

基于深度学习的齿轮视觉微小缺陷检测^①



韩明¹, 吴庆祥^{1,2}, 曾雄军¹

¹(福建师范大学 光电与信息工程学院, 福州 350007)

²(医学光电科学与技术教育部重点实验室, 福州 350007)

通讯作者: 吴庆祥, E-mail: qxwu@fjnu.edu.cn

摘要: 针对齿轮视觉微小缺陷, 采用一种基于深度学习算法的 Mask R-CNN 网络进行检测, 并对网络进行相应地优化调整. 首先, 通过比较 5 种残差神经网络检测效果, 选择 resnet-101 作为图像共享特征提取网络. 然后, 剔除特征金字塔网络中对特征图 P_5 进行的不合理的 3×3 卷积, 缺齿检出率指标相应得到提升. 最后, 为了对候选区域网络进行有效的训练, 根据设计的样本标注方案中小范围波动的标注尺寸, 设置合适的 anchors 大小以及宽高比. 最终, 经过优化的 Mask R-CNN 网络达到了 98.2% 缺齿检出率.

关键词: 深度学习; 微小缺陷; 齿轮; 残差神经网络; 特征金字塔网络

引用格式: 韩明, 吴庆祥, 曾雄军. 基于深度学习的齿轮视觉微小缺陷检测. 计算机系统应用, 2020, 29(3): 100-107. <http://www.c-s-a.org.cn/1003-3254/7323.html>

Visual Detection of Minor Gear Defect Based on Deep Learning

HAN Ming¹, WU Qing-Xiang^{1,2}, ZENG Xiong-Jun¹

¹(College of Photonic and Electronic Engineering, Fujian Normal University, Fuzhou 350007, China)

²(Key Laboratory of OptoElectronic Science and Technology for Medicine (Ministry of Education), Fuzhou 350007, China)

Abstract: The optimized Mask R-CNN network based on deep learning is used to visual detection of the tiny defects on gears. Firstly, by comparing the detection effects of five kinds of residual neural network, resnet-101 is selected as the image sharing feature extraction network. Then, the detection rate for missing tooth is correspondingly improved by eliminating the unreasonable 3×3 convolution of feature map P_5 in the feature pyramid network. Finally, in order to effectively train the region proposal network, the appropriate anchor size and aspect ratio are set according to small fluctuation of annotation box in the designed sample labeling scheme. The optimized Mask R-CNN network eventually achieved 98.2% detection rate for missing tooth on gears.

Key words: deep learning; tiny defects; gears; residual neural network; feature pyramid network

1 引言

随着人工智能的不断发展, 人工智能技术被广泛地应用在各个领域, 不断推进社会的进步^[1]. 其中, 深度学习领域的发展, 加快了人工智能的步伐. 深度学习属于机器学习领域的一个新的研究分支, 该算法可分为监督式学习、非监督式学习、半监督式学习、以及强化学习^[2]. 其中, 监督式学习是指通过人工标记的数据

来推测出目标函数. 目前的分类, 目标检测等任务大多均采取的是监督式的学习方式. 基于深度学习算法的 Mask R-CNN 网络亦是通过对监督式的学习方式来对网络权重进行迭代更新.

近年来, 针对齿轮齿形缺陷检测已存在一些检测方案. 2019 年, 李凯等通过将齿轮图像二值化来获取齿轮中心点坐标, 利用中心点坐标构建环形滤波器获取

① 基金项目: 中央引导地方科技发展专项 (2017L3009)

Foundation item: Special Fund of Central Government for Local Science and Technology Development (2017L3009)

收稿时间: 2019-08-06; 修改时间: 2019-09-05; 采用时间: 2019-09-18; csa 在线出版时间: 2020-02-28

只含轮齿部分的感兴趣区域,最终通过对灰度均值小于阈值的轮齿进行细分来判定该轮齿是否为缺齿^[3]. 2018年,杨亚等通过检测得到的关键点将标准齿轮与待检测齿轮进行配准,对配准图像采用绝对差分的方式得到绝对差分图像,通过绝对差分图像判定待检测齿轮是否存在缺齿^[4]. 2017年,郑硕文等构建基于机器视觉与多传感器的齿轮缺陷检测系统^[5]. 上述的齿轮缺陷检测方法以及张钰婷、李盼等^[6,7]更早提出的缺陷检测方案大多是围绕传统的图像处理算法来进行展开的. 采用深度学习算法实现齿轮缺陷检测的特点在于神经网络可以自主对齿轮图像进行特征提取,通过将提取到的低层特征进行组合形成高层具有较深语意的特征,使得网络模型可以自动学习到从样本图像输入到结果输出的复杂映射关系^[8]. 鉴于 Mask R-CNN^[9]网络具有对微小目标进行有效检测的优势,在原网络基础上,针对齿轮微小视觉缺陷检测目标对原网络进行改进,提高了 Mask R-CNN 网络对于缺齿对象的召回率指标. 基于改进的 Mask R-CNN 齿轮缺齿检测方法可应用到对流水线上齿轮视觉缺陷或者类似的产品进行检测,具有广阔的应用场景.

