

DOI: 10.11992/tis.201710005

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20180413.0946.004.html>

卷积神经网络的贴片电阻识别应用

谌贵辉, 何龙, 李忠兵, 亢宇欣, 江泉宇

(西南石油大学 电气信息学院, 四川 成都 610500)

摘要: 贴片电阻由于其体积微小、性能稳定等独特的性质, 在当今智能化的电子设备中被广泛使用。为保证贴片电阻的出厂质量, 需要对其进行缺陷识别、极性方向识别、正反面识别和种类识别, 目前很大程度上依靠人工肉眼进行识别检测, 效率低、容易误检、成本高。本文针对传统图像识别方法的局限性, 结合近年来卷积神经网络在图像识别方面所取得的巨大成就, 基于 AlexNet 模型、GoogLeNet 模型、ResNet 模型思想设计了 3 种深度适宜、可训练参数约 4×10^6 (百万) 的卷积神经网络, 克服了当前主流卷积神经网络模型由于可训练参数过多、模型层数太深导致在贴片电阻识别应用中识别速度不能满足实时性要求、泛化识别准确率低的问题。实验表明, 3 种模型的识别准确率均超过 90%, 最高识别准确率达到 95%, 识别速度达到 0.203 s/张 (256 像素 \times 256 像素, CORE I5)。因此, 本文设计的 3 种卷积神经网络可根据具体实际需求进行选用, 在实践中具有极强的可行性和可推广性, 同时也在提升企业生产效率和产品质量方面具有重要意义。

关键词: 贴片电阻识别; 卷积神经网络; AlexNet 模型; GoogLeNet 模型; ResNet 模型

中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2019)02-0263-10

中文引用格式: 谌贵辉, 何龙, 李忠兵, 等. 卷积神经网络的贴片电阻识别应用[J]. 智能系统学报, 2019, 14(2): 263-272.

英文引用格式: CHEN Guihui, HE Long, LI Zhongbing, et al. Chip resistance recognition based on convolution neural network[J]. CAAI transactions on intelligent systems, 2019, 14(2): 263-272.

Chip resistance recognition based on convolution neural network

CHEN Guihui, HE Long, LI Zhongbing, KANG Yuxin, JIANG Xiaoyu

(School of Electrical Information, Southwest Petroleum University, Chengdu 610500, China)

Abstract: Chip resistors are widely used in intelligent electronic devices because of their unique properties such as small size and stable performance. The chip resistors produced by the factory must be identified for defects in both front and back faces, polarity, and type in order to guarantee the quality. However, such identification largely relies on the eye detection, which is inefficient, prone to error, and costly. In this paper, considering the limitation of the traditional image recognition methods and the great achievements of convolutional neural network (CNN) in image recognition in recent years, three CNN models, AlexNet model, GoogLeNet model, and ResNet model, with appropriate depth and training parameters of about 4M (million) are designed to overcome the demerits of low speed that results in the inability to meet the real-time requirement. These models overcome the low accuracy problem of generalization recognition associated with the prevailing CNN models, which is caused by many trainable parameters and many layers of model. Experiments show that the recognition accuracy of these three models exceeds 90%. The highest recognition accuracy rate is 95%, and the recognition speed is 0.203 s/piece (256×256 pixels, CORE I5). Therefore, these three CNN models can be adopted in practice and have a strong feasibility and replicability; thus, they have a great potential to improve the production efficiency and product quality for chip resistors.

Keywords: Chip resistance recognition; convolution neural network; AlexNet model; GoogLeNet model; ResNet model

收稿日期: 2017-10-11. 网络出版日期: 2018-04-13.

基金项目: 四川省科技支撑计划项目 (2016GZ0107); 四川省教育厅重点项目 (16ZA0065); 南充市重点科技项目 (NC17SY4001).

通信作者: 何龙. E-mail: 396024902@qq.com.

当今正处于信息智能时代, 电子元器件犹如构筑起这个时代的一块块砖瓦, 唯有对这些砖瓦的质量进行保证, 才能坚固地铸就属于这个时代

的智能大厦,贴片电阻也不例外。工厂生产出来的贴片电阻,首先需要对其进行缺陷识别、极性方向识别、正反面识别和种类识别,才能保证贴片电阻的出厂质量,因此,生产企业迫切需要一种可行的贴片电阻识别检测方法,截至目前,虽然已经有论文提出针对贴片电阻缺陷进行识别的方法,但生产企业由于其识别性能原因依旧没有广泛的使用,贴片电阻在实际生产过程中的缺陷识别、方向识别等还是依靠人工肉眼进行识别检测,不仅检测速度慢,而且长期成本高、误检率高。

