

基于改进 Faster R-CNN 的子弹外观缺陷检测

马晓云^{1,2,3,4,5*}, 朱丹^{1,2,3,4,5}, 金晨^{1,2,4,5}, 佟新鑫^{1,2,4,5}

¹中国科学院沈阳自动化研究所, 辽宁 沈阳 110016;

²中国科学院机器人与智能制造创新研究院, 辽宁 沈阳 110016;

³中国科学院大学, 北京 100049;

⁴中国科学院光电信息处理实验室, 辽宁 沈阳 110016;

⁵辽宁省图像理解与视觉计算重点实验室, 辽宁 沈阳 110016

摘要 为了实现子弹外观缺陷的自动检测, 解决传统机器视觉方法在缺陷检测方面手工设计目标特征耗时和泛化能力差的问题, 针对子弹外观缺陷数据集, 采用 K -means++ 算法改进锚框的生成方法, 提出了 Faster R-CNN 子弹外观缺陷检测模型。该模型采用卷积神经网络, 可以自动提取目标特征, 泛化能力强。将该检测模型分别与 ZFNet、VGG_CNN_M_1024 和 VGG16 结合, 结果表明, 与 VGG16 结合的检测模型的检测精度高于其他两种模型方案, 并且在所提算法的基础上, 精度提升到了 97.75%, 速度达到 $28 \text{ frame} \cdot \text{s}^{-1}$ 。

关键词 测量; 目标检测; 子弹外观缺陷; 卷积神经网络

中图分类号 TP391

文献标识码 A

doi: 10.3788/LOP56.151202

Bullet Appearance Defect Detection Based on Improved Faster Region-Convolutional Neural Network

Ma Xiaoyun^{1,2,3,4,5*}, Zhu Dan^{1,2,3,4,5}, Jin Chen^{1,2,4,5}, Tong Xinxin^{1,2,4,5}

¹Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

²Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China;

³University of Chinese Academy of Sciences, Beijing 100049, China;

⁴Key Laboratory of Opto-Electronic Information Processing, CAS, Shenyang, Liaoning 110016, China;

⁵The Key Lab of Image Understanding and Computer Vision, Shenyang, Liaoning 110016, China

Abstract To realize automatic detection of bullet appearance defects and to overcome the limitations associated with traditional machine vision methods, i. e., excessive time required to manually design a target feature and generalization ability is poor in defect detection, we use the K -means++ algorithm to improve the anchor frame generation method and propose a bullet appearance defect detection model based on the improved faster region-convolutional neural network (R-CNN). The proposed model uses a CNN that can automatically extract target features and has strong generalization ability. The detection model is combined with ZFNet, VGG_CNN_M_1024, and VGG16, respectively. Results demonstrate that the detection accuracy of the detection model combined with VGG16 is higher than the others. The results show that that of the improved model demonstrates 97.75% accuracy and the speed reaches $28 \text{ frame} \cdot \text{s}^{-1}$.

Key words measurement; target detection; bullet appearance defect; convolutional neural network

OCIS codes 120.5010; 100.4996; 100.3008

1 引言

在子弹生产制造过程中,由于原材料、加工工艺、生产设备等因素的影响,弹壳可能会出现壳口不齐、口部裂缝等缺陷,这些缺陷是直接影响子弹精度及安全性的关键因素。目前,国内外的大多数军工企业仍然采用传统的人工视觉检测方法,虽然成本较低,但检测结果因人而异,标准难以统一,检测精度和效率都较低,不利于子弹的自动化生产。因此,提高子弹外观缺陷检测的精度、效率以及自动化程度,是急需解决的问题。

目前国内对子弹外观缺陷检测也有许多相关研究,如张亚军等^[1]在2012年采用了基于图像灰度值形态学的枪弹钢芯表面检测方法;史进伟等^[2]在2013年分析了采用灰度阈值及单一颜色模型分割缺陷的局限性,提出了基于混合颜色模型的缺陷图像分割方法,该方法采用优化的BP神经网络作为缺陷分类器,具有较好的准确性和稳健性;史进伟等^[3]又采用基于改进的Hough变换和双峰法的子弹缺陷检测方法,快速、准确地提取了子弹外观缺陷图像,并将该方法拓展到带钢、玻璃、路面等图像灰度分布不均的缺陷自动检测中;王鹏等^[4]在2015年针对子弹外观缺陷图像特征,采用一种基于Otsu的快速迭代算法,实现了缺陷的自动分割;王鹏等^[5]又提出了一种基于支持向量机(SVM)的子弹外观缺陷自动识别和分类模型,该模型具有较好的分类性能。