2 Mask R-CNN 网络架构及优化

Mask R-CNN 架构(见图 1)主要由两部分构成,第一部分为区域候选网络,用于获取候选框. 第二部分利用 RoIAlign 网络层对感兴趣区域进行池化操作,产生固定尺寸的特征图,后续网络层通过对该固定尺寸特征图进行处理,实现对目标的分类、检测、以及分割. 本文通过理论分析并结合实验结果对共享特征提取过程进行优化,并针对标注方案中标注框的大小对候选区域网络中的锚框大小进行重新设计.

(1) 共享特征提取网络

Mask R-CNN 的共享特征提取网络采用残差神经网络与特征金字塔网络的组合,对输入图像进行特征提取.

1) 残差神经网络

为了尽可能提高网络的缺齿检出率,同时兼顾网络的训练与检测速度,应优选一种残差神经网络. 通用的残差神经网络^[10]主要有 5 种,分别为: ResNet-18、ResNet-34、ResNet-50、ResNet-101、ResNet-152,缺齿检出率公式表达如下:

$$\text{缺齿检出率} = \frac{\text{检出缺齿的总个数}}{\text{缺齿的总个数}} \times 100\% \quad (1)$$

在通用的目标检测评价指标中,一般采用召回率与精确率是两个指标,计算方法如表 1.

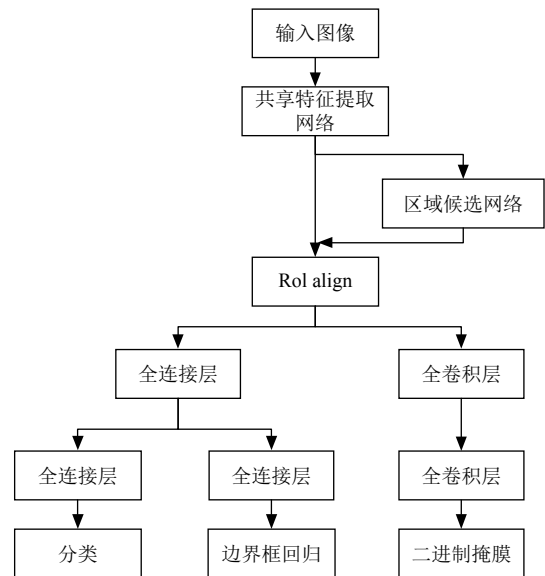


图 1 Mask R-CNN 结构

表 1 精确率与召回率

	Positive (P)	Negative (N)
True (T)	True Positive (TP)	True Negative (TN)
Flase (F)	Flase Positive (FP)	Flase Ngtive (FN)

表 1 中, P 表示缺陷轮齿对象, N 表示为正常轮齿对象, T 与 F 表示检测结果是否正确. TP : 缺陷轮齿对象被正确识别为缺陷轮齿对象的数量; FP : 正常轮齿对象被错误识别为缺陷轮齿对象的数量; FN : 缺陷轮齿对象被错误识别为正常轮齿对象的数量. 精确率与召回率表达式如下:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

由于检测目标为缺陷轮齿,故将缺陷轮齿视为正样本. 召回率与缺齿检出率两个指标在数值上是完全相等的,其中, TP 等于检出缺齿的总个数, TP 与 FN 之和等于缺齿的总个数. 由于训练的样本标注类别只有缺齿这一类,因此 FP 的值将趋近于 0, Precision (精确度) 将趋近于 1 这个常数. 故实验从通用的目标检测评价指标角度,还可使用召回率来对网络的检测效果进

行评价。

实验比较 5 种通用残差神经网络作为共享特征提取网络的检测效果, 采用迁移学习方式对 Mask R-CNN 网络进行训练, 网络初始权重为基于 coco 数据集训练得到的权重. 训练集数量为 200 张, 缺齿标注个数为 500 个, 测试集的数量为 150 张, 总齿数为 1800 个, 缺齿数为 280 个, 检测结果如表 2.