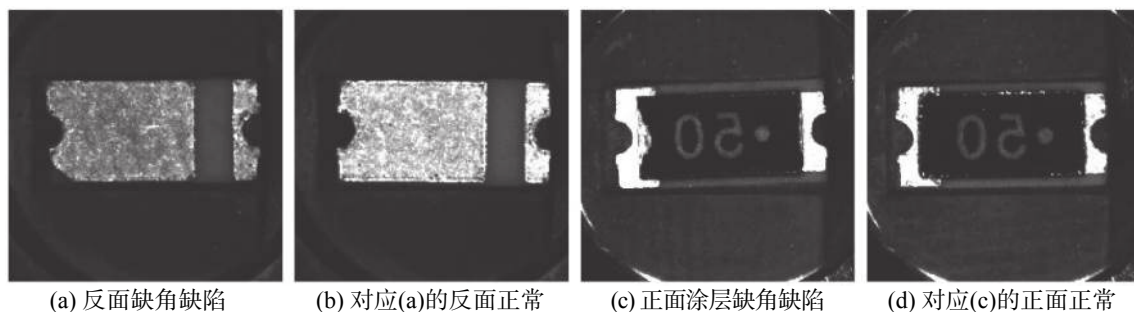
针对上述问题,学者们提出的很多理论方法均可在贴片电阻识别中进行应用,比如模板匹配算法^[1-2]、PCA(主成分分析)法^[3-4]、Canny 边缘检测算法^[5]、B 样条小波多尺度积边缘检测算法^[6]、自适应阈值 SUSAN 边缘算法^[7]等,通过使用这些方法首先对贴片电阻图片进行特征提取,然后将提取出的特征通过一个分类器进行识别分类,比如支持向量机^[8]、决策树^[9]、贝叶斯分类器^[10]等进行分类。上述过程从本质上来讲,就是通过人为的方式构造专家系统和知识库对贴片电阻图片进行特征抽取,将高度相关的数据编码形式解耦成易于分类的低维数据形式,随后将提取出的低维数据向量通过一个分类器进行识别分类。上述系统的识别性能很大程度上取决于前半部分专家系统的构造,即对特定论域问题是否提取出了有用的特征和是否将高维数据转换到易于分类的低维有效空间,通常要构造出一个专家知识系统是极其困难的,而且当研究论域一旦改变或者扩展,之前构造的专家系统又不得不重新进行构造,这样的再构造过程不仅是极其枯燥和乏味的,而且通常情况下其实际效果也并不是事先所预想的那样。

然而,上述问题可以通过卷积神经网络的方法进行特征提取,从而避免繁琐的专家知识系统的构造。神经网络早在 20 世纪 40 年代就已经被提出,其作为连接主义智能实现的典范,是当今人工智能深度学习方法的重要思想源泉,也是当前类脑智能研究中的有效工具^[11]。自 2006 年以来, Hinton 等在《科学》上发表文章将神经网络推向了深度学习时代^[12],指出一个训练好的多层神经网络可以将高维的数据转换成低维的数据编

码形式,提出一种 Pre-Training^[13]的方法初始化权重,使其更加接近全局最小,随后再通过梯度下降法来 Fine-Tuning 网络权重值,证明了采用这样的方法训练的网络比 PCA 方法更好;同时也证明了在参数相同的情况下,一个更深的网络结构在测试数据集上具有更低的错误识别率,但这种优势会随着权值参数的增多而逐渐消失。Hinton 所证明的结论为后续网络模型设计提供了理论依据。而 LeCun 所提出的卷积神经网络架构,是一个真正意义上的深度网络架构^[14],其不同于传统的全连接式的受限制波尔兹曼机,并在图像识别领域中取得了许多振奋人心的结果,通过端到端的训练,能够自学习出易于分类的低维特征,完全避免了人为专家知识系统的构造,与其相反所带来的代价是训练样本数据的收集。Hinton 所证明的结论对于真正意义的深度网络模型是有效的,对于全连接式的深度网络模型并不适用,全连接网络总是会陷入过拟合现象,使得实际泛化效果并不理想。因此,选择识别性能更优的卷积神经网络用于贴片电阻识别,充分借鉴前人思想设计卷积神经网络,尽可能精简整个卷积神经网络的可训练参数从而保证识别速度,基于此,对比研究得出了 3 种不同卷积神经网络结构在贴片电阻识别任务中的实际效果。

1 贴片电阻识别任务分析

从企业中调研可以发现,贴片电阻识别问题是一个复杂的多分类任务问题。图 1 所示贴片电阻图片来源于工厂实际拍摄,图中贴片电阻分类情况由工厂中经验丰富的工人所提供,图 1(a) 为 50 型号反面缺角缺陷,图 1(c) 为 50 型号正面涂层缺角缺陷,图 1(e) 为 50 型号反面腐蚀缺陷,图 1(g) 为正面字迹不清缺陷,其余 4 张图片为无缺陷电阻面。另一方面,假设贴片电阻已经做了外观缺陷检查,通过的合格电阻在自动编带时还需确保正反面、方向的一致性,因此需要对贴片电阻正反面和方向进行识别,由于企业生产的贴片电阻型号种类繁多,如图 2 所示,尺寸一样表示同一种封装形式的贴片电阻,故还需要对贴片电阻种类进行识别。



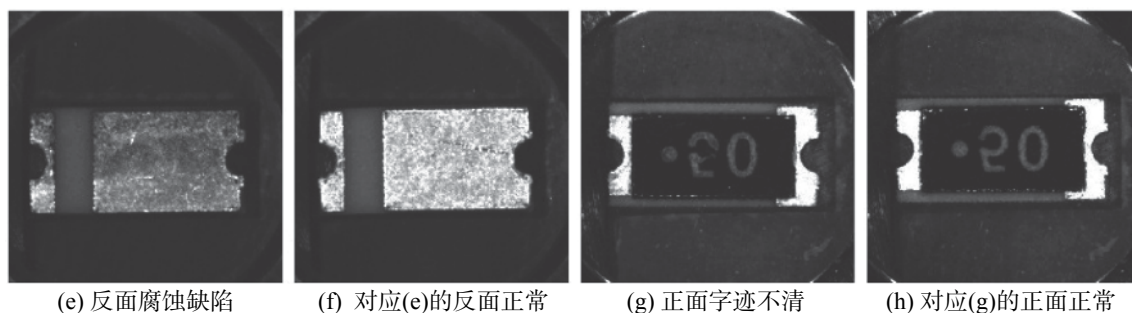


图 1 一种典型电阻型号需识别种类数

Fig. 1 A typical resistance type is required to identify the number of categories