上述针对子弹外观缺陷检测的方法均是传统的基于机器视觉的方式,用于检测的特征是人为选择和设计的,过程中存在较多的主观导向,需要研究人员具有较丰富的专业知识和经验,设计过程比较耗时,而且最终得到的特征也参差不齐,泛化能力较差^[6]。

深度学习的出现解决了这一问题。深度学习中的卷积神经网络(CNN)^[7]可以根据训练数据集,自

动地提取目标特征而无需人工设计,具有更强的表达能力^[8]。近几年来,基于卷积神经网络模型的目标检测算法层出不穷,并且在目标检测方面具有良好的性能。Girshick等^[9]采用区域卷积神经网络(R-CNN)的目标检测算法,利用Selective Search算法在图像上获得1000~2000个候选区域,分别将这些候选区域输入卷积神经网络中提取特征,并用SVM训练的分类器将候选区域分为目标或者背景。为了解决R-CNN中的大量重复计算问题,Girshick等^[10]又采用Fast R-CNN目标检测算法,将候选区域映射到卷积神经网络的最后一层特征图中,图像只需要提取一次特征,极大地提高了目标检测的性能。Ren等^[11]提出的由区域建议网络(RPN)和Fast R-CNN组成的Faster R-CNN目标检测算法,进一步提高了检测精度和速度。

本文将Faster R-CNN目标检测算法引入到子弹外观缺陷检测中,将其分别与3种卷积神经网络ZFNet^[12]、VGG_CNN_M_1024^[13]和VGG16^[13]相结合,通过实验对比3种网络的性能。针对子弹外观缺陷数据集,采用K均值聚类算法(K-means++)对RPN网络中的锚框(anchor)尺寸和数量进行改进,以提高检测准确率,同时调整非极大值抑制(NMS)后提取的anchor数量,以提高检测速度。

2 检测网络结构模型

图1为Faster R-CNN的检测流程,主要由卷积神经网络、区域建议网络、感兴趣区域(RoI)池化和目标检测4个部分组成。图1中,Conv layers表示卷积层,FC表示全连接层。

2.1 卷积神经网络

卷积神经网络专门针对图像识别问题,具有类似于生物神经网络的全连接和权重共享的网络结构,网络模型的复杂度较低,权值数量较少。与传统识别

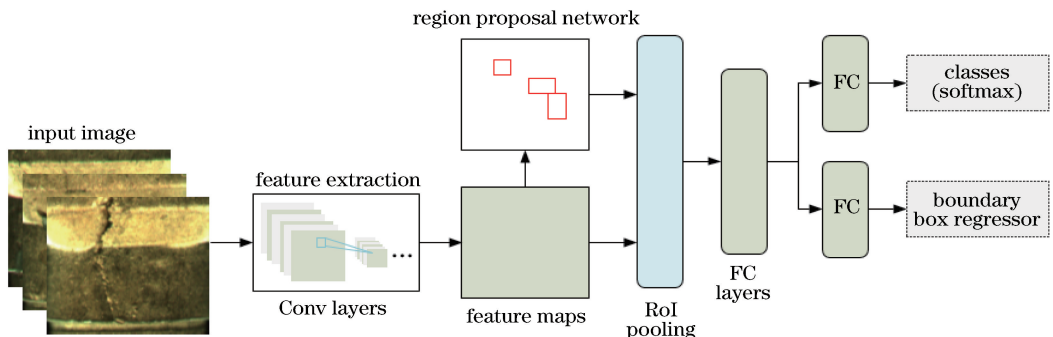


图1 基于Faster R-CNN的目标检测框架

Fig. 1 Target detection framework based on Faster R-CNN

的人工设计特征相比,卷积神经网络能够根据训练数据集自动提取特征,具有更强的泛化能力^[14]。

将输入图像大小调整为 $M \times N$,使用卷积神经网络(ZFNet、VGG_CNN_M_1024、VGG16)中一组基础的 Conv+ReLU+Pooling 层提取输入图像的特征图,该特征图被共享用于后续的 RPN 层和全连接层。