表 2 5 种残差神经网络检测结果比较

	TP/检出缺齿个数	召回率/缺齿检出率 (%)
ResNet-18+FPN	232	0.829/82.9
ResNet-34+FPN	246	0.879/87.9
ResNet-50+FPN	251	0.896/89.6
ResNet-101+FPN	259	0.925/92.5
ResNet-152+FPN	263	0.939/93.9

由以上检测结果可以得出, 随着残差神经网络深度的增加, 召回率表现为递增, 更深的网络同时需要更多的训练时间, 且测试时间也相应增加. 实验结果显示, ResNet-101 与 ResNet-152 召回率指标的差距相对不是很大, 综合召回率与网络耗时考量, 选择使用 ResNet-101 作为特征提取网络.

2) 特征金字塔网络

高层特征图虽然含有较深语意的特征表达, 其分辨率却很低, 当对小尺度目标进行特征提取时, 重要的特征信息容易丢失. 特征金字塔网络^[11]利用自上而下通道以及侧向连接将高层高语意低分辨率特征图与低层低语意高分辨率的特征图进行融合, 从而获得不同尺度兼具高语意与高分辨率的特征图. 图像共享特征提取过程如图 2, ResNet-101 具有 5 个阶段, 每个阶段是根据输出特征图的大小是否改变进行划分. 分别将各个阶段最后网络层输出的特征图称为 $\{C_1, C_2, C_3, C_4, C_5\}$, 图 2 中, C_5 是经过 ResNet-101 第五阶段生成的特征图, 也是网络学习到的具有最深语意的特征图, 该特征图未进行融合操作, 分辨率较低.

特征金字塔网络将 C_5 进行 1×1 卷积操作得到 P_5 , 对 P_5 进行上采样操作与 C_4 经过 1×1 卷积得到的特征图进行融合, 得到 P_4 , P_4 既具有 P_5 的深层语意而且还保留了 C_4 中微小对象较高分辨率特征. P_3, P_2 的获取方式与 P_4 相同. 最后, 还需对融合得到的特征图进行一次 3×3 卷积操作, 目的是降低对特征图进行上采样操作之后的混叠效应. 在此, 实验对用于后续网络进行预测的特征图进行优化调整, 剔除对 P_5 进行的 3×3 卷

积操作, 如图 3, 分析如下: ① P_5 并非是经过上采样与融合得到, 既对 P_5 而言不存在混叠效应. ② 该 3×3 卷积操作采用“padding=same”的方式, 即卷积前后特图的尺寸不发生改变, 剔除该 3×3 卷积对后续特征图尺寸匹配不会造成影响. ③ 经实验测试, 该 3×3 卷积操作的剔除相应地提高了网络的缺齿检出率指标.

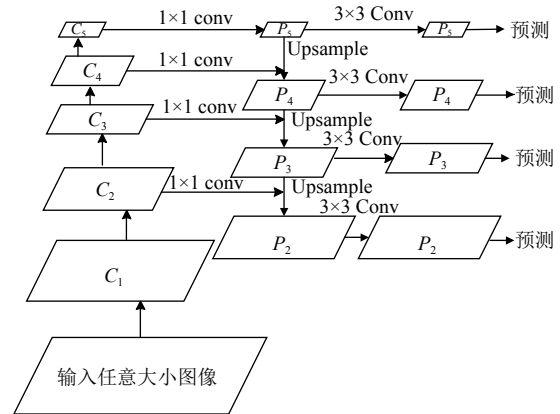


图 2 图像特征提取过程

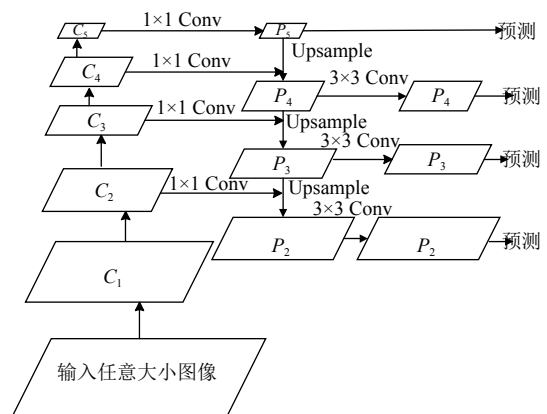


图 3 优化的图像特征提取过程

(2) 候选区域网络

候选区域网络^[12]实质是一个基于滑窗的无类别目标检测器. 用于生成目标候选框. 候选区域网络始应用于 Faster R-CNN 中, 作用对象为单尺度特征图. 具体操作: 将一个 3×3 的滑窗置于单尺度特征图上进行滑动, 在滑窗滑过的位置上以滑窗的中心作为锚点, 产生 k 个锚框. Faster R-CNN 中, k 参数为 9, 以像素为单位, 大小为 $\{128^2, 256^2, 512^2\}$, 在该 3 种尺寸的基础上分配 $\{1:2, 1:1, 2:1\}$ 宽高比. 通过计算锚框与标注框的 IoU 值来判断锚框与标注框的重叠度, 数值越大表示重

叠度越高,最大值为1,其计算公式如下:

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (4)$$

即锚框与标注框交集和锚框与标注框并集的比值.

