

基于深度学习的极性电子元器件目标检测与方向识别方法

陈文帅^{1,2} 任志刚^{1,2,3,4} 吴宗泽^{1,2,3,4} 付敏跃^{1,5}

摘要 极性电子元器件的类别、方向识别、定位在工业生产、焊接和检测等领域发挥着至关重要的作用. 本文首先将极性电子元器件的方向识别问题转化为一个分类问题, 然后, 采用 Faster RCNN (Region convolutional neural network) 与 YOLOv3 方法实现了极性电子元器件的准确分类、方向识别和精准定位. 实验取得良好的效果, 两种算法的平均准确率 (Mean average precision, mAP) 分别达到 97.05%、99.22%. 此外, 我们通过数据集目标框的长宽分布, 利用 K-means 算法对 Faster RCNN 和 YOLOv3 的 Anchor boxes 进行了改进设计, 使准确率分别提高了 1.16%、0.1%, 并提出针对大目标检测的网络结构: YOLOv3-BigObject, 在提高准确率的同时, 将检测单张图片的时间大幅缩减为原来检测时间的一半, 并最终用焊接有元器件的电路板进行检测, 得到了很好的实验结果.

关键词 电子制造, 深度学习, 方向识别, 目标检测, Faster RCNN

引用格式 陈文帅, 任志刚, 吴宗泽, 付敏跃. 基于深度学习的极性电子元器件目标检测与方向识别方法. 自动化学报, 2021, 47(7): 1701–1709

DOI 10.16383/j.aas.c190037

Detecting Object and Direction for Polar Electronic Components via Deep Learning

CHEN Wen-Shuai^{1,2} REN Zhi-Gang^{1,2,3,4} WU Zong-Ze^{1,2,3,4} FU Min-Yue^{1,5}

Abstract The category, direction identification and positioning of polar electronic components play an important role in the industrial production, welding and inspection. In this paper, we first successfully transform the original problem of directional identification of polar electronic components into a classification problem. Then, the Faster RCNN (region convolutional neural network) and YOLOv3 methods are used to realize the correct classification, direction identification and accurate positioning of the polar electronic components. The experiments validate the effectiveness of our proposed method and the mAP (mean average precision) of the two proposed algorithms can reach 97.05%, 99.22%. In addition, we improve the anchor boxes of the Faster RCNN and YOLOv3 by K-means algorithm through the length and width distributions of the target frames of the datasets, the accuracy can be improved by 1.16%, 0.1%. We also propose the YOLOv3-BigObject network structure for the large target detection, while improving the accuracy, the cost time for detecting a single picture is also greatly reduced. Finally, the board with the electronic components is chosen to test and good experimental results are obtained.

Key words Electronic manufacturing, deep learning, direction recognition, object detection, Faster RCNN

Citation Chen Wen-Shuai, Ren Zhi-Gang, Wu Zong-Ze, Fu Min-Yue. Detecting object and direction for polar electronic components via deep learning. *Acta Automatica Sinica*, 2021, 47(7): 1701–1709

收稿日期 2019-01-15 录用日期 2019-04-30
Manuscript received January 15, 2019; accepted April 30, 2019
国家自然科学基金 (61673126, 61703114, U1701261, U1701264)
资助

Supported by National Natural Science Foundation of China (61673126, 61703114, U1701261, U1701264)

本文责任编辑 金连文

Recommended by Associate Editor JIN Lian-Wen

1. 广东工业大学自动化学院 广州 510006 中国 2. 广东省物联网信息技术重点实验室 广州 510006 中国 3. 精密电子制造技术与装备国家重点实验室 广州 510006 中国 4. 广东省离散制造知识自动化工程技术研究中心 广州 510006 中国 5. 纽卡斯尔大学电气工程与计算机科学学院 纽卡斯尔 2308 澳大利亚

1. School of Automation, Guangdong University of Technology, Guangzhou 510006, China 2. Guangdong Key Laboratory of IoT Information Technology, Guangzhou 510006, China 3. State Key Laboratory of Precision Electronic Manufacturing Technology and Equipment, Guangzhou 510006, China 4. Guangdong Discrete Manufacturing Knowledge Automation

电子元器件是电子设备的重要基础组成部分, 在其生产、回收、焊接与检测中, 首先需要对元器件进行正确分类. 此外, 在焊接和检测元器件的过程中, 还需要对极性电子元器件的方向进行准确识别. 因此, 对电子元器件正确分类的同时能够对极性电子元器件的方向进行识别具有重要的研究意义^[1]. 但随着电子元器件制造工艺的提高与需求的多样化, 电子元器件变得更加微型化、多样化, 这为电子元器件的分类与检测方向等问题带来了更严峻的挑战.

2012 年之前, 图像分类和目标检测一度进入瓶颈期, 发展相对停滞. 随着计算机视觉的快速发展,

Engineering Technology Research Center, Guangzhou 510006, China 5. School of Electrical Engineering and Computer Science, the University of Newcastle, Newcastle 2308, Australia

深度学习在目标检测方面也取得了巨大的成功. 得益于此, 深度学习在计算机视觉方向的算法不断更新迭代, 在文字识别^[2]、医学图像识别^[3-4]等方面得到了广泛运用. 2012 年, Krizhevsky 等^[5]提出了一个称为 AlexNet 的深度卷积神经网络 (Deep convolutional neural network, DCNN), 其在大规模的视觉识别挑战 (ImageNet large-scale visual recognition challenge, ILSVRC) 中实现了破纪录的图像分类精度, 成为深度学习发展的一个重要里程碑. 从此, 深度学习进入高速发展阶段, 许多计算机视觉应用领域的研究也都集中在深度学习方法上^[6], 其中包括目标检测及图片背景消减技术^[7].