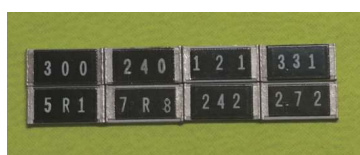


图 2 需要识别的不同型号的贴片电阻示意

Fig. 2 Different types of chip resistors that need to be identified

因此, 为了覆盖上述识别的所有情况, 通过在实际工厂中拍摄大量的贴片电阻图片, 再经过人工严格筛选和预处理算法处理之后, 本文制作出了包含 29 种类别的贴片电阻数据样本集。类别标签 1~13 主要是电阻缺陷识别, 14~29 主要是方向、正反面以及种类识别, 总的贴片电阻样本数量为 1 044 张, 每一类别包含 36 张样本, 经过预处理算法处理之后的图片, 每张图片被剪切为 320×320 像素大小, 贴片电阻居于图像正中位置, 贴片电阻以外其余的像素点全部被置零为黑色背景, 即清除了贴片电阻图像背景特征。在后续的实验, 每一类别均随机选取 5 张电阻图片用于准确率测试, 其余的贴片电阻均作为训练样本, 详细数据集及标签类别对应情况如图 3 和表 1 所示, 图 3 中相邻两张图片为一个类别, 以阅读的顺序与表 1 中的标签依次对应 (如图 3 中第一行前两张图片对应标签为 1, 所属类别为 50 型号反面弹珠缺陷)。

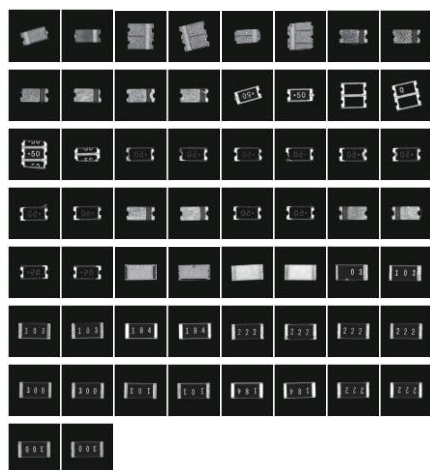


图 3 预处理后的 29 种贴片电阻图片样例

Fig. 3 Samples of 29 kinds of chip resistors after preprocessing

2 卷积神经网络相关数学理论

当前主流的 3 种用于模式识别的卷积神经网络模型——AlexNet 模型^[15]、GoogLeNet 模型^[16-17]、ResNet 模型^[18-19], 分别在 2012 年、2014 年、2015 年的 ILSVRC 竞赛中取得第一名的成绩, 其模型架构思想可以用图 4 进行概括, 图中卷积层实现特征提取, 全连接层实现特征分类。

2.1 卷积层模型

卷积神经网络模型首先由 LeCun 所提出, 其卷积过程类似于对图像抽取特征的各种算子, 如 Sobel 算子、Laplace 算子。算子的卷积过程:

$$O(m, n) = I(m, n) * H(m, n) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(i, j) H(m-i, n-j) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} I(m-i, n-j) H(i, j) \quad (1)$$

式中: I 表示输入灰度图像; H 表示图像操作算子; M 、 N 和 I 、 J 表示卷积核大小, 其操作过程为一固定大小的卷积核在图像像素点上进行了一次动态扫描。但卷积神经网络的卷积过程有其自己独有的特点, 引入了通道的概念, 则式 (1) 变为

$$O^l(m, n) = I^k(m, n) \otimes H^{kl}(m, n) = \sum_{k=0}^{K-1} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} I^k(m+i, n+j) H^{kl}(i, j) \quad (2)$$

式中: k 表示输入通道; l 表示输出通道。卷积神经网络在进行逐层特征抽取的过程中, 每个通道仍然按照二维卷积方式计算, 输入的多个通道与多个卷积核分别进行二维卷积, 得到多通道输出, 需要“合并”为一个通道, 则每层总的卷积核个数为 kl 个。通过卷积的方式, 卷积层在输出特征图维度实现了权值共享, 相比全连接的方式, 不仅极大减少了训练参数, 而且二维卷积核的卷积过程也符合图像的结构特征。

表1 实验详细数据表

Table 1 Experimental detailed data sheet

标签	训练样本数	测试样本数	种类名称
1	31	5	50 型号反面弹珠缺陷
2	31	5	50 型号反面两块整电阻缺陷
3	31	5	50 型号反面无规则裂角缺陷
4	31	5	50 型号反面腐蚀缺陷
5	31	5	50 型号反面缺角缺陷
6	31	5	50 型号反面桥接缺陷
7	31	5	50 型号正面弹珠缺陷
8	31	5	50 型号正面两块整电阻缺陷
9	31	5	50 型号正面无规则裂角缺陷
10	31	5	50 型号正面腐蚀缺陷
11	31	5	50 型号正面缺角缺陷
12	31	5	50 型号正面黑色涂层缺角
13	31	5	50 型号正面字迹不清
14	31	5	50 型号反面无缺陷方向朝东
15	31	5	50 型号正面无缺陷方向朝东
16	31	5	50 型号反面无缺陷方向朝西
17	31	5	50 型号正面无缺陷方向朝西
18	31	5	103,184,222,型号反面有缺陷
19	31	5	103,184,222,型号反面无缺陷
20	31	5	103 型号正面有缺陷
21	31	5	103 型号正面无缺陷
22	31	5	184 型号正面无缺陷
23	31	5	222 型号正面电极不对称缺陷
24	31	5	222 型号正面无缺陷
25	31	5	300 型号正面无缺陷
26	31	5	103 型号正面无缺陷字迹反向
27	31	5	184 型号正面无缺陷字迹反向
28	31	5	222 型号正面无缺陷字迹反向
29	31	5	300 型号正面无缺陷字迹反向