Mask R-CNN 中候选区域网络的作用对象为由特征金字塔网络生成的具有金字塔层次的多尺度特征图,分别为 $\{P_2, P_3, P_4, P_5, P_6\}$,其中 P_6 是对 P_5 进行降采样得到的,尺寸缩小了一倍.与Faster R-CNN不同的是,对每个层次的特征图仅分配一种尺寸的锚框,这样可以有效减少冗余锚框的数量.以像素为单位,Mask R-CNN 网络将 $\{P_2, P_3, P_4, P_5, P_6\}$ 锚框的尺寸分别设置为 $\{32^2, 64^2, 128^2, 256^2, 512^2\}$,每个尺寸的锚框均分配以 $\{1:2, 1:1, 2:1\}$ 的宽高比,故每个层次特征图的锚框的尺寸种数为3种,5种具有金字塔层次特征图的锚框尺寸的种数为15种,采用式(4)来计算锚框与标注框的IoU数值.

实验检测对象为有缺陷的轮齿,齿轮样本如图4.



图4 齿轮样本

当齿轮旋转以及缺陷存在于轮齿位置的不同使得对有微小缺陷的轮齿进行标注的尺寸会发生改变,即标注框的尺寸发生改变.但无论齿轮转动多少角度,标注框的尺寸最大值不会超过图5中线段1与线段2的长度,实际标注框的宽与高最大值分别为线段1、线段2长度的正弦、余弦值,通过计算得到,以像素为单位,线段1与线段2的长度为32个像素.根据标注框的尺寸,假设各层次特征图使用锚框的尺寸为 $\{32^2, 64^2, 128^2, 256^2, 512^2\}$,设标注框的尺寸为最大值 32×32 ,则锚框的尺寸均大于等于标注框的尺寸.当标注框包含于锚框时会有最大IoU值,位置关系如图6,各尺寸锚框与标注框的最大IoU值如表3.特别地,当锚框尺

寸为 32×32 ,锚框与标注框完全重合时,IoU数值取得最大值1.

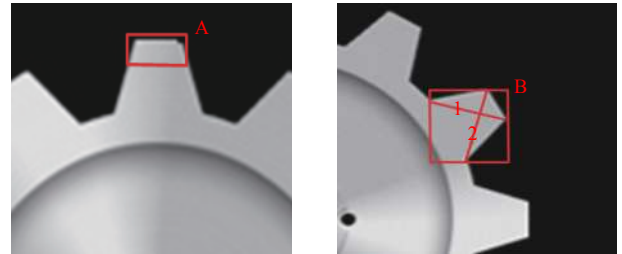


图5 标注尺寸波动示意图

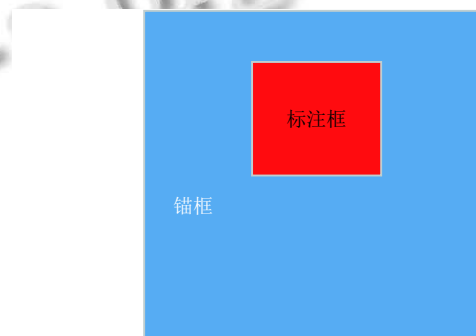


图6 IoU达到最大值时锚框与标注框的包含关系

表3 5种锚框尺寸对应最大IoU数值

特征图	锚框尺寸	锚框宽高比	$IoU_{(\max)}$
P_2	32^2	$\{1:2, 1:1, 2:1\}$	1
P_3	64^2	$\{1:2, 1:1, 2:1\}$	0.5
P_4	128^2	$\{1:2, 1:1, 2:1\}$	0.125
P_5	256^2	$\{1:2, 1:1, 2:1\}$	0.031 25
P_6	512^2	$\{1:2, 1:1, 2:1\}$	0.007 8125

计算各个尺寸锚框与标注框的IoU数值与设定阈值0.7做比较,IoU数值大于0.7的锚框被分配一个正标签,IoU数值小于0.3的锚框被分配一个负标签,介于0.3与0.7之间的锚框将直接被舍弃.由表3可得仅有 P_2 特征图的锚框会被分配正标签,通过对分配正标签的锚框进行修正使得训练候选区域网络的损失值达到最低.训练候选区域网络的损失函数如下:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*) \quad (5)$$