深度学习在目标检测领域的目标检测方法^[8]可分为基于候选区域的目标检测方法和基于回归的目标检测算法. 基于候选区域的检测方法采用了卷积神经网络代替人工特征的提取^[9], 常用的候选区域提取办法有 Selective Search^[10] 和 Edge Boxes^[11]. 常用来提取特征图的网络结构有 VGG (Visual geometry group)^[12]、ResNet^[13] 等. Faster RCNN (Region convolutional neural network)^[14] 是最具有代表性的算法之一, Faster RCNN 为了减少 Selective Search 和 Edge Boxes 提取候选区域的时间, 结合全卷积设计的候选区域思想^[15], 提出了用区域候选网络 (Region proposal networks, RPN) 计算候选框, 不仅提高了候选区域的质量与准确率, 还将检测时间减少为原来的 1/10. 基于候选区域的目标检测方法需要分两步去分别训练两个复杂的卷积结构, 虽然检测精确, 但其有在训练与检测速度上缓慢的缺点. 基于回归的目标检测算法为了解决这个缺点, 不通过生成区域建议的方式, 直接将目标检测的相关问题转化为回归问题^[16], 其结构相对简单, 实时性更好, 最具有代表性的算法有 SSD (Single shot multibox detector)^[17] 和 YOLO (You only look once) 系列算法^[18-20]. YOLO 算法简单地将图像均匀分为 7×7 个部分, 再利用卷积神经网络检测每一个部分是否有目标, 最后对目标类别与位置进行预测. YOLO 就是通过节省生成区域建议的时间来大幅提高检测速度, 但这也会对其精度带来一定的影响.

目前对于电子元器件分类研究的算法, 主要分为两大类, 如图 1 所示, 一类是基于传统的图像分类的方法, 首先输入图像、然后经过复杂的程序提取图像的特征, 再根据提取出来的特征进行分类. 如杜思思等采取将最小二乘法和 Hough 变换相结合^[21], 得到一种改进的直线段精确检测算法. 另外一类是基于深度学习的分类方法, 首先输入图像、然后经过卷积神经网络自动提取特征生成特征图、再根据特征图进行目标的定位与分类. 如陈翔等^[22]提出了一种

基于卷积神经网络 (AlexNet) 的电子元器件分类方法, 其可以直接在输入图像上进行自动特征学习和分类, 弥补了传统元件分类方法存在部分缺陷, 但只能对图片进行分类, 不能运用到单幅有多目标的场景并对目标进行定位. 第 1 类传统的图像分类方法需要提取图像的多种特征, 要对每一种类别的元器件找出其表示的特征, 如蜂鸣器的特征用圆表示等. 其有以下缺点: 1) 传统图像分类方法对于图像的对比度要求很高, 模糊、灰度化不明显、灰度不均匀等情况都会增加识别的误差; 2) 对特征相似的目标识别效果不佳, 比如同样类别的不同型号的元器件分辨准确率低; 3) 需要人工提取特征, 无法对包含上百种海量数据进行分类, 每多一种分类, 计算复杂度就上升一个等级. 第 2 类基于深度学习的分类算法, 虽然有随着深度学习网络的加深计算量不断加大的缺点, 但融入了反向传播^[23]、权值共享、稀疏连接、池化、Dropout^[24] 等思想, 简化了操作, 并且大大降低了神经网络参数的量级, 一定程度上减轻了计算量, 使其更易于训练, 而且深度学习算法能够自动的提取图像特征, 自动学习各种各样的表征特征纹理, 甚至能够表达完整的目标物, 比人工提取的特征更具有描述力, 其对于目标的理解的概念分布在神经网络不同的层中, 最后形成了对不通目标物体的整体感知, 并根据感知同时对目标进行分类与定位.

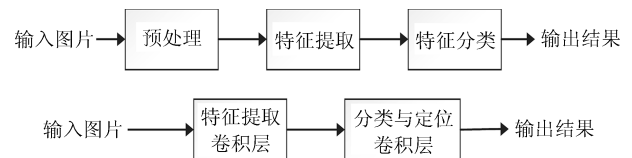


图 1 基于传统和深度学习图像分类方法模型

Fig. 1 Traditional and deep learning image classification method based on model

本文的主要贡献主要有以下三个方面: 1) 首次提出基于深度学习极性电子元器件类别与方向识别方法. 通过把方向识别问题转化为分类问题, 利用先进的目标检测算法进行分类, 平均准确率 (Mean average precision, mAP) 最优高达 99.44%. 2) 提出一种基于数据集设计遍历框 (Anchor boxes) 的方法. 根据训练集目标框长宽的分布结合算法提供的框来选取 Anchor boxes 的个数与大小. 3) 提出 YOLOv3-BigObject 算法模型. 对于大目标的数据集, 通过对 Anchor boxes 大小与个数的分析, 压缩 YOLOv3 的算法框架, 减少了训练复杂度与时间, 大幅减少对单张图片的检测时间.

本文组织架构如下: 首先, 在第 1 节对数据集的构建与内容进行了详细介绍. 在第 2 节利用 K-means 算法, 基于数据集和目标检测算法设计了

Anchor boxes 的大小与个数. 在第 3 节提出用于大目标检测的 YOLOv3-BigObject 网络模型. 在第 4 节, 我们通过进一步实验分析, 验证了深度学习在该场景下的极性元器件方向识别的效果, 分析了用 K-means 算法设计 Anchor boxes 的效果以及 YOLOv3-BigObject 模型在检测时间上的提升效果, 并用焊接有元器件电路板视频检测算法实际效果, 最后, 对全文工作进行了总结和展望.