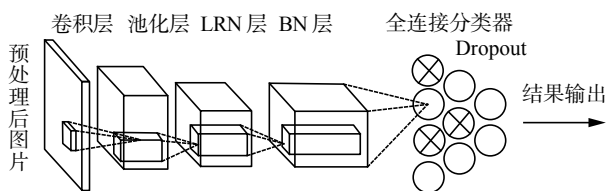


图4 常见卷积模型图

Fig. 4 Common convolution model diagram

2.2 ReLU 非线性函数

根据神经科学的研究,存在一个阈值来控制神经元的选择性输出,从而增加整个神经系统的

非线性特性和稀疏特性。数学家们将这样的阈值函数称为激活函数,早期的激活函数主要有 Sigmoid 函数和 tanh 函数两种,但二者均由于使梯度在反向传播过程中衰减过快而被人们所放弃,过快的梯度衰减与构造一个更深的卷积神经网络是矛盾的。因此人们提出了 ReLU 非线性激活函数^[20],其表达式为

$$f(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (3)$$

式(3)的激活函数不仅能够减少计算量,在一定程度上克服了梯度反向传播消失问题,并且增加了网络的稀疏性,因此而被广泛的使用。

2.3 Dropout(随机丢弃)层

Dropout 层是为了增加神经网络的稀疏性而特别设计的,其背后的理论依据是为了模仿大脑的稀疏特性。人脑的神经元大多处于抑制状态,只有不到 10% 的神经元处于活动状态,因此,人脑是一个典型的稀疏结构,每次训练网络通过选择性的丢弃某些神经元,使其不参与前向传播和反向传播过程,一方面可以增加网络的稀疏性,另一方面提供了正则化一大类模型的方法。

2.4 Pooling(池化)层

Pooling 层是为了快速降低特征图维度而设计的,从而减轻计算的负担。Pooling 层通过汇总上一层特征图相邻神经元的输出,输出一个维度被降低的特征图。通过 Pooling 层的操作,模型计算量被减少,网络过拟合问题得到一定的缓解,网络整体性能也有所提升。常见的汇总操作有 MAX 和 AVE, MAX 操作为局部相邻神经元的最大值作为输出,AVE 操作为局部相邻神经元的平均值作为输出,与图像均值滤波类似。

2.5 LRN(local response normalization)层

LRN 层为局部响应值归一化层,LRN 层跟随在 Pooling 层之后,通过对局部神经元的活动创建竞争机制,使得其中响应比较大的值变得相对更大,并抑制其他反馈较小的神经元,保留关键特征,抑制非重要特征^[15]。LRN 层具体归一化公式如式(4)所示:

$$b_{x,y}^i = a_{x,y}^i / \left(k + \alpha \sum_{j=\max(0,i-n/2)}^{\min(N-1,i+n/2)} (a_{x,y}^j)^2 \right)^\beta \quad (4)$$

式中: $a_{x,y}^i$ 表示第 i 个特征图像素位置为 x, y 的值; N 表示特征图的总个数; n 表示沿着特征图(也称通道)方向需要被归一化的像素点个数; k, α, β 为可自由设置的超参数。求和函数中的 min 和 max 用于避免数据溢出特征图个数(通道个数)之外。

2.6 BatchNorm 层

BatchNorm 层为步长归一化层,卷积神经网络在读入训练数据和测试数据时,并不是每次只读取一张图片,而是一次读取一个 Batch(步长)值

的图片数据。这样就带来一个问题, 深度神经网络每一层的输入分布在训练时会发生改变, 导致卷积神经网络难以学习到数据中蕴藏的本质, 这使得要训练这样的网络变得困难复杂, 而这样的问题可以通过归一化每一层的输入来解决。BN 的本质作用原理是强行将每一层输入分布的变化进行矫正, 使得每一层输入在进行卷积之前保持零均值和单位方差^[19]。使用 BN 层的深度网络结构, 可以使用更高的学习率进行训练, 而且也不用过多的关心训练权值初始化问题。实践证明加入 BN 层的深度网络不仅训练收敛速度极大提高, 而且识别准确率也得到改善。

3 卷积模型架构设计

Hinton 证明了在参数相同的情况下, 一个更深的网络结构在测试数据集上具有更低的错误识别率, 但这种优势会随着权值参数的增多而逐渐消失, 本文基于此理论依据, 针对传统图像识别方法的局限性, 结合 AlexNet 模型、GoogLeNet 模型、ResNet 模型思想和其突出的图像识别分类能力, 设计了 3 种卷积神经网络模型用于识别贴片电阻图像, 需同时保证设计模型的识别准确率和识别速度。

起初 AlexNet 模型被设计的目的是用于参加 ILSVRC 图像分类竞赛, 同时消除人们对卷积神经网络识别性能的质疑。网络总共包含 60M 的可训练参数, 主要是为了提升网络容量, 应对包含千万张图像的 ImageNet 数据集, 但贴片电阻识别并不需要如此巨大的可训练参数, 这是因为贴片电阻数据并不像 Imagenet 数据集那样蕴含大量的特征, 并且在实践中企业迫切希望使用少量的训练样本去达到非常理想的识别准确率, 因此需要对已有模型可训练参数进行减少, 防止过拟合现象。同时, 贴片电阻识别既需要保证高识别准确率, 也需要保证高识别速度, 这就需要将卷积神经网络的深度控制在一个合理的范围之内。因此, 设计的网络需从整体上进行“瘦身”, 同理, 另外 2 种模型的修改思路亦是如此, 将修改后的 3 种模型可训练参数保持一致, 可以方便比较他们的综合性能。

基于上述思想, 由 AlexNet 模型思想修改后的网络依旧为 8 层深度, 结构与原来保持不变, 做出的改动是对原模型可训练参数进行了大量精简, 可训练参数从原来的 60×10^6 降低为 4×10^6 左右, 训练参数只有原来的 1/15, 详细参数设置如图 5 所示, 图中参数设置含义为, 比如 (11×11 卷积, 48, /4) 表示卷积核大小为 11×11, 输出特征图个数为 48, 卷积时滑动步长为 4, 没有写滑动步长表示默认为 1, 其他同理, 后续网络中参数设置含义与此类似。