2.2 区域建议网络

RPN 是一种全卷积网络(FCN)^[15],与检测网络共享卷积层。为了生成区域建议框,在最后一个共享的卷积层生成的卷积特征图上用一个小网络进行滑动。该网络将卷积层生成的卷积特征图的 d

个 $n \times n$ 空间窗口作为输入(本文中 $n=3$),映射到一个 d 维向量上(ZFNet: 256 维,VGG: 512 维)。这个 d 维向量同时输出给两个同级的全连接层——框回归层(reg layer)和框分类层(cls layer)。在每一个滑动窗口的位置,同时预测 k 个区域建议框,所以回归层有 $4k$ 个输出,用于编码 k 个区域建议框的坐标,分类层输出 $2k$ 个得分,用于估算每个区域建议框是目标或非目标的概率。相对于被称为 anchor 的 k 个参考边界框, k 个区域建议框是参数化的。anchor 的中心位于所讨论的滑动窗口的中心。图 2 为区域建议网络结构。

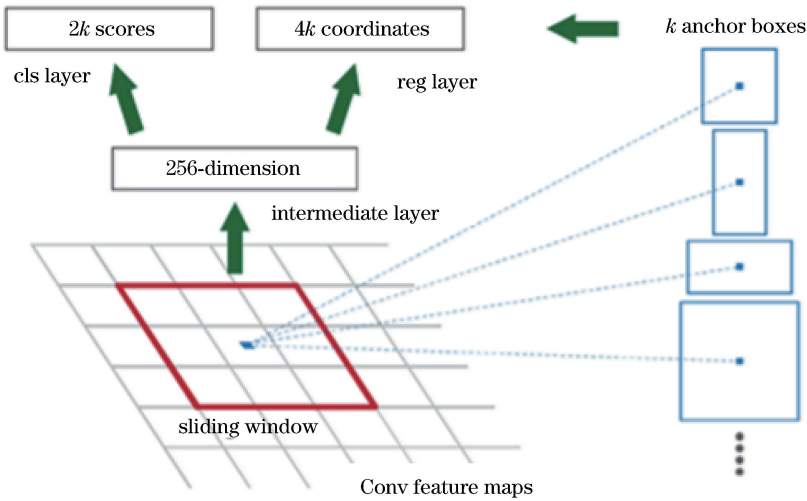


图 2 区域建议网络结构

Fig. 2 Structure of region proposal network

RPN 的总体损失函数定义为

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

式中: L_{cls} 为分类损失; L_{reg} 为回归损失; λ 为权重平衡参数; i 表示一个 mini-batch 中 anchor 的索引; p_i 为 anchor i 属于目标的预测概率; p_i^* 为对应的真值(GT)预测概率; τ_i 为预测的包围盒的 4 个参数化坐标; τ_i^* 为与正 anchor 对应的 GT 包围盒的坐标。

$$\begin{cases} t_x = (x - x_a)/w_a \\ t_y = (y - y_a)/h_a \\ t_w = \ln(w/w_a) \\ t_h = \ln(h/h_a) \\ t_x^* = (x^* - x_a)/w_a \\ t_y^* = (y^* - y_a)/h_a \\ t_w^* = \ln(w^* - w_a) \\ t_h^* = \ln(h^* - h_a) \end{cases}, \quad (2)$$

式中: x, y 为预测的包围盒的中心坐标; w, h 为预测的包围盒的宽度和高度; x_a, y_a 为 anchor 的中心坐标; w_a, h_a 分别为 anchor 的宽度和高度; x^*, y^* 为 GT 包围盒的中心坐标; w^*, h^* 分别为 GT 包围盒的宽度和高度。

分类损失 L_{cls} 是目标与非目标的对数损失,回归损失 L_{reg} 定义为

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

式中: R 为鲁棒损失函数(smooth L_1)^[10]。 $p_i^* L_{\text{reg}}$ 表示仅 anchor 为正 ($p_i^* = 1$) 时有回归损失, anchor 为负 ($p_i^* = 0$) 时没有。(1) 式中的分类项和回归项分别通过 N_{cls} (一般取值为 256) 和 N_{reg} (一般取值为 2400) 归一化, 并设 $\lambda = 10$, 使这两项权重大致相等。