1 数据集构建及预处理

深度学习的算法模型需要大量的数据进行训练, 但目前还没有公开的、带类别标签的电子元件的数据集, 为了能够实现本文所提出的基于深度学习的极性电子元器件类别与方向识别, 实验采取人工拍摄的方式采集数据集. 采集设备为海康威视 MV-

CE013-50 GM 130 万像素的 CCD 工业相机, 数据集 1 一共采集 2000 幅大小为 1280×960 像素、包含 3 种型号的二极管图片, 如图 2 所示. 数据集 2 一共采集 813 幅大小为 1280×960 像素、包含 2 种型号且带有引脚的二极管图片, 如图 3 所示. 目前深度学习算法还没有能够识别目标的具体倾斜角度, 但可以通过把方向问题转换为分类问题达到粗略的方向识别的效果. 根据二极管正极的朝向将其分为 4 类: 上、下、左、右, 每类角度范围为 90° , 具体分类细节如图 4 所示. 为了训练出分类效果比较好的算法模型, 首先对原始数据集进行数据增强的预处理, 经过对图片分别逆时针旋转 90° 、 180° 、 270° 与水平翻转、垂直翻转、水平垂直翻转等 6 种方式将数据集扩增到原数据集的 7 倍, 其中, 数据集 1 具体分布如表 1 所示, 数据集 2 具体分布如表 2 所示, 括号外的数字表示数据集包含该类的图片数量, 括号内

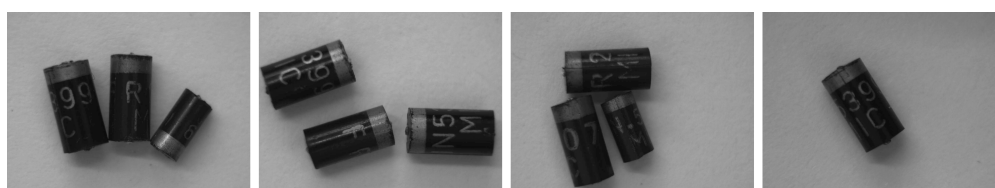


图 2 数据集 1 图片示例

Fig. 2 Image examples of the dataset 1

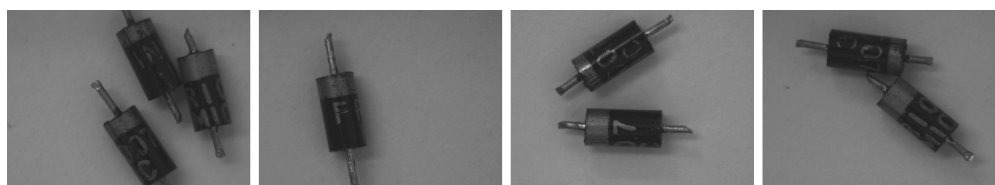


图 3 数据集 2 图片示例

Fig. 3 Image examples of the dataset 2

表 1 数据集 1 包含各类的图片数量和目标数量

Table 1 The dataset 1 contains the number of images and targets of targets in each category

二极管型号	方向 1 (上)	方向 2 (右)	方向 3 (下)	方向 4 (左)	总计
型号 1	2775 (2934)	3129 (3300)	2739 (2899)	3264 (3460)	9793 (12593)
型号 2	2833 (2970)	3123 (3315)	2813 (2946)	3180 (3369)	9800 (12600)
型号 3	2840 (2993)	3090 (3309)	2827 (2977)	3108 (3321)	9800 (12600)
总计	7032 (8897)	7570 (9924)	6975 (8822)	7676 (10150)	14000 (37793)

表 2 数据集 2 包含各类的图片数量和目标数量

Table 2 The dataset 2 contains the number of images and targets of targets in each category

二极管型号	方向 1 (上)	方向 2 (右)	方向 3 (下)	方向 4 (左)	总计
型号 1	1278 (1365)	1349 (1448)	1322 (1461)	1329 (1441)	4949 (5714)
型号 2	1282 (1380)	1380 (1467)	1307 (1400)	1344 (1430)	4970 (5677)
总计	2155 (2745)	2289 (2915)	2214 (2861)	2673 (2871)	5691 (11391)

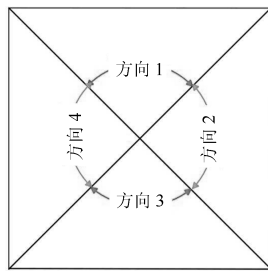


图 4 方向识别转化为分类

Fig. 4 Direction recognition change to classification

的数字表示数据集包含该类别的目标数量。

2 Anchor boxes 的选取方法

在目标检测的深度学习算法中, 每种算法都有基于某些数据集给出 Anchor boxes 的大小与个数, 如在 Faster RCNN 给出的 9 个 Anchor boxes^[14], YOLOv2 提供的 5 个 Anchor boxes^[19], YOLOv3 给出的 9 个在不同大小特征图 (Feature map) 的 Anchor boxes^[20], 但默认算法提供的 Anchor boxes 不一定适用在某些特定场景下的数据集, 如果过多, 会造成训练时间过长、检测速度过慢、资源浪费等问题. 因此, 选取适合特定数据集的 Anchor boxes 至关重要. 因此, 本文结合具体应用背景, 提出一种基于 K-means 算法设计 Anchor box 的方法, 并且在 Faster RCNN 与 YOLOv3 上取得良好的效果. 具体步骤如下: 1) 计算训练集中每个目标框 (Ground truth boxes, GT boxes) 与算法给出的 Anchor boxes 的交并比 (Intersection over union, IoU); 2) 每个 GT Boxes 选择与其 IoU 最大的 Anchor box 并统计每个 Anchor box 被选到的个数, 剔除计数为 0 的 Anchor boxes, 选择计数不为 0 的 n 个 Anchor boxes 为初始化的聚类中心; 3) 用 K-means 算法进行聚类迭代计算, 最终得到 n 个适合该数据集的 Anchor boxes.