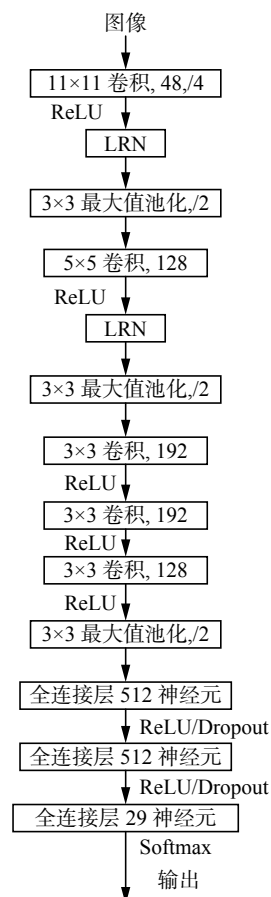


图 5 识别贴片电阻的 8 层经典卷积网络结构, 可训练参数为 4.07×10^6

Fig. 5 The 8-layer classical convolution network structure of identify the chip resistors with a training parameter of 4.07×10^6

基于谷歌 Inception 结构思想设计的 16 层 4-Inception 网络如图 6 所示, 模型中将两个 Inception 结构组合构成一个基本的结构单元, 取名为 4-Inception(参见图 7), 这样可以方便网络结构呈现。在 4-Inception 的卷积结构当中, 卷积核大小全部使用 1×1、3×3、5×5, 参数设计思想是随着深度的增加, 提取出的特征越来越抽象, 每个特征所涉及的感受野 (receptive field) 越来越大。故网络越到后面, 3×3、5×5 卷积核比例相应增加, 但卷积核的增加会带来巨大的训练参数和计算量, 因此在进行卷积之前先进行 1×1 卷积降低特征图 (通道) 个数。为了克服连接器拼接特征图维度对齐问题, 1×1、3×3、5×5 卷积时要分别设定填充大小为 0、1、2, 则卷积后的尺寸保持不变, 更易对齐。采用此结构设计的 16 层深度卷积网络, 由于 4-Inception 基本结构单元比较复杂, 需要利用多尺度的训练方法对其进行训练, 从而提高网络训练收敛速度。图 7 中 3 个 4-Inception 结构参数详细设置见表 2。

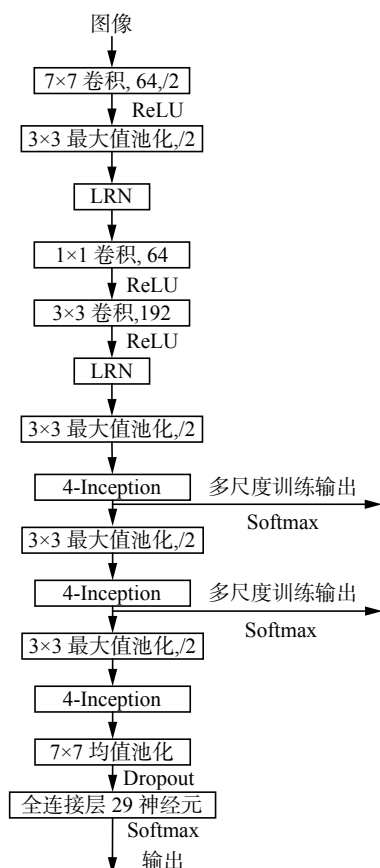


图6 识别贴片电阻的16层深度4-Inception模型图,可训练参数为 3.8×10^6

Fig. 6 The 16-layer depth 4-Inception model of identify the chip resistors with a training parameter of 3.8×10^6

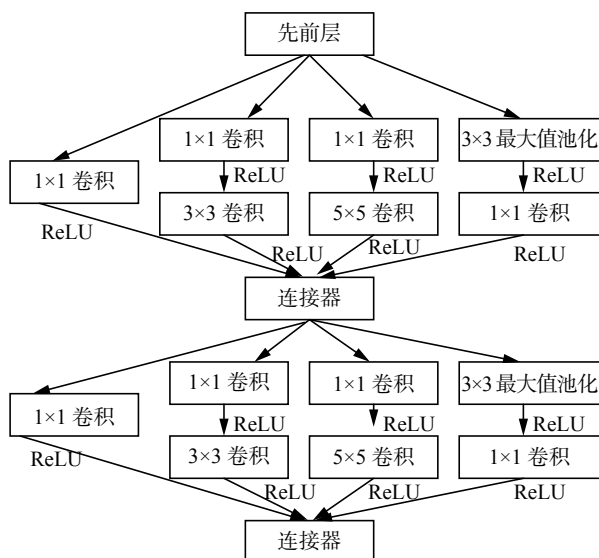


图7 谷歌Inception卷积结构(4-Inception)

Fig. 7 Google Inception convolution structure (4-Inception)

基于残差思想设计的16层4-ResNet网络如图8所示,图中4-ResNet详细结构如图9所示,该结构的主要思想是为了加深网络结构,克服梯度反向传播过程中梯度消失问题。对比本文16层深度4-Inception结构需要采用多尺度训练的方法,

虽然能够很好地训练出网络模型,但同样深度的残差网络结构,直接采用端到端的训练方式,即可使网络快速收敛并取得很好的训练结果。从图9中很容易看出,梯度在反向传播的时候,求和后的函数的导数为

$$\frac{\partial(f(x)+x)}{\partial x} = \frac{\partial f(x)}{\partial x} + 1 \quad (5)$$

表2 4-Inception详细参数设置
Table 2 4-Inception detailed parameter settings

Inception-4/1	上半参数	下半参数
1×1 卷积	64	128
1×1 卷积	96	128
3×3 卷积	128	192
1×1 卷积	16	32
5×5 卷积	32	96
3×3 池化	无	无
1×1 卷积	32	64
Inception-4/2	上半参数	下半参数
1×1 卷积	192	160
1×1 卷积	96	224
3×3 卷积	208	224
1×1 卷积	16	24
5×5 卷积	48	64
3×3 池化	无	无
1×1 卷积	64	64
Inception-4/3	上半参数	下半参数
1×1 卷积	256	384
1×1 卷积	160	192
3×3 卷积	320	384
1×1 卷积	32	48
5×5 卷积	128	128
3×3 池化	无	无
1×1 卷积	128	128