2.3 RoI 池化

由于 Faster R-CNN 使用全连接层, 所以利用 RoI 池化将不同大小的 RoI 转换为固定大小。

图 3 为 RoI 池化过程。池化层的输入分为两部

分;RPN 网络输出的目标建议框(proposals)和最后一个卷积层输出的图像的特征图。Proposals 对应 $M \times N$ 尺度,所以使用空间尺度数 1/16 将其映射回 $(M/16) \times (N/16)$ 大小的特征图尺度,将每个 proposals 水平方向和垂直方向都进行 7 等分,对每一份进行最大值池化处理。处理后,大小不同的 proposals,输出结果都是 7×7 大小,从而实现了固定长度输出。

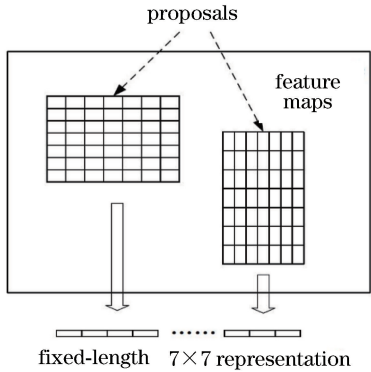


图 3 RoI 池化过程

Fig. 3 RoI pooling process

2.4 目标检测

将 RoI 池化获得的 proposal 特征图作为后续全连接层的输入,然后利用 softmax 函数计算每个 proposal 具体属于哪个类别(壳口不齐、口部裂缝),并输出类别概率向量。同时利用边界框回归获得每个 proposal 的位置偏移量,用于回归更加精确的目标检测框。

3 Anchor 生成方法改进

3.1 平移不变的 anchor

Anchors 实际上是一组由 RPN 生成的矩形,在文献[11]中,Faster R-CNN 目标检测算法采用 3 种尺度和 3 种长宽比(1:1,1:2,2:1),在每一个滑动窗口的中心生成 $k = 9$ 个 anchor,如图 4 所示。

通过 anchor 引入检测中常用的多尺度方法,anchor 的尺寸是根据检测图像设置的。任意大小的原始图像输入检测网络后均被调整为 $800 \text{ pixel} \times 600 \text{ pixel}$,而 anchor 在长宽比为 2:1 中的最大尺寸为 $736 \text{ pixel} \times 384 \text{ pixel}$,在长宽比为 1:2 中的最大尺寸为 $352 \text{ pixel} \times 704 \text{ pixel}$,这 9 个 anchor 基本覆盖了 $800 \text{ pixel} \times 600 \text{ pixel}$ 的各个尺度和形状。

3.2 基于 K-means++ 改进 anchor 的生成方法

使用 anchor 对子弹缺陷数据集进行检测时遇

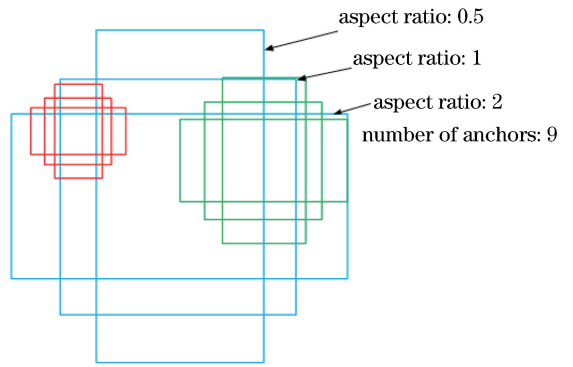


图 4 Anchor 比例示意图

Fig. 4 Diagram of anchor proportion

到了问题。Anchor boxes 的宽高维度针对的是 PASCAL VOC^[16] 数据集的精选先验框,虽然在训练过程中网络能够继续学习调整框的宽高维度,最终得到准确的 bounding boxes,但是,对于子弹缺陷数据集,这些先验框维度并不具有代表性。

使用 K-means++ 聚类方法生成新的 anchor boxes 宽高维度来改进 anchor 的生成方法,提升对子弹缺陷的检测性能。利用训练集中所有目标的 bounding boxes 左上角和右下角的坐标,得到所有目标的 bounding boxes 的宽高信息。这些宽高信息组成 K-means++ 的输入数据集 D ,设定聚类个数为 g ,通过 K-means++ 聚类方法,得到 g 个聚类中心点,设为新的 g 个 anchor boxes 的初始宽高维度。