2.1 基于 Faster RCNN 算法的 Anchor boxes

Faster RCNN 算法^[14] 是根据设定的 3 个长宽比例和 3 种面积生成 9 个 Anchor boxes, 这样设置难于直观调节大小, 而且部分 Anchor boxes 之间有关联, 在特定数据集下, 几乎得不到与训练集平均 IoU 最大的 Anchor boxes. 相对来说, 以长宽设置 Anchor box 大小更为直观, 这样能得到与训练集平均 IoU 最大的 Anchor box. 在极性电子元器件的数据集 1 和数据集 2 中, 其训练集需检测的目标长宽的分布如图 5 所示, 图中横坐标、纵坐标分别表示 GT boxes 的宽度和高度, 训练集所有 GT boxes 用圆表示, 星号为 Faster RCNN 提供的 9 个 Anchor boxes, 根据上述计算步骤, 最远离原坐标的 3 个 Anchor boxes 计数为 0, 说明相比较其他

Anchor box, 这 3 个距离 GT boxes 比较远, 训练集与其 IoU 没有比其他 6 个 Anchor boxes 大. 所以, 对于该数据集, Faster RCNN 只要 6 个 Anchor boxes 足以满足数据集目标的大小. 故选取算法提供的其他 6 个 Anchor boxes 作为 K-means 算法初始聚类的点, 经过多次迭代, 最终得到适用于该数据集的 Anchor boxes, 即图 5 中正方形的点. 通过计算, 算法提供的 9 个 Anchor boxes 与训练集的平均 IoU 为 0.6294, IoU 的均方差为 0.08289. 经过 K-means 算法得到的 6 个 Anchor boxes 与训练集的平均 IoU 为 0.8, IoU 的均方差为 0.0555.

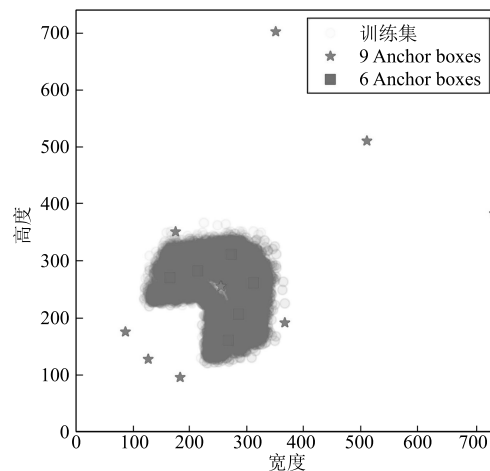


图 5 基于 Faster RCNN 的 Anchor boxes 和训练集的分布
Fig. 5 Distribution of training set and anchors boxes based on Faster RCNN

2.2 基于 YOLOV3 算法的 Anchor boxes

YOLOv3 算法^[20] 采用多个 Scale 融合的方式, 利用 3 种大小不一样的特征图分别检测大、中、小的目标, 并在每个特征图上选择 3 个 Anchor boxes 进行遍历, 在对 GT boxes 分布比较广的、目标比较小的数据集上效果比较好, 但在特定数据集不一定需要 3 种尺度的特征图. YOLOv3 虽然提出用 K-means 算法根据数据集目标的长宽来选取 Anchor box, 但其初始化 K-means 算法的聚类中心一般是在训练集中随机选取, 而且聚类中心的个数需要自己设定. 这样选取的 Anchor box 随机性比较大, 不一定是最佳选择, 而且个数难于设定. 在有极性电子元器件的数据集 1 和数据集 2 中, 其训练集的 GT boxes 转化后的分布如图 6 所示. 图中横坐标、纵坐标分别为 GT boxes 的宽度和高度, 训练集所有 GT boxes 用圆表示, 星号为 YOLOv3 提供的 9 个 Anchor boxes 的分布, 根据上述计算步骤, 只有 3 个 Anchor boxes 的计数不为 0. 所以, 对于该数据集, YOLOv3 只要 3 个 Anchor boxes 足以满足该数据集. 故选取这 3 个 Anchor boxes 作为 K-means 算

法初始聚类的点. 经过多次迭代, 最终得到适用于该数据集的 Anchor boxes, 即图 6 中正方形的点. 通过计算, 算法提供的 9 个 Anchor boxes 与训练集的平均 IoU 为 0.44768, IoU 的均方差为 0.1136. 经过 K-means 算法得到的 3 个 Anchor boxes 与训练集的平均 IoU 为 0.6645, IoU 的均方差为 0.0870.

3 YOLOv3-BigObject

YOLOv3 算法的网络结构以 Darknet-53 为基

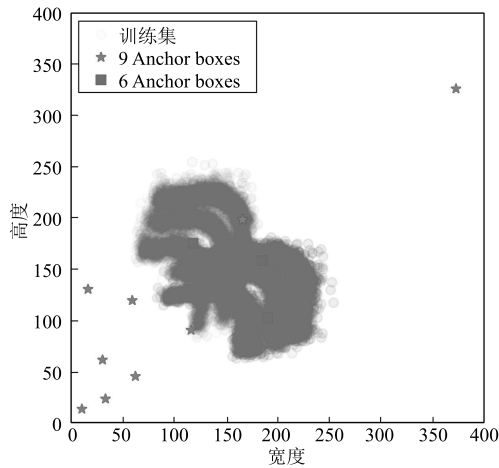


图 6 基于 YOLOv3 的 Anchor boxes 和训练集的分布
Fig. 6 Distribution of training set and anchor boxes based on YOLOv3

础, 在第 36, 61, 79 层分别形成大小为 52×52 , 26×26 , 13×13 的特征图, 如图 7 所示, 3 种特征图对应的 Anchor boxes 大小分别为 $((10, 13), (16, 30), (33, 26))$, $((30, 61), (62, 45), (59, 119))$, $((116, 90), (156, 198), (373, 326))$. 结合特征金字塔网络 (Feature pyramid network, FPN)^[25] 的思想, 后面生成的特征图融合前面生成的特征图, 可以提高对中小物体检测的准确率, 最终在 82, 94, 106 层分别经过输出对大, 中, 小目标的检测结果. 根据前面 YOLOv3 在极性电子元器件的数据集上得到适用于该数据集的 3 个 Anchor boxes: $((118, 175), (190, 102), (185, 154))$. 显而易见, 该数据集都是大目标, 这 3 个 Anchor boxes 与算法在 13×13 特征图对应的 Anchor boxes $((116, 90), (156, 198), (373, 326))$ 最为接近, 在 13×13 特征图上进行遍历能有最好的结果. 其他两种特征图在该数据集上明显多余, 基于此, 我们对 YOLOv3 结构进行改进, 删减生成 52×52 , 26×26 特征图的结构, 保留中间的下采样结构, 把原来 106 层的结构缩减为 46 层的结构, 如图 8 所示.