导数为1不会对梯度有任何的衰减,且直接跨越多个卷积层进行反向传播,和多尺度进行训练的方法有着异曲同工之妙,这就是残差网络能够快速收敛并取得良好训练结果的原因。图9中输入引出的1×1卷积可以使得第一次求和的特征图维度相同,残差结构的本质精髓就在于图9中所示的圆弧线,其能同时跨越多个(不一定是两个)

卷积层。从图 8 中可以详细的看出模型的参数设置, 4-ResNet 结构中的卷积核参数全部设置成 148。

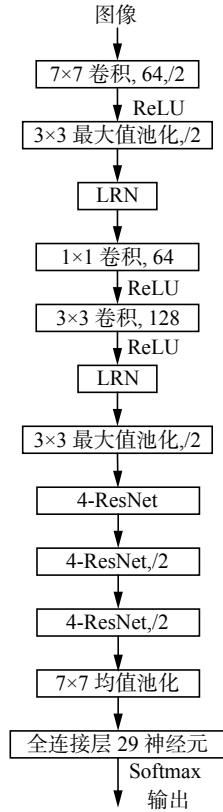


图 8 识别贴片电阻的 16 层深度 4-ResNet 模型, 可训练参数为 4.0×10^6

Fig. 8 The 16-layer depth 4-ResNet model of identify the chip resistors with a training parameter of 4.0×10^6

3 种模型的输出神经元均被设置成 29, 这是由实验数据集包含 29 个类别所导致的, 训练时均采用随机梯度下降法进行训练, 选择 Softmax 回归模型对网络进行优化, 权值更新公式为

$$v_{i+1} = 0.9 \times v_i - 0.0005 \times \varepsilon \times w_i - \varepsilon \times \left\langle \frac{\partial L}{\partial w} \right\rangle_{D_i} \quad (6)$$

$$w_{i+1} = w_i + v_{i+1}$$

式中: i 是迭代次数; v 是运动常量; ε 是学习率; $\left\langle \frac{\partial L}{\partial w} \right\rangle_{D_i}$ 是损失函数相对于 w 的偏导在第 i 次步长输入数据 D_i 上的平均。

4 实验及讨论

4.1 贴片电阻识别指标定义

假设贴片电阻存在 m 个类别, 每种类别训练样本数为 n_1, n_2, \dots, n_m , 测试样本数为 t_1, t_2, \dots, t_m , 识别错误数为 e_1, e_2, \dots, e_m , 则易得总的训练样本数为

$$N = \sum_{i=1}^m n_i \quad (7)$$

总的测试样本数为

$$T = \sum_{i=1}^m t_i \quad (8)$$

总的识别错误数为

$$E = \sum_{i=1}^m e_i \quad (9)$$

因此, 总的贴片电阻识别准确率为

$$\text{acc_all} = \left(1 - \frac{E}{T}\right) \times 100\% \quad (10)$$

任意贴片电阻类别的总识别错误率为

$$\text{err_all} = \frac{e_i}{T} \times 100\% \quad (11)$$

任意贴片电阻类别 i 的识别准确率为

$$\text{acc_i} = \left(1 - \frac{e_i}{t_i}\right) \times 100\% \quad (12)$$

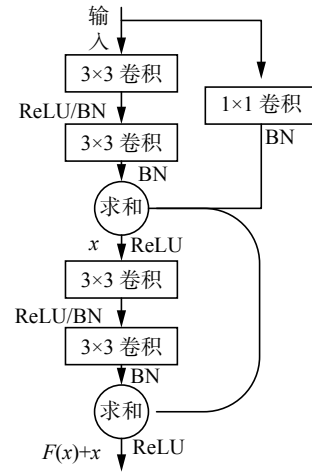


图 9 残差网络结构, 取名 4-ResNet

Fig. 9 Residual network structure, named 4-ResNet

4.2 实验过程

在贴片电阻姿态固定而位置随机情况下, 设计的 3 种卷积神经网络模型对贴片电阻进行识别的准确率分析, 实际中贴片电阻会被限制在一个导轨上进行运动, 所拍摄的贴片电阻位置并不会都严格居于图像正中间, 因此对贴片电阻随机位置的识别是与实际情况更加吻合的。

为了模拟贴片电阻在图像中位置随机的情况, 每次迭代训练都进行随机截取 256 像素 \times 256 像素大小图片输入网络进行训练。原始数据集中每张图片为 320 像素 \times 320 像素大小是事先设计好的, 这样总可以保证每次随机截取的 256 像素 \times 256 像素大小图片完全将贴片电阻信息包含在图像中。

整个实验过程基于 Caffe 深度学习框架, 计算 GPU 显卡为 GeForce GTX 1060 6 GB。执行训练时, 训练步长 (batch) 为 32 张图片, 迭代 4 次输出一次训练损失值, 每训练 20 次执行一次准确率测试。测试步长 (batch) 为 16 张图片, 测试迭代次数为 10 次。网络训练动量因子为 0.9, 初始学

习率为 0.01, 权重衰减因子为 0.000 5, 学习率更新策略为默认方法。总的训练次数为 7 200 次, 即训练周期约为 240 个 (epoch)。

4.3 实验结果分析

3 种模型的训练结果分别如图 10~11 所示, 可以看出, 由于贴片电阻位置的随机性, 识别准确率最高的为 16 层 4-Inception 网络, 识别准确率达到 95%, 这说明层数更深的 4-Inception 网络更加擅长处理复杂情况下的图像识别, 特征解耦更好, 更适合贴片电阻识别分类任务。残差网络是收敛最快的网络结构, 但其测试损失波动非常大并在迭代 2 500 次后趋于平稳, 继续训练则是过拟合训练, 因此实验时训练迭代次数未到 7 200 次则将其停止。