K-means++ 方法使用的是欧氏距离函数,具体步骤如下。

输入:数据集 D (包含训练集中所有目标的 bounding boxes 的宽高)、聚类个数 g 。

输出: g 个聚类中心点。

1) 从 D 中随机选择一个聚类中心点 c_1 ;

2) 计算 D 中每个样本点 $X_p (p \in 1, 2, \dots, m, m$ 为样本点个数)与最近聚类中心 $c_j (j \in 1, 2, \dots, g)$ 的距离 $d(X)$,表达式为

$$d(X) = \sqrt{\sum_{p=1}^m (X_p - c_j)^2}; \quad (4)$$

3) 计算样本点 X 作为下一个初始质心的概率 $P(X)$,表达式为

$$P(X) = \frac{d^2(X)}{\sum_{X \in D} d^2(X)}; \quad (5)$$

4) 重复步骤 2) 和 3) 直到产生 g 个初始聚类中心 $C = c_1, c_2, \dots, c_g$;

5) 分别计算每个样本点 X 到各个聚类中心 c_j

的距离,把对象分配到距离最近的聚类中,表达式为

$$d_{pj} = \|X_p - c_j\|_2^2; \quad (6)$$

6) 所有对象分配完成后,重新计算 g 个聚类的中心;

7) 与上一次计算得到的 g 个聚类中心比较,如果聚类中心发生变化,转到步骤 5),否则转到步骤 8);

8) 输出聚类结果和 g 个聚类中心点。

图 5 所示为 $g=3,4,5,6,7$ 的聚类结果。图中横、纵坐标轴 w', h' 分别表示目标的 bounding boxes 的宽和高,聚类中心点用圆圈标示。每个子图中的 g 个聚类中心点就是新的 g 个 anchor boxes 的初始宽高维度。

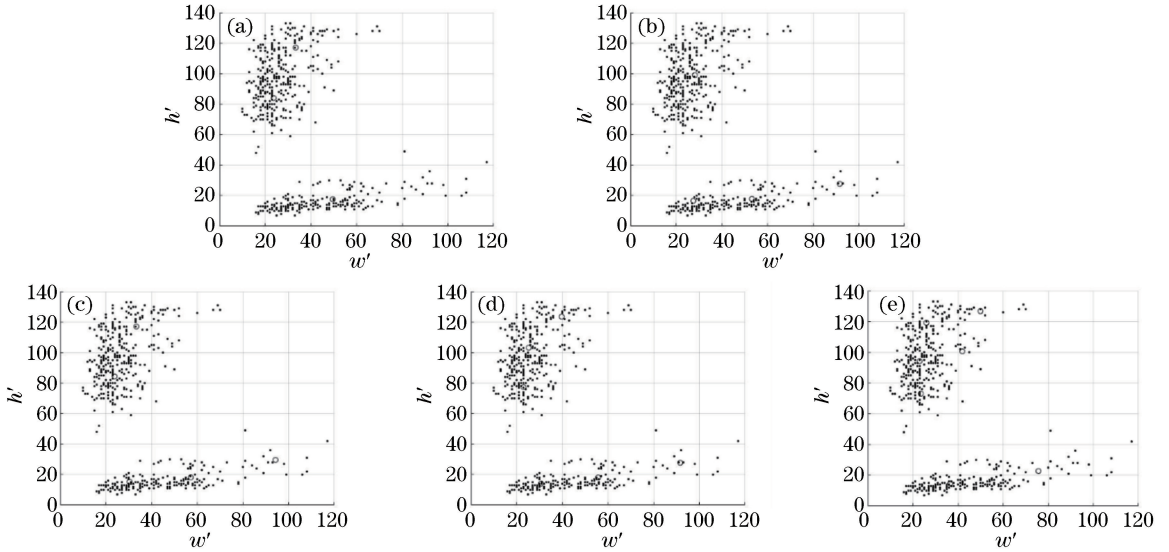


图 5 基于不同 g 值的 K-means++ 聚类结果。(a) $g=3$; (b) $g=4$; (c) $g=5$; (d) $g=6$; (e) $g=7$

Fig. 5 K-means++ clustering results based on different g values. (a) $g=3$; (b) $g=4$; (c) $g=5$; (d) $g=6$; (e) $g=7$