YOLOv3-BigObject 具体的网络结构如图 9 所示, 由 YOLOV3 的 Darknet-53 结构改进得到, 它主要由 5 个下采样层、11 个残差块、1 个平均池化层、1 个全连接层组成. 下采样层、残差块中的卷积层依次包含有二维卷积 (Conv2d)、批归一化 (Batch normalization, BN)、LeakyReLU. 其中, BN 的目的是

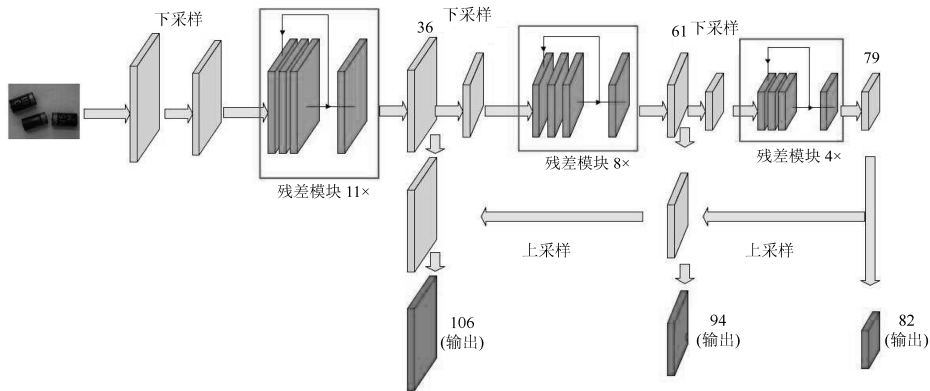


图 7 YOLOv3 的网络结构
Fig. 7 YOLOv3 network structure

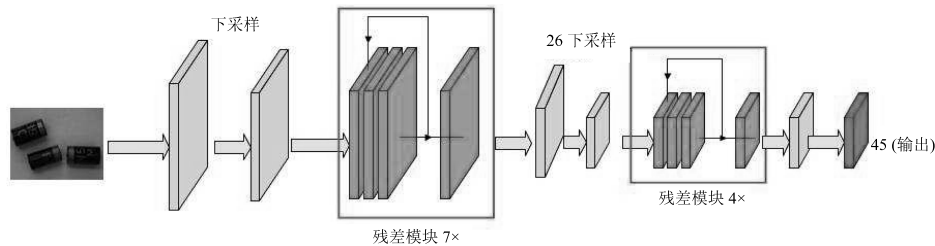


图 8 YOLOv3-BigObject 的网络结构
Fig. 8 YOLOv3-BigObject network structure

为了防止梯度消失,在不破坏特征分布的情况下将输入控制在一个稳定的范围内,使网络中每层输入数据分布比较稳定,并且可以加速模型的学习速率,式(1)为其计算公式,其中 y 为经过 BN 后的输出, x 为 BN 的输入, γ, β 分别为可学习的缩放参数与偏移参数, ε 是趋向于 0 的参数,预防式中分母为 0.

	类型	滤波器	尺寸	输出
	Convolutional	32	3 × 3	416 × 416
	Convolutional	64	3 × 3/2	208 × 208
1 ×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			208 × 208
	Convolutional	128	3 × 3/2	104 × 104
2 ×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			104 × 104
	Convolutional	256	3 × 3/2	52 × 52
2 ×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			52 × 52
	Convolutional	512	1 × 1/2	26 × 26
2 ×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			26 × 26
	Convolutional	1 024	1 × 1/2	13 × 13
4 ×	Convolutional	512	1 × 1	
	Convolutional	1 024	3 × 3	
	Residual			13 × 13
	Avgpool		Global	
	Connected		1 000	
	Softmax			

图 9 YOLOv3-BigObject 的内部网络参数设置结构
Fig.9 The Internal network parameter setting structure of the YOLO v3-BigObject

$$y = \frac{\gamma x}{\sqrt{\text{Var}(x) + \varepsilon}} + \left(\beta - \frac{\gamma E(x)}{\sqrt{\text{Var}(x) + \varepsilon}} \right) \quad (1)$$

在经过 BN 后, BN 所得的结果输入激活函数中,在此算法结构中的激活函数为 LeakyReLU,如式(2)所示, LeakyReLU 是一个非饱和激活函数,而且给所有负值赋予一个非零斜率,以解决梯度消失、加快收敛速度、避免神经元死亡等问题.

$$\text{LeakyReLU}(z) = \begin{cases} z, & z > 0 \\ \alpha z, & z \leq 0 \end{cases} \quad (2)$$

4 实验分析

4.1 实验设置

本文实验的显卡设备为 NVIDIA GeForce GTX 1080 Ti. 实验分为三部分: 实验 1 用于证

明本文所选取 Anchor boxes 的方式在准确率上有一定的提升,我们分别用 Faster RCNN 和 YOLOv3 算法对其进行验证,并在测试集中进行测试. 在 Faster RCNN 的实验中,设置的 Batch size 为 1, 训练步数为 140 000; 在 YOLOv3 的实验中,设置的 Batch size 为 64, 训练步数为 50 000. 实验 2 用于证明 YOLOv3-BigObject 网络结构有显著的提升效果,我们分别用 YOLOv3 的网络结构和 YOLOv3-BigObject 进行训练和测试,训练方式和实验 1 相同,然后对比两者的训练与测试情况. 实验 3 用于验证算法的性能,检验增加类别后算法的鲁棒性,在数据集 1 的基础上增加数据集 2 作为该实验的数据集.