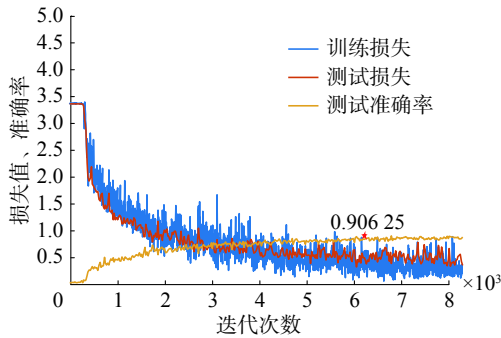
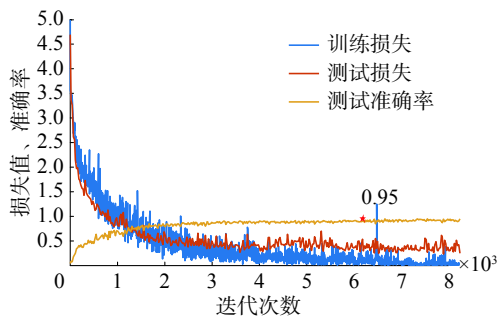
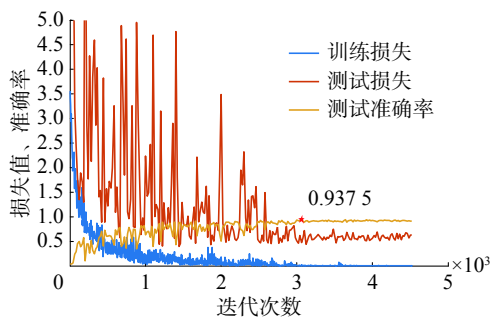


图 10 贴片电阻位置随机 8 层深度卷积模型训练结果

Fig. 10 Chip resistor position random 8-layer depth convolution model training results



(a) 贴片电阻位置随机 16 层 4-Inception 网络



(b) 贴片电阻位置随机 16 层 4-ResNet 网络

图 11 网络训练结果

Fig. 11 Network training results

3 种模型性能详细对比如表 3 所示, 识别时间最快的是 8 层深度卷积网络, 低至 0.203 秒/张 (256×256 像素/COREI5), 识别准确率最高的是 16 层 4-Inception 网络, 识别准确率达 95%, 残差网络性能折中。

表 3 3 种模型识别时间及准确率对比

Table 3 Comparison of 3 model identification times and accuracy

测试图片 大小/平台	8 层深度 卷积网络	16 层 4- Inception 网络	16 层 4- ResNet 网络
256 像素×256 像素/ COREI5	0.203 秒/张	0.556 秒/张	0.337 秒/张
总识别准确率/%	90.63	95	93.75

3 种模型的泛化识别错误率分布如图 12~13 所示, 可以看出, 标签 1~13 主要属于缺陷类别识别, 是识别错误率的主要来源。标签 14~29 主要识别正反面、方向和种类, 3 种模型均取得了非常理想的识别效果。图 13(a) 中的 16 层 4-Inception 模型, 其仅在缺陷识别上存在一定的错误率, 在种类、正反面和方向识别上取得了 100% 的识别准确率。

综上所述, 当贴片电阻在图像中位置随机时, 基于 AlexNet 思想“瘦身”后的 8 层深度卷积神经网络识别时间最短, 16 层 4-Inception 网络准确性更加突出, 进一步提高识别准确率可适当增加缺陷样本的数量。

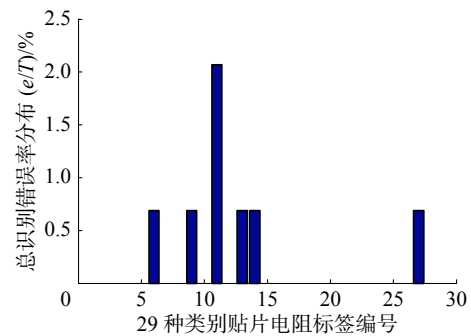
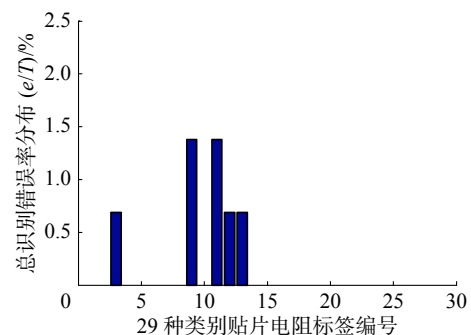


图 12 8 层深度卷积网络总识别错误率分布

Fig. 12 Total error rate distribution of 8-layer deep convolutional networks



(a) 16 层深度 4-Inception 网络

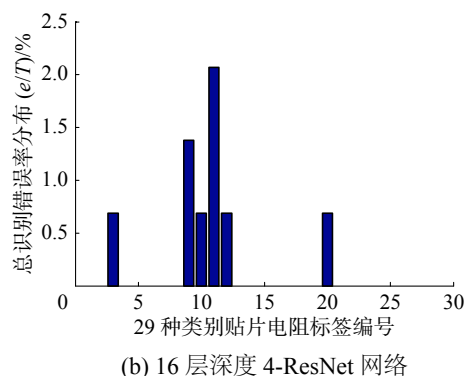


图 13 网络总识别错误分布

Fig. 13 Networks total error rate distribution

5 结束语

本文基于 AlexNet 模型、GoogLeNet 模型、ResNet 模型思想设计了 3 种卷积神经网络模型, 用于解决实际工业生产中贴片电阻识别问题, 包括缺陷识别、方向识别、正反面识别和种类识别。实验表明, 设计的 3 种卷积神经网络模型中, 基于 AlexNet 模型思想“瘦身”后的卷积神经网络识别时间最短, 识别时间低至 0.203 秒/张 (256×256 像素, CORE I5), 16 层 4-Inception 网络准确性更加突出, 总识别准确率达 95%; 16 层 4-ResNet 网络性能折中, 克服了当前主流卷积模型由于可训练参数过多、模型层数太深导致在贴片电阻识别应用中识别速度不能满足实时性要求、泛化识别准确率低的问题。上述 3 种卷积模型可根据具体实际需求进行选用, 有望解决工业界贴片电阻分类依靠人工方式的难题。