其中, L_{cls} 表达式如下:

$$L_{\text{cls}}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (6)$$

其中, t_i 代表对锚框进行不断修正得到的候选框的 4 个参数化坐标, p_i 代表锚框中含有目标的概率或置信度, i 代表锚框索引. 当锚框被赋予正标签时, p_i^* 为 1, 此时 L_{reg} 函数值对损失函数值 L 有效, 通过反向传输与梯度下降的优化方法, 对 t_i 进行不断修正, 最终得到较理想的候选框坐标, 使得损失值 L 降到最低; 当锚框均被赋予负标签时, p_i^* 为 0, L_{reg} 函数值对损失函数值 L 无效, 即 t_i 不影响 L 值, 此时将式 (6) 代入式 (5) 中得:

$$L = \frac{1}{N_{cls}} \sum_i [-\log(1 - p_i)] \quad (7)$$

式中, N_{cls} 为常数, 此时的 L 随 p_i 同增同减, 为使 L 优化为最低, p_i 应为最低, 此时锚框 i 均被视为背景来对候选区域网络进行训练.

根据以上推导, 针对标注框可能的最大尺度 32×32 (图 5), 仅有尺寸为 32×32 的锚框会被分配正标签, 但当标注框的尺寸为图 5 中 A 框 (20×10) 时, 此时使用 32×32 的锚框得到最大的 IoU 值为 0.4, 既 A 框中对象即使参与了候选区域网络的训练, 也是以一种不存在缺齿对象的身份来对候选区域网络训练的, 这种问题导致的最终结果是造成 Mask R-CNN 网络更高的漏检率. 针对以上问题, 将图 5 中标注框 A 框的尺寸设为锚框的最小尺寸, 将图 5 中线段 1 与线段 2 的长度设为锚框的最大尺寸, 以像素为单位, A 框尺寸近似为 20×10 , 则以 1:1 的锚框宽高比, 使用的最小锚框尺寸 20×20 , 最大锚框尺寸为 32×32 . 同样对每个尺寸的锚框分配 {1:2, 1:1, 2:1} 的宽高比, 则各个层次特征图锚框尺寸分配情况如表 4, 选择表 4 中的锚框尺寸提高了可容忍标注框尺寸的下限, 从而实现对候选区域网络进行更加有效的训练.

表 4 根据标注框尺寸设计相应的锚框尺寸

特征图	锚框尺寸	锚框宽高比
P_2	20^2	{1:2, 1:1, 2:1}
P_3	23^2	{1:2, 1:1, 2:1}
P_4	26^2	{1:2, 1:1, 2:1}
P_5	29^2	{1:2, 1:1, 2:1}
P_6	32^2	{1:2, 1:1, 2:1}

(3) 多任务训练

采用多任务训练机制, 利用 Mask R-CNN 的掩膜分支损失函数, 使得微小目标检测效果更加精确. 多任务训练损失函数公式定义如下:

$$L = L_{cls} + L_{box} + L_{mask} \quad (8)$$

式中, L_{cls} 为分类损失函数, L_{box} 为边框损失函数, L_{mask} 为掩膜损失函数.

3 制作数据集与样本标注方案设计

(1) 数据集制作

目前缺乏公共的齿轮数据集, 且数量较少. 因此采用数据扩增的手段来建立数据集. 采用的数据扩增方式为对缺齿齿轮图像平移、360 度随机旋转、添加随机椒盐噪声. 对图像进行平移与随机旋转是为了实现 Mask R-CNN 网络能够对图像内任意位置、旋转任意角度的缺齿齿轮进行有效的检测, 添加随机椒盐噪声是为了提高网络对含有椒盐噪声干扰的齿轮图像的鲁棒性. 本实验齿轮图像数据集统一为 273×265 (像素) RGB 图像, 含有缺齿的样本展示如图 7.