4.2 基于 K-means 算法设计 Anchor boxes 的实验结果

在实验 1 中,用数据集 1 进行实验,经过调节学习率、数据增强等步骤,我们设计的深度学习算法能够很好地识别极性电子元器件的类别与方向,如图 10 所示,标签第 1 个数字是表示种类,后面的数字表示方向. 在 Faster RCNN 算法中,我们用算法提供的 9 个 Anchor boxes 与 K-means 算法设计的 Anchor boxes 进行实验对比,原 Anchor boxes 与训练集的平均 IoU 为 0.7037, Faster RCNN (VGG-16)、Faster RCNN (ResNet101) 在测试集中的 mAP 分别为 93.77% 和 97.05%. 经过 K-means 算法设计 Anchor boxes 的大小与个数,所设计的 6 个 Anchor boxes 与训练集的平均 IoU 为 0.8577, Faster RCNN (VGG-16)、Faster RCNN (ResNet101) 在测试集中的 mAP 分别为 94.95% 和 97.36%, mAP 分别提高了 1.18%、0.31%. 在 YOLOv3 算法中,用算法提供的 9 个 Anchor boxes 进行训练,Anchor boxes 与训练集的平均 IoU 为 0.5728. 在测试集中的 mAP 为 99.22%. 经过 K-means 算法选取 Anchor boxes 的大小与个数,我们所选取的 Anchor boxes 与训练集的平均 IoU 为 0.7362,最终在测试集中的 mAP 为 99.31%. 图 11

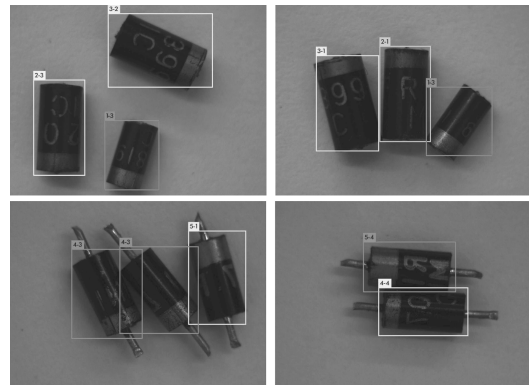


图 10 测试集结果展示

Fig.10 The display of test sets result

和图 12 分别是 Faster RCNN 和 YOLOv3 训练过程中损失函数 (Loss) 的下降曲线. 由上面的实验结果可知, 经过 K-means 算法对特定数据集提取 Anchor boxes 的个数与大小在测试集上表现得更优异, 两种算法都大大提高了与训练集的平均 IoU. 这样选取的方式使网络在学习 Anchor boxes 回归时复杂度降低, 进一步提高了检测的精度. 在 Faster RCNN 中, 经过 K-means 算法选取的 Anchor boxes 在训练时的 Loss 的收敛速度与收敛效果比 9 个 Anchor boxes 表现得更好.

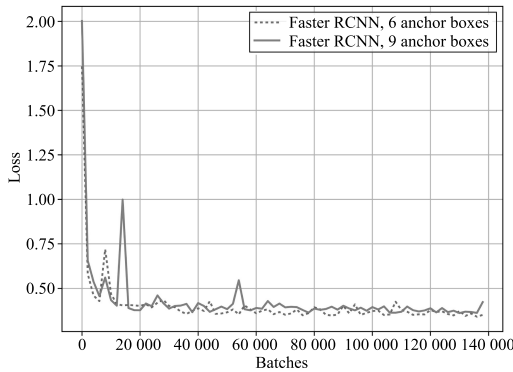


图 11 Faster RCNN (ResNet101) 的 Loss 曲线

Fig. 11 The loss curves of Faster RCNN (ResNet101)

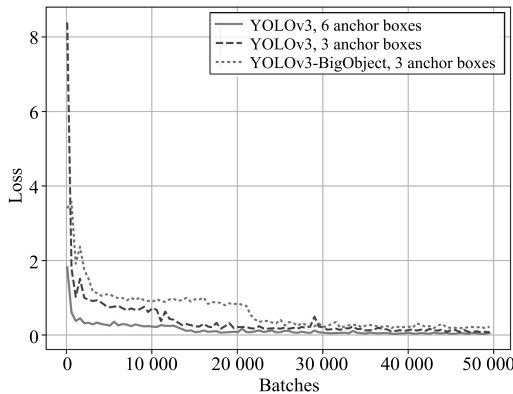


图 12 YOLOv3 和 YOLOv3-BigObject 的 Loss 曲线

Fig. 12 The loss curves of YOLOv3 and YOLOv3-BigObject

4.3 YOLOv3-BigObject 网络的实验结果

在实验 2 中, 用数据集 1 进行实验, 检验 YOLOv3-BigObject 在极性元器件分类与方向识别的效果, 通过对 YOLOv3 和 YOLOv3-BigObject 网络结构的训练和测试, 结果如表 3 所示. 在经过 50 000 次迭代后, YOLOv3-BigObject 在测试集中的 mAP 为 99.44%. 相比于 YOLOv3 网络结构, mAP 提高 0.13%, 两者 mAP 相近. 但 YOLOv3-BigObject 的检测速度几乎是 YOLOv3 的一半. YOLOv3-BigObject 在训练过程的 Loss 下降得比较缓慢, 如图 10, 其主要原因是, 预训练模型中每一层负责感知的特征已经学习形成, 改变结构后, 对目标物体的整体感知不全, 相当于每层需重新学习感知对应的特征, 这是 Loss 下降缓慢的最主要原因. 综上所述, YOLOv3-BigObject 网络结构在大目标检测上有与原来网络结构相当的准确率的基础上, 将检测速度大约提高到原来的两倍.