4 实验与分析

4.1 实验环境

使用 Caffe^[17] 框架进行检测,并采用图形处理器(GPU)加速,具体实验环境如表 1 所示。

表 1 实验环境

Table 1 Experimental environment

Name	Type
System	LINUX64 Ubuntu14.04
Frame	Caffe
Language	Python,C++,Protobuf
CPU	Intel Core i7-7700
GPU	GTX1080Ti
Memory/GB	11
RAM/GB	16
Hard disk/GB	250

4.2 实验数据集

采用多台摄像机从不同角度对子弹外观进行拍摄,得到大量的子弹外观原始图片集,从中筛选出具有壳口不齐和口部裂缝缺陷的图片。将图片格式统一为 bmp,图片名统一为 6 位长度的数字,例如,000001.bmp,序号连续。

按照 PASCAL VOC 数据集格式标准构建子弹

缺陷检测数据集,利用 labelImg 软件在图片上标注画框,标注后的数据集样例如图 6 所示,并为每张图片自动生成一个 xml 文件,记录图片的各种信息(文件名、存储路径、宽度、高度、深度、目标信息)以及 bounding box 坐标信息(左上、右下点)等。

训练数据集共 1024 张,分为 2 类,有标签;测试数据集共 256 张,分为 2 类,无标签。

4.3 实验结果与分析

实验采用平均精度均值(mAP)^[18] 作为子弹缺陷检测性能的精度评价指标,其数值越高,表示算法的检测精度越高。

分别使用基于 ZFNet、VGG_CNN_M_1024 和 VGG16 网络的 Faster R-CNN 模型对子弹缺陷数据集进行检测,检测结果如表 2 所示。比较 3 种网络结构可以发现,使用小型的 ZFNet 训练时,mAP 为 91.09%,使用中型的 VGG_CNN_M_1024 训练时,mAP 也只达到 93.82%,而使用大型的 VGG16 训练时,mAP 能达到 96.37%。可以看出,采用较深层网络 VGG16 的 Faster R-CNN 模型的检测精度高于其他两种模型。由此得出,卷积神经网络模型越深,检测精度越高。后续实验都将采用 VGG16 网络进行图像的特征提取。

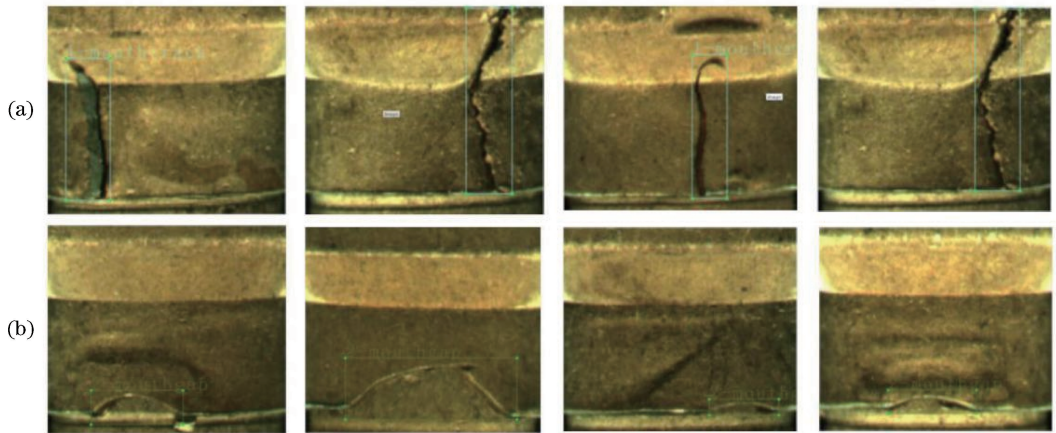


图6 子弹外观缺陷数据集样例。(a)口部裂缝;(b)壳口不齐

Fig. 6 Examples of bullet appearance defect dataset. (a) Mouthcrack; (b) mouthgap

表2 基于不同卷积网络的 Faster R-CNN 模型
检测结果比较Table 2 Comparison of detection results of Faster R-CNN
model based on different convolutional networks

Method	mAP / %
Faster R-CNN+ZFNet	91.09
Fater R-CNN+ VGG_CNN_M_1024	93.82
Fater R-CNN+ VGG16	96.37

