度学习的图像目标检测算法综述[J]. 电信科学, 2020, 36(07): 92-106..
- [66] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]. Conference on Computer Vision & Pattern Recognition. IEEE, 2017. 6517-6525.
- [67] Brami S, Lamarche J P, F Souyris. Treatment of facial asymmetries by one-stage maxillary and mandibular bilateral osteotomies[J]. International Journal of Oral Surgery, 1974, 3(5): 239-242.
- [68] Gkioxari G, Hariharan B, Girshick R, et al. R-CNNs for Pose Estimation and Action Detection[J]. Computer ence, 2014 (07): 57-64.
- [69] Purkait P, Zhao C, Zach C. SPP-Net: Deep Absolute Pose Regression with Synthetic Views[C]. British Machine Vision Conference, 2017. 115-132.
- [70] Girshick R. Fast R-CNN[J]. Computer Science, 2015 (17): 31-37.
- [71] Sun X, Wu P, Hoi S. Face Detection using Deep Learning: An Improved Faster RCNN Approach[J]. Neurocomputing, 2018, 299(19): 42-50.
- [72] 耿振宇. 基于 Faster R-CNN 的 PCB 表面缺陷检测[J]. 科学技术创新, 2021 (11): 86-87.
- [73] 王森. 基于改进 RPN+PCANet 的雾天光学卫星图像船舶目标识别算法研究[D]:[硕士学位论文]. 天津: 天津大学, 2019.

附录：作者在攻读硕士学位期间发表的论文

- [1] 于振中, 秦岭. 基于深度学习的布料图像检索方法研究[J]. 科技与创新, 2021, 21(00): 50-53.