4.4 增加数据集的实验结果

我们在数据集 1 的基础上, 进一步增加了数据集 2 对算法进行实验验证, 实验结果如表 4 所示, 结果显示运用 K-means 算法设计 Anchor boxes 后, Faster RCNN 与 YOLOv3 的算法性能依然得到了一定程度的提升, 而且实验效果比在 FPN、SSD 等算法中表现得更好. 在本实验中, YOLOv3-BigObject 的 mAP 高达 0.9920, 证明了 YOLOv3-BigObject 即使在增加种类后的情况下, 也具有比较好的性能. 为了进一步更好地检验算法的性能, 我们选了一个焊接有数据集 1 和数据集 2 中的 5 种型号的元器件电路板, 但不是用来拍摄制作数据集的元器件对焊有元器件的电路板进行视频拍摄, 并用 YOLOv3-BigObject 生成的算法模型对视频进行检测, 结果如图 13 所示. 实验结果显示我们的算法能够很好地检测出电路板上元器件的类别与方向, 进一步证明了我们的算法具有很好的鲁棒性和实用性.

表 3 数据集 1 的实验结果

Table 3 Experimental result of dataset 1

	Backbone	Anchor boxes	Anchor boxes 与训练集的平均 IoU	mAP (%)	测试速度 (s)
Faster RCNN	VGG-16	9	0.7037	93.77	-
Faster RCNN	VGG-16	6	0.8577	94.95	-
Faster RCNN	ResNet101	9	0.7037	97.05	-
Faster RCNN	ResNet101	6	0.8577	97.36	-
YOLOv3	Darknet-53	9	0.5728	99.22	-
YOLOv3	Darknet-53	3	0.7362	99.31	0.0217
YOLOv3-BigObject	-	3	0.7362	99.44	0.0118

表 4 数据集 1 和数据集 2 的实验结果

Table 4 Experimental result of datasets 1 and 2

	Backbone	Anchor boxes	Anchor boxes 与训练集的平均 IoU	mAP (%)	测试速度 (s)
Faster RCNN	ResNet101	9	0.6294	97.28	—
Faster RCNN	ResNet101	6	0.8000	97.53	—
FPN	ResNet101	9+	—	98.70	—
SSD	ResNet101	9+	—	76.25	—
YOLOv3	Darknet-53	9	0.4476	99.26	—
YOLOv3-BigObject	—	3	0.6645	99.20	0.0118

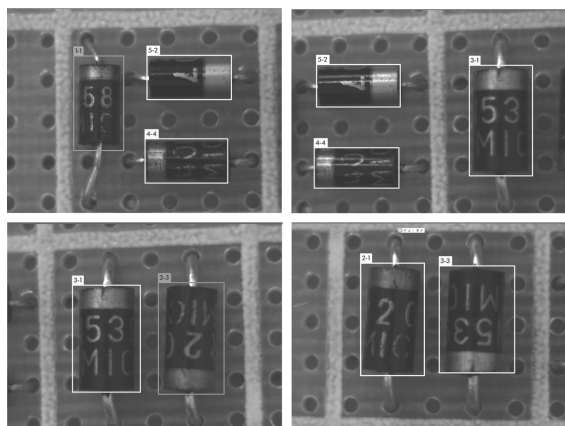


图 13 视频检测结果

Fig. 13 Video experimental result

5 总结与展望

本文提出了一种基于深度学习的算法, 实现了极性电子元器件的类别与方向检测. 实验结果表明, 我们所改进的深度学习算法在电子元器件的类别分类上比传统算法和卷积网络上表现得更好, 而且还能进一步识别类别的方向. 通过 K-means 算法设计 Anchor boxes 的方式提高了深度学习算法检测的准确率和速度. 本文也对 YOLOv3 网络结构进行了压缩改进, 其在检测大目标提高准确率的同时, 将检测时间几乎缩小到一半, 进一步达到实时检测的效果, 最后将算法模型运用到焊接有元器件的电路板上检测, 得到很好的效果. 本文所提出的方法在极性电子元器件类别、方向识别、定位、焊接和检测等方面也同样有很大的应用潜力. 虽然本文在极性电子元器件目标检测与方向识别上能够得到很好的识别效果, 但其仍然有很大的改进空间, 主要有两个方面: 一是类别增多问题, 二是方向识别提升精度问题. 在第一个方面, 类别增多使分类更加复杂, 未来可以进行如下展望: 1) 建立更加庞大规范的数据集; 2) 在达到算法模型能够高准确率分类的最大类别数时, 对电路板进行划区域或划种类对模型进行训练. 在第二个方面, 如何进一步将方向识别的精度提升到具体多少度, 可以考虑下面几个方面: 1) 与其他回归

方法结合的方式得到具体角度; 2) 借鉴文本检测的思想, 用带角度 (如: 与水平线夹角每转 30° 为一个 Anchor) 的 Anchor 加上回归方法得到具体的角度; 3) 用语义分割的方法识别出每个元器件的区域, 再对元器件区域用传统图像处理的方法进行边缘检测等操作得到具体角度.