参考文献:

- [1] GIUNCHIGLIA F, YATSKEVICH M. Element level semantic matching. Technical Report DIT-04-035[R]. Trento, Italy: Information Engineering and Computer Science, 2004.
- [2] OMACHI M, OMACHI S. Fast two-dimensional template matching with fixed aspect ratio based on polynomial approximation[C]//Proceedings of the 9th International Conference on Signal Processing. Beijing, China, 2008: 757-760.
- [3] JOLLIFFE I T. Principal component analysis[M]. New York, USA: Springer-Verlag, 1986.
- [4] ZHANG Xingfu, REN Xiangmin. Two dimensional principal component analysis based independent component analysis for face recognition[C]//Proceedings of 2011 International Conference on Multimedia Technology. Hangzhou, China, 2011: 934-936.
- [5] PELLEGRINO F A, VANZELLA W, TORRE V. Edge detection revisited[J]. *IEEE transactions on systems, man, and cybernetics, part b (cybernetics)*, 2004, 34(3): 1500-1518.
- [6] DUCOTTET C, FOURNEL T, BARAT C. Scale-adaptive detection and local characterization of edges based on wavelet transform[J]. *Signal processing*, 2004, 84(11): 2115-2137.
- [7] SMITH S M, BRADY J M. SUSAN-a new approach to low level image processing[J]. *International journal of computer vision*, 1997, 23(1): 45-78.
- [8] CORTES C, VAPNIK V. Support-vector networks[J]. *Machine learning*, 1995, 20(3): 273-297.
- [9] GUO H, GELFAND S B. Classification trees with neural network feature extraction[J]. *IEEE transactions on neural networks*, 1992, 3(6): 923-33.
- [10] GOLDSZMIDT M. Bayesian network classifiers[J]. *Machine learning*, 2011, 29(2/3): 598-605.
- [11] 焦李成, 杨淑媛, 刘芳, 等. 神经网络七十年: 回顾与展望[J]. *计算机学报*, 2016, 39(8): 1697-1717.
- [12] JIAO Licheng, YANG Shuyuan, LIU Fang, et al. Seventy years beyond neural networks: retrospect and prospect[J]. *Chinese journal of computers*, 2016, 39(8): 1697-1717.
- [13] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [14] ERHAN D, MANZAGOL P A, BENGIO Y, et al. The difficulty of training deep architectures and the effect of unsupervised pre-training[C]//Appearing in Proceedings of the 12th International Conference on Artificial Intelligence and Statistics. Florida, USA, 2009: 153-160.
- [15] LECUN Y, BENGIO Y. Convolutional networks for images, speech, and time series[M]//ARBIB M A. The Handbook of Brain Theory and Neural Networks. Cambridge, MA, USA: MIT Press, 1998.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems. Montreal, Canada, 2012: 1097-1105.
- [17] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolutions[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA, 2014: 1-9.
- [18] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[J]. arXiv: 1602.07261, 2016.
- [19] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 770-778.
- [20] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on Machine Learning. Lille, France, 2015: 448-456.

- [20] GLOROT X, BORDES A, BENGIO Y. Deep sparse rectifier neural networks[C]//Proceedings of the 14th International Conference on Artificial Intelligence and Statistics. Fort Lauderdale, FL, USA, 2011: 315–323.

作者简介:



湛贵辉, 男, 1971 年生, 教授, 主要研究方向为 MEMS 集成器件及传感器、智能仪表、计算机仿真及模拟技术及图像处理及模式识别技术。



何龙, 男, 1991 年生, 硕士研究生, 主要研究方向为智能控制、模式识别。



李忠兵, 男, 1987 年生, 博士, 主要研究方向为图像处理、精密仪器及现代信号处理。

第 15 届中国智能系统会议 (CISC2019) 征文通知

中国智能系统会议是由中国人工智能学会智能空天系统专业委员会发起的系列学术会议, 其宗旨是为本领域的专家学者、研究生以及工程技术人员提供一个学术交流的平台, 以推动我国智能系统相关理论、技术与应用的发展。第 15 届中国智能系统会议 (CISC2019) 将于 2019 年 10 月 26—27 日在海南省海口市召开。

本次会议由中国人工智能学会主办, 中国人工智能学会智能空天系统专业委员会协办, 海南大学与北京航空航天大学联合承办。会议论文集将由 Springer 出版社在 Lecture Notes in Electrical Engineering 系列正式出版, EI 收录。热忱欢迎海内外广大同仁踊跃投稿并出席本届会议, 交流学术成果。

投稿要求:

1. 论文未曾在国内外杂志或会议上发表;
2. 稿件写作必须使用英文, 并严格按照模板要求进行排版;
3. 所有论文采用网上投稿, 投稿系统网址为 <https://easychair.org/conferences/?conf=cisc2019>

如有疑问请邮件咨询 cisc2019@126.com

论文模板:

会议提供 LaTeX 和 Microsoft Word 模 (分别见附件 2 和附件 3)。我们强烈建议您的论文使用会议模板进行编排。您也可以不使用这些模板, 但必须严格按照模板中的规范编排您的论文。如果您使用 MS Word 2007 及以上版本编排论文, 请使用 MS Word 2003 兼容格式。

重要日期:

论文投稿截止日期: 2019 年 4 月 30 日。

论文录用通知日期: 2019 年 6 月 20 日。

会议注册/终稿提交截止日期: 2019 年 7 月 10 日。