将 Faster R-CNN 结合 K -means++ 的改进模型在子弹缺陷检测数据集上进行测试,不同 k 值下子弹缺陷检测精度的实验结果如表 3 所示。

表3 不同 anchor 生成方法测试结果对比

Table 3 Comparison of test results of different
anchor generation methods

Method	Parameter	mAP / %
Anchor boxes	$k=9$	96.37
K -means++	$g=3$	90.03
K -means++	$g=4$	94.68
K -means++	$g=5$	97.24
K -means++	$g=6$	98.06
K -means++	$g=7$	98.02

由表 3 可以看出,使用 Faster R-CNN 目标检测算法进行检测,检测精度为 96.37%,而使用结合 K -means++ 的 Faster R-CNN 改进模型,当 $g=5$ 时,检测精度就优于 Faster R-CNN 目标检测算法。结果表明,使用改进后的 anchor 生成方法,生成的 anchor 更具有代表性,提升了子弹外观缺陷的检测精度。选用 $g=6$ 时,结合 K -means++ 的改进模型的 mAP 提升至 98.06%。

Faster R-CNN 目标检测算法在非极大值抑制后选用了 300 个 anchor 进行检测^[11],本文在 K -means++ 改进 anchor 生成方法的基础上,调整非

极大值抑制后选用的 anchor 数目为 100,测试结果如表 4 所示。

表4 不同 anchor 数目测试结果对比

Table 4 Comparison of test results of different
anchor numbers

Method	Parameter	Proposals	mAP / %	Speed / (frame · s ⁻¹)
Anchor boxes	$k=9$	300	96.37	17
K -means++	$g=6$	300	98.06	19
K -means++	$g=6$	100	97.75	28

Faster R-CNN 目标检测算法的检测精度为 96.37%,速度为 17 frame · s⁻¹。改进 anchor 生成方法,但不调整非极大值抑制后 anchor 的数目,检测精度为 98.06%,速度为 19 frame · s⁻¹。既改进 anchor 生成方法,又调整 anchor 数目为 100 后,检测精度为 97.75%,速度为 28 frame · s⁻¹,虽然精度相比 98.06%略有下降,但是检测速度极大提升,检测效果较好。图 7 为基于改进后 Faster R-CNN 模型的部分检测效果图,图 7(a)为口部裂缝检测效果图,图 7(b)为壳口不齐检测效果图。图 7 中方框代表模型所检测出缺陷的位置,方框上方文字为模型对于方框区域目标的分类结果:子弹缺陷分类和类别得分(属于此子弹缺陷类别的概率值)。

5 结 论

针对传统机器学习方法在子弹外观缺陷检测中手工设计特征耗时和泛化能力差的问题,提出了一种基于 Faster R-CNN 的子弹外观缺陷检测方法,利用深度卷积网络自动提取目标特征,有效避免了传统方法依赖手工设计缺陷特征的问题。将检测模型分别与 3 种卷积神经网络结合进行对比实验,结果

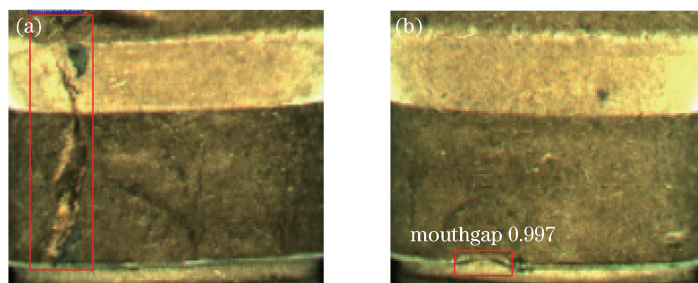


图7 基于改进后的 Faster R-CNN 模型的部分检测结果图。(a)口部裂缝;(b)壳口不齐

Fig. 7 Partial test results based on improved Faster R-CNN model. (a) Mouthcrack; (b) mouthgap

表明,VGG16 相对于 ZFNet 和 VGG_CNN_M_1024 的检测效果更佳。在 Faster R-CNN 目标检测算法的基础上结合 K-means++ 聚类方法改进 anchor 的生成方法,并且调整检测中 anchor 的数量,提高了子弹外观缺陷检测的精度和效率。由于子弹外观缺陷数据集的获取较为困难,目前只对已有的子弹外观缺陷数据集验证了算法的性能,未来将获取包含更多类别的子弹外观缺陷数据集,以便进一步研究。