References

- Du Si-Si, Li Zhuo-Qiu, Zhang Xue-Feng. Research on accurate positioning technology of image features of electronic components. *Journal of Wuhan University of Technology*, 2008, **30**(3): 359–362
(杜思思, 李卓球, 张雪峰. 电子元器件图像外形特征的精确定位技术研究. *武汉理工大学学报*, 2008, **30**(3): 359–362)
- Li Wen-Ying, Cao Bin, Cao Chun-Shui, Huang Yong-Zhen. A deep learning based method for bronze inscription recognition. *Acta Automatica Sinica*, 2018, **44**(11): 2023–2030.
(李文英, 曹斌, 曹春水, 黄永祯. 一种基于深度学习的青铜器铭文识别方法. *自动化学报*, 2018, **44**(11): 2023–2030)
- Luo Jian-Hao, Wu Jian-Xin. A survey on fine-grained image categorization using deep convolutional features. *Acta Automatica Sinica*, 2017, **43**(8): 1306–1318
(罗建豪, 吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述. *自动化学报*, 2017, **43**(8): 1306–1318)
- Tian Juan-Xiu, Liu Guo-Cai, Gu Shan-Shan, Ju Zhong-Jian, Liu Jin-Guang, Gu Dong-Dong. Deep learning in medical image analysis and its challenges. *Acta Automatica Sinica*, 2018, **44**(3): 401–424
(田娟秀, 刘国才, 谷珊珊, 鞠忠建, 刘劲光, 顾冬冬. 医学图像分析深度学习方法研究与挑战. *自动化学报*, 2018, **44**(3): 401–424)
- Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. In: *Proceedings of the 2014 Advances in Neural Information Processing Systems*, Lake Tahoe, Spain, 2012. 1097–1105
- Zheng Wen-Bo, Wang Kun-Feng, Wang Fei-Yue. Background subtraction algorithm with bayesian generative adversarial networks. *Acta Automatica Sinica*, 2018, **44**(5): 878–890
(郑文博, 王坤峰, 王飞跃. 基于贝叶斯生成对抗网络的背景消减算法. *自动化学报*, 2018, **44**(5): 878–890)
- Zhang Hui, Wang Kun-Feng, Wang Fei-Yue. Advances and perspectives on applications of deep learning in visual object detection. *Acta Automatica Sinica*, 2017, **43**(8): 1289–1305
(张慧, 王坤峰, 王飞跃. 深度学习在目标视觉检测中的应用进展与展望. *自动化学报*, 2017, **43**(8): 1289–1305)

- 8 Yu Jin-Yong, Ding Peng-Cheng, Wang Chao. Overview: Application of convolution neural network in object detection. *Computer Science*, 2018, **45**(Z11): 17–26
(于进勇, 丁鹏程, 王超. 卷积神经网络在目标检测中的应用综述. *计算机科学*, 2018, **45**(Z11): 17–26)
- 9 Girshick R, Donahue J, Darrell T, Jitendra M. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA: IEEE, 2014. 580–587
- 10 Uijlings J R R, Van De Sande K E A, Gevers T, Smeulders A W M. Selective search for object recognition. *International Journal of Computer Vision*, 2013, **104**(2): 154–171
- 11 Zitnick C L, Dollár P. Edge boxes: Locating object proposals from edges. In: Proceedings of the 13th European Conference on Computer Vision, Zurich, Switzerland: Springer, 2014. 391–405
- 12 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognitions [Online], available: <https://arxiv.org/abs/1406.2952>, April 10, 2015
- 13 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016. 770–778
- 14 Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. In: Proceedings of the 2015 Advances in Neural Information Processing Systems, Montreal, Canada, 2015: 91–99
- 15 Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA: IEEE, 2015. 3431–3440
- 16 Girshick R. Fast R-CNN. In: Proceedings of the 2015 IEEE International Conference on Computer Vision, Santiago, Chile: IEEE, 2015. 1440–1448
- 17 Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, Berg A C. SSD: Single shot multibox detector. In: Proceedings of the 2016 European Conference on Computer Vision, Berlin, Germany: Springer-Verlag, 2016. 21–37
- 18 Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA: IEEE, 2016. 779–788
- 19 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, Hawaii, USA: IEEE, 2017. 7263–7271
- 20 Redmon J, Farhadi A. Yolov3: An incremental improvement [Online], available: <https://arxiv.org/abs/1804.02767>, April 8, 2018
- 21 Du Si-Si, San Zhong-De, Huang Zhang-Can, Liu Hong-Qi, Liu Shi-Li. Algorithm research and system development for automatic classification of electronic components. *Mechanical and Electrical Product Development and Innovation*, 2008, **21**(6): 133–135
(杜思思, 单忠德, 黄樟旭, 刘红旗, 刘世立. 电子元器件自动分类的算法研究与系统开发. *机电产品开发与创新*, 2008, **21**(6): 133–135)
- 22 Chen Xiang, Yu Jian-Ding, Chen Xiao-Ai, Zhai Ying. Classification of electronic components based on convolutional neural networks. *Wireless Communication Technology*, 2018, **27**(2): 7–12
(陈翔, 俞建定, 陈晓爱, 翟影. 基于卷积神经网络的电子元器件分类研究. *无线通信技术*, 2018, **27**(2): 7–12)
- 23 Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors. *Nature*, 1986, **323**(6088): 533–536
- 24 Hinton G E, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov R R. Improving neural networks by preventing co-adaptation of feature detectors [Online], available: <https://arxiv.org/abs/1207.0580>, July 3, 2012
- 25 Lin T Y, Dollár P, Girshick R, He K M, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA: IEEE, 2017. 2117–2125



陈文帅 广东工业大学自动化学院硕士研究生. 主要研究方向为深度学习, 计算机视觉.

E-mail: 2111704068@mail2.gdut.edu.cn
(**CHEN Wen-Shuai** Master student at the School of Automation, Guangdong University of Technology. His research interest covers deep learning and

computer vision.)



任志刚 广东工业大学自动化学院副教授. 主要研究方向为最优控制, 知识自动化, 人工智能.

E-mail: renzhigang@gdut.edu.cn
(**REN Zhi-Gang** Associate professor at the School of Automation, Guangdong University of Technology. His research interest covers optimal control, knowledge automation, and artificial intelligence.)



吴宗泽 广东工业大学自动化学院教授. 主要研究方向为自动控制, 信号处理, 大数据, 知识自动化, 人工智能. 本文通信作者. E-mail: zzwu@gdut.edu.cn

(**WU Zong-Ze** Professor at the School of Automation, Guangdong University of Technology. His research interest covers automation control, signal

processing, big data, knowledge automation, and artificial intelligence. Corresponding author of this paper.)



付敏跃 广东工业大学自动化学院教授, 澳大利亚纽卡斯尔大学终身教授. 主要研究方向为自动控制, 信号处理和通信.

E-mail: minyue.fu@newcastle.edu.au
(**FU Min-Yue** Professor at the School of Automation, Guangdong University of Technology and tenured professor at University of Newcastle, Australia. His research interest covers automation control, signal

processing, and communications.)