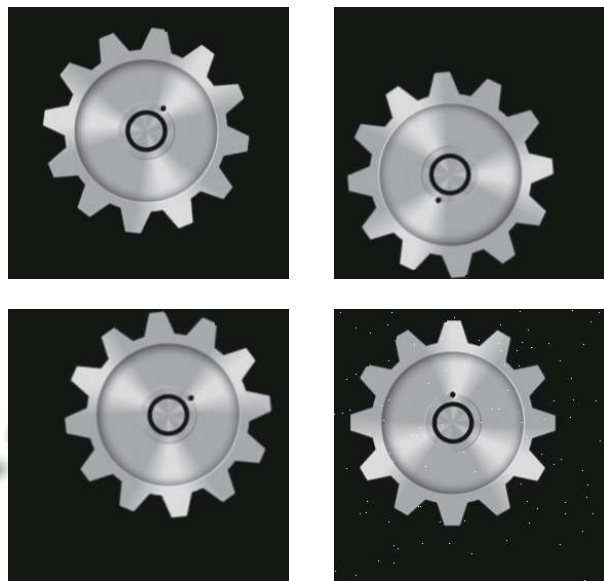


图 7 缺齿样本展示

(2) 标注方案设计

训练集使用 VGG Image Annotator (VIA) 进行标注, 标注完另存后缀名为 json 文件. 在进行齿轮图像标注时, 设计两种标注方案, 针对不同的标注方案使用对应的锚框尺寸. 方案 1: 对整个齿轮对象进行标注, 然后二分类, 既对标注对象分类为合格齿轮与不合格齿轮. 不合格的齿轮是指带有缺齿等缺陷的齿轮 (如图 8). 方案 2: 直接对齿轮缺陷部分进行标注 (如图 8). 方案 1 的优点在于可忽略具体缺陷类型, 分类类别单一, 但是

利用方案 1 训练得到的 Mask R-CNN 网络在精度判别方面存在一个问题, 对于有微小缺陷的齿轮, 其判别能力较差. 具体来说, 一个有几个像素缺齿的齿轮, 其对于整个齿轮而言是微不足道的, 从而导致微小缺陷容易被忽略, 将不合格齿轮判定为合格. 采用方案 2 对缺陷部分进行标注, 就好比在提醒网络更应该关注微小缺陷本身, 使得处于像素级别缺陷可以被检测出来. 两种标注方案在设计理念上各有优缺点, 只是方案一的标注思路使得带有像素级别缺齿的齿轮与合格齿轮相比并无明显差别, 召回率较低, 且标注工作量更大, 而方案 2 相对于方案 1 而言应当注意的是, 当待检测工件缺陷类型较多时, 需要对缺陷类别进行逐一分类, 既多分类. 综合样本标注工作量以及网络检测效果考量, 针对微小缺陷对象, 实验采用第 2 种样本标注方案.

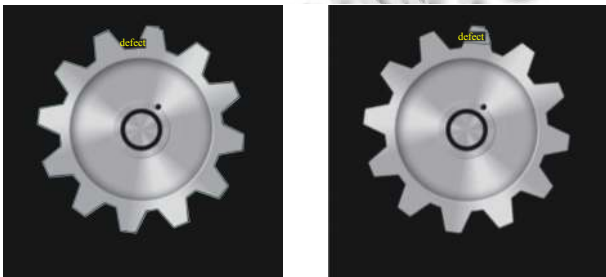


图 8 缺齿样本标注

4 实验与分析

特征提取网络采用 ResNet-101 与 FPN 的组合, 加载基于 coco 数据集预训练好的网络模型权重文件, 使用经过样本扩增的齿轮数据集对网络模型进行迁移学习. 实验训练集为 200 张微小缺齿图像, 图像规格统一为 273×265 的三通道 RGB 图像, 每张齿轮图像至少有一个带有缺陷的轮齿, 同时部分齿轮图片含有多个不同尺度的微小缺齿, 保证训练得到的模型能同时对多种尺度范围的微小缺齿进行检测.

使用显卡型号为 1080Ti 的服务器对 Mask R-CNN 网络模型进行训练, 显卡显存 11 GB. Python 版本 3.6.0、Tensorflow-gpu 版本 1.6.0、Keras 版本 2.1.6. 网络一次迭代训练读取两张图像, 一个 epoch 迭代次数为 100 次, 经过一次 epoch 训练之后, 训练集中的 200 张图像对网络模型进行了一次完整训练, 参数 epoch 设置为 500 次, 每训练完一个 epoch, 便将最后一次迭代训练得到的损失值保存在日志中, 经过 500 次

epoch 训练之后, 改进的 Mask R-CNN 网络损失值得到有效收敛, 使用 tensorboard 对损失函数曲线进行显示, 如图 9.