参 考 文 献

- [1] Zhang Y J, Qi X L, Dong C J. Surface detection of gun steel core based on image gray value morphology under machine vision [J]. Journal of Sichuan Ordnance, 2012, 33(6): 59-61.
张亚军, 齐杏林, 董成基. 机器视觉下基于图像灰度值形态学的枪弹钢芯表面检测[J]. 四川兵工学报, 2012, 33(6): 59-61.
- [2] Shi J W, Guo C Y, Liu H N. Study on detection system of bullet surface defect based on machine vision [J]. Modular Machine Tool & Automatic Manufacturing Technique, 2013(4): 59-64.
史进伟, 郭朝勇, 刘红宁. 基于机器视觉的枪弹外观缺陷检测系统研究[J]. 组合机床与自动化加工技术, 2013(4): 59-64.
- [3] Shi J W, Guo C Y, Liu H N, *et al.* Bullet surface defect extraction based on Hough transformation and two-peak algorithm [J]. Fire Control & Command Control, 2013, 38(5): 129-132.
史进伟, 郭朝勇, 刘红宁, 等. 基于 Hough 变换和双峰法的枪弹外观缺陷提取[J]. 火力与指挥控制, 2013, 38(5): 129-132.
- [4] Wang P, Guo C Y, Liu H N. Research on automatic online detection system of bullet surface defect [J]. Journal of Ordnance Engineering College, 2015, 27(4): 50-53.

王鹏, 郭朝勇, 刘红宁. 枪弹外观缺陷在线自动检测系统研究[J]. 军械工程学院学报, 2015, 27(4): 50-53.

- [5] Wang P, Guo C Y, Liu H N. Bullet surface defect recognition and classification based on support vector machine [J]. Computer Engineering & Science, 2016, 38(9): 1943-1949.
王鹏, 郭朝勇, 刘红宁. 基于支持向量机的枪弹外观缺陷识别与分类[J]. 计算机工程与科学, 2016, 38(9): 1943-1949.
- [6] Wang L, Zhang H H. Application of faster R-CNN model in vehicle detection [J]. Journal of Computer Applications, 2018, 38(3): 666-670.
王林, 张鹤鹤. Faster R-CNN 模型在车辆检测中的应用[J]. 计算机应用, 2018, 38(3): 666-670.
- [7] Zhang H Y, Wang S N, Hu W B. Improved method for estimating number of people based on convolution neural network [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121503.
张红颖, 王赛男, 胡文博. 改进的基于卷积神经网络的人数估计方法[J]. 激光与光电子学进展, 2018, 55(12): 121503.
- [8] Tian Q, Yuan T Y, Yang D, *et al.* A pedestrian detection method based on dark channel defogging and deep learning [J]. Laser & Optoelectronics Progress, 2018, 55(11): 111007.
田青, 袁瞳阳, 杨丹, 等. 基于暗通道去雾和深度学习的行人检测方法[J]. 激光与光电子学进展, 2018, 55(11): 111007.
- [9] Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 580-587.
- [10] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision

- (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1440-1448.
- [11] Ren S Q, He K M, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [12] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks [M] // Fleet D, Pajdla T, Schiele B, *et al.* Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8689: 818-833.
- [13] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J/OL]. (2015-04-10)[2018-12-05]. <https://arxiv.org/abs/1409.1556>.
- [14] Li J N, Zhang B H. Face recognition by feature matching fusion combined with improved convolutional neural network [J]. Laser & Optoelectronics Progress, 2018, 55(10): 101504. 李佳妮, 张宝华. 特征匹配融合结合改进卷积神经网络的人脸识别[J]. 激光与光电子学进展, 2018, 55(10): 101504.
- [15] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [16] Everingham M, Eslami S M A, van Gool L, *et al.* The Pascal visual object classes challenge: a retrospective[J]. International Journal of Computer Vision, 2015, 111(1): 98-136.
- [17] Jia Y Q, Shelhamer E, Donahue J, *et al.* Caffe[C] // Proceedings of the ACM International Conference on Multimedia, November 3-7, 2014, Orlando, Florida, USA. New York: ACM, 2014: 675-678.
- [18] Feng X Y, Mei W, Hu D S. Aerial target detection based on improved faster R-CNN [J]. Acta Optica Sinica, 2018, 38(6): 0615004. 冯小雨, 梅卫, 胡大帅. 基于改进 Faster R-CNN 的空中目标检测 [J]. 光学学报, 2018, 38(6): 0615004.