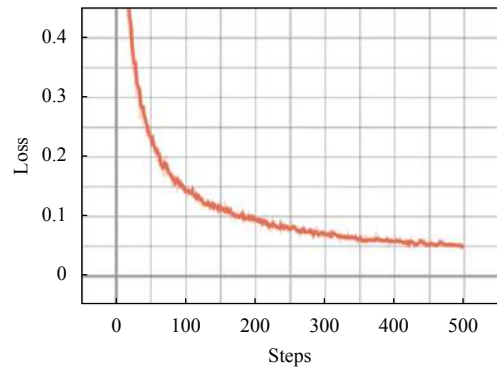


图 9 损失函数曲线

5 实验结果

部分检测结果如图 10, 图中上侧为使用未经调整 Mask R-CNN 网络检测结果, 下侧为使用经过优化改进之后的 Mask R-CNN 网络检测结果. 从图 10 的检测结果可以看出, 通过对图像特征提取过程进行优化以及根据标注框的尺寸设置合适的锚框尺寸, 可以提高召回率, 改善缺齿检出率指标, 实现对采集范围内任意位置的缺齿进行更加有效地检测. 实验目标为实现缺齿目标的检测, 故去除掩膜输出, 得到图 11 中的检测结果. 去除掩膜输出之后更便于对检测对象进行观察. Mask R-CNN 网络改进前与改进后的检测数据比较如表 5. 由表 5 得出, 经过优化的 Mask R-CNN 网络对比未优化 Mask R-CNN 网络, 缺齿检出率提高了 5.7%, 召回率提高了 0.057, 召回率与缺齿检出率分别达到了 0.982 与 98.2%. 利用优化的 Mask R-CNN 网络进行检测仍存在 5 个缺齿漏检, 可对训练集进一步地扩增, 使得缺齿样本覆盖更广的特征空间, 进一步提高召回率与缺齿检出率指标. 同时本方法与传统的图像处理与机器学习算法^[13]进行了比较, 识别率提高了 0.4 个百分点, 比较结果见表 6.

6 结语

采用优化后的 Mask R-CNN 目标检测网络不仅可以对齿轮微小缺齿进行检测, 譬如齿轮划痕、齿面腐蚀、齿面污渍等视觉缺陷均可以采用这一方法. 通过

将优选的残差神经网络与对特征提取过程进行优化的特征金子塔网络进行结合作为共享特征提取网络,使得整个网络模型能够更好地学习到微小对象的特征表达.该方法无需进行大量图像处理工作,有效的避免了传统图像处理方法对微小缺陷图像进行处理过程中造成的误差,且可实现对采集范围内任意位置的齿轮微

小缺陷的检测,提高了检测手段的智能化水平.该方法具有通用性,无需针对某一特定缺陷设定一个算法去解决.基于分割的样本标注过程相对比较繁琐,要求样本标注精确以避免手工标注带来的误差.尽量使用像素比较高的样本,优质的数据集可以训练出检测效果更好网络模型.

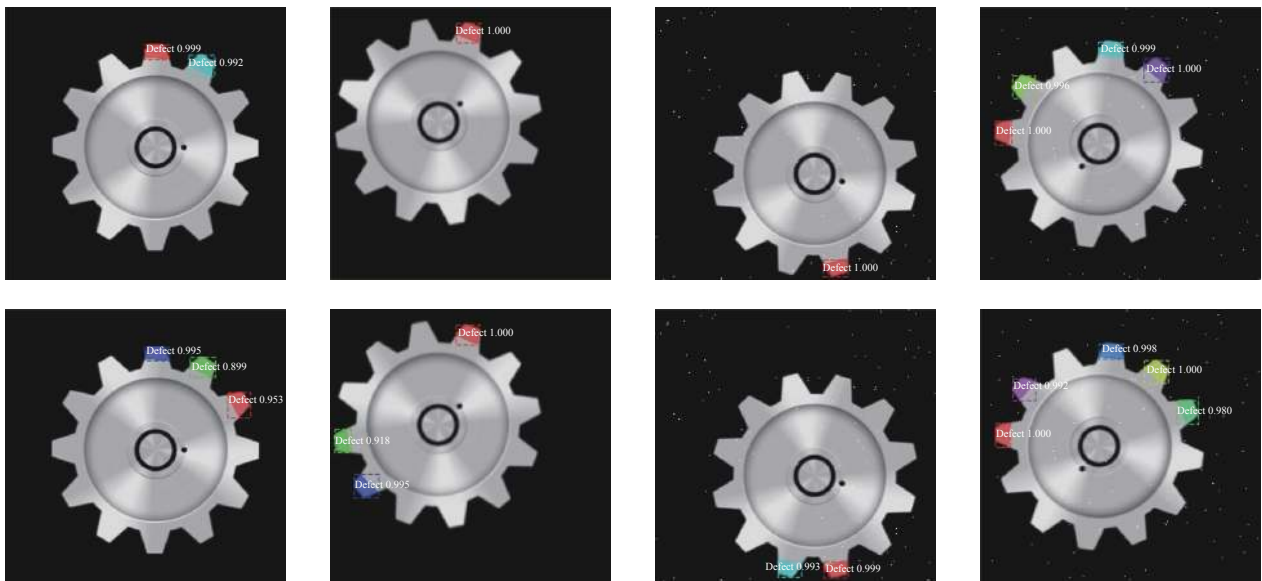


图 10 检测结果比较

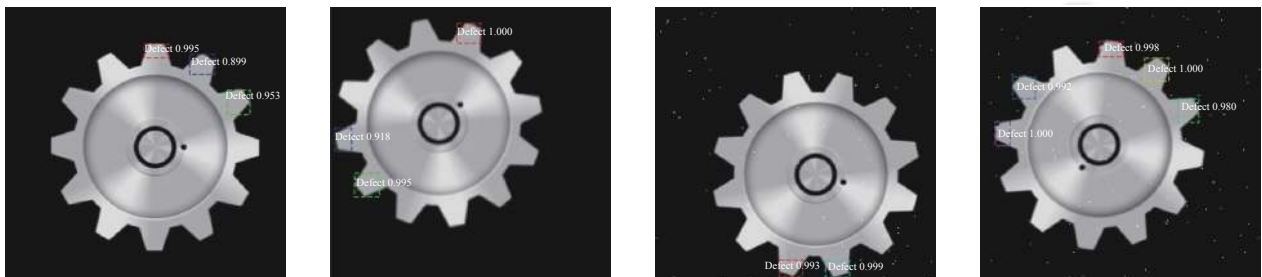


图 11 去除掩膜输出

表 5 检测结果比较

锚框尺寸	缺齿数	检出缺齿的个数	召回率/缺齿检出率 (%)
未优化, {32 ² , 64 ² , 128 ² , 256 ² , 512 ² }	280	259	0.925/92.5
优化后, {20 ² , 23 ² , 26 ² , 29 ² , 32 ² }	280	275	0.982/98.2

表 6 检测方法比较

检测方法	识别率 (%)
基于图像处理与支持向量机方法	97.8
基于改进的 Mask R-CNN 网络方法	98.2

致谢

作者感谢中央引导地方科技发展专项 (2017L3009) 以及福建省光电传感工程研究中心对本研究的支持。

参考文献

- 张妮, 徐文尚, 王文文. 人工智能技术发展及应用研究综述. 煤矿机械, 2009, 30(2): 4-7. [doi: 10.3969/j.issn.1003-0794.2009.02.002]
- 周晋. 关于人工智能的深度学习算法. 电脑编程技巧与维护, 2019, (7): 115-117. [doi: 10.3969/j.issn.1006-4052.2019.

- 07.044]
- 3 李凯, 王化明, 沈玮, 等. 粉末冶金齿轮的视觉缺陷检测方法研究. 机械制造与自动化, 2019, 48(2): 173-176.
 - 4 杨亚, 陶红艳, 余成波. SURF 与灰度差分在小模数塑料齿轮缺陷检测中的研究与应用. 机械传动, 2018, 42(5): 156-160.
 - 5 郑硕文, 刘泓滨, 段先云, 等. 基于机器视觉的齿轮缺陷检测系统设计. 机械制造, 2017, 55(10): 96-97, 103. [doi: [10.3969/j.issn.1000-4998.2017.10.027](https://doi.org/10.3969/j.issn.1000-4998.2017.10.027)]
 - 6 张钰婷, 邵勇, 顾桂鹏, 等. 基于机器视觉的齿轮缺陷检测方法研究. 工业控制计算机, 2017, 30(7): 90-91, 94. [doi: [10.3969/j.issn.1001-182X.2017.07.037](https://doi.org/10.3969/j.issn.1001-182X.2017.07.037)]
 - 7 李盼, 禹翼. 基于图像处理的齿轮缺陷检测. 黑龙江科技信息, 2017, (8): 45.
 - 8 连自锋. 基于深度神经网络的图像识别算法研究[博士学位论文]. 北京: 北京邮电大学, 2017.
 - 9 He KM, Gkioxari G, Dollár P, *et al.* Mask R-CNN. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 2961-2969.
 - 10 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. 2016. 770-778.
 - 11 Lin TY, Dollár P, Girshick R, *et al.* Feature pyramid networks for object detection. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 936-944.
 - 12 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149. [doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031)]
 - 13 贺秋伟, 王龙山, 于忠党, 等. 基于图像处理和支撑向量机的微型齿轮缺陷检测. 吉林大学学报(工学版), 2008, 38(3): 565-569.