

基于改进 SRGAN 和 GoogLeNet 模型的安全帽佩戴安全性智能检测

张 旺¹

ZHANG Wang

摘 要

在变电站作业中,正确佩戴安全帽可以保障工作人员的人身安全,利用摄像头采集数据可以检测工作人员是否正确佩戴安全帽,但是拍摄的图像存在分辨率低、含噪成分严重和图像模糊等问题,严重影响安全帽佩戴安全性智能检测的效果。针对该问题,提出一种结合改进的超分辨率生成式对抗网络(SRGAN)和 GoogLeNet 模型的安全帽佩戴安全性检测。首先,引入注意力机制、残差密集块网络、渐进式上采样和 Wasserstein 距离,使得 SRGAN 网络能够更好地获得图像的高频特征,生成高分辨率图像;其次,利用生成的高清图像训练已预训练的 GoogLeNet 模型,对 GoogLeNet 网络参数进行微调,重新训练后的 GoogLeNet 模型可以自主挖掘图像的特征;最后,在 Faster RCNN 框架下实现图像中安全帽佩戴安全性智能检测。实验结果表明,利用改进的 SRGAN 模型得到的高分辨率图像在主观视觉和客观评价指标上均优于现有方法,安全帽佩戴安全性智能检测的准确率可以达到 98.10%,召回率可以达到 98.30%。所提出的方法具有较好的智能检测效果,并且具有一定的泛化能力与鲁棒性。

关键词

安全帽;生成式对抗网络;GoogLeNet 模型;智能检测

doi: 10.3969/j.issn.1672-9528.2024.10.049

0 引言

变电站工作人员若不按照规定佩戴安全帽,将会存在严重的安全隐患^[1]。利用无人机或巡检机器人的摄像头可以采集图像数据。若采用人工判断工作人员是否违规,存在以下三个缺点:(1)耗费大量的人力财力;(2)监控人员长时间监控现场容易出现疲劳,影响检测效果;(3)人工判断具有一定的主观性,缺乏客观评价。

深度学习方法可以通过构建深层网络自主挖掘图像的特征^[2]。近年来,该方法在智能巡检、目标检测、图像分割和自然语言处理等领域得到了广泛的应用。许多学者利用深度学习实现安全帽佩戴与否检测。文献[3]采用数据增强方式平衡样本,利用 YOLO-S 模型实现安全帽检测。文献[4]结合网格搜索算法和交叉验证改进了随机森林算法,并将其应用到了安全帽检测上。文献[5]采用混合高斯模型实现前景检测,通过连通域检测定位到人体头部实现安全帽佩戴与否检测。文献[6]利用 OpenPose 姿态估计模型确定头部区域,利用 RetinaNet 模型检测安全帽是否正确佩戴。上述方法将深度学习引入安全帽佩戴与否检测中,具有一定的效果,但是没有考虑拍摄的图像存在分辨率低、含噪成分严重和图

像模糊等问题,而这些问题会严重影响深度学习算法的检测效果。

本文提出一种结合改进的超分辨率生成式对抗网络(super resolution generative adversarial network, SRGAN)和 GoogLeNet 模型的安全帽佩戴安全性检测。在 SRGAN 中引入注意力机制、残差密集块网络、渐进式上采样和 Wasserstein 距离,提高网络的性能。利用改进的 SRGAN 模型生成高分辨率图像作为 GoogLeNet 模型的输入,实现图像特征自主挖掘,在 Faster RCNN 框架下完成安全帽的自动标注。该方法可以在图像低分辨率、含噪成分严重的情况下实现安全帽检测,对于没有正确佩戴安全帽的工作人员发出警告,为变电站安全智能检测提供了技术保障。

1 超分辨率图像重建模型

2014 年,Goodfellow 等人^[7]提出了生成式对抗网络(generative adversarial network, GAN)模型。GAN 利用生成器与判别器之间的相互博弈,通过生成器生成逼真图像^[8]。GAN 的结构图如图 1 所示。生成器的目标函数为:

$$V_G = E_{z \sim p_z(z)}[\log \{1 - D(G(z))\}] \quad (1)$$

判别器的目标函数为:

$$V_D = E_{x \sim p_d(x)}[\log D(x)] + E_{z \sim p_z(z)}\{\log(1 - D(G(z)))\} \quad (2)$$

1. 国网吉林省电力有限公司长春供电公司 吉林长春 130000

求式(1)的最小值和式(2)的最大值^[9],将式(1)和式(2)合并,得到 GAN 整体训练过程的目标函数为:

$$\min_G \max_D V(D, G) = E_{x \sim P_d(x)} [\log D(x)] + E_{z \sim P_z(z)} [\log(1 - D(G(z)))] \quad (3)$$

式中: E 表示数学期望, $G(z)$ 表示生成器生成的图像, $D(G(z))$ 表示判别器的输出结果, x 为真实数据集, z 为噪声数据集, $P_d(x)$ 表示真实数据集的概率分布, $P_z(z)$ 表示噪声的概率分布。

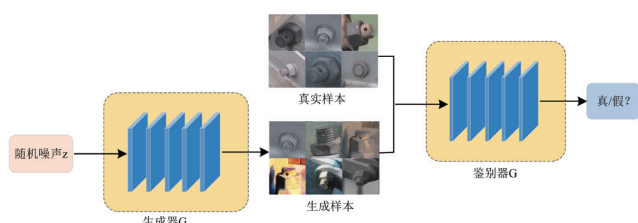


图 1 GAN 结构图

本文利用 SRGAN 生成高分辨率图像,实现拍摄图像去模糊、去噪,其网络函数为^[10]:

$$\min_G \max_D E_{I^H \sim P_G(I^H)} [\log D(I^H)] + E_{I^L \sim P_G(I^L)} [\log(1 - D(G(I^L)))] \quad (4)$$

式中: I^L 表示模糊或含噪图像, I^H 表示生成的高分辨率图像, I^H 表示原始高清图像。通过式(4)可以尽可能地生成接近原始高分辨的图像。

本文在生成器模块中引入了注意力机制模块^[11],其模型结构如图 2 所示。注意力机制模块由两部分组成,其中一部分由一个 mean-polling 层、两个卷积层和一个 ReLU 层组成,另一部分由一个 max-pooling 层、两个卷积层和一个 ReLU 层。两个部分的输出结果相加得到残差模块。

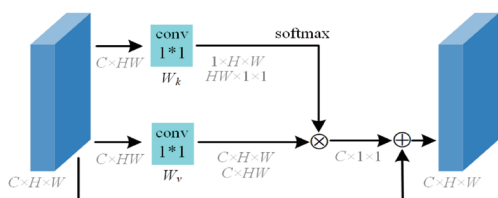


图 2 注意力机制

本文采用生成器和判别器的网络结构如图 3 所示,其中生成器的模块包括残差密集网络、卷积层和跳跃层^[12]。残差密集网络中的每个卷积层有 16 个 7*7 卷积核,步长为 1。生成器网络的后半部分由渐进式上采样、卷积层和 tanh 激活函数组成。后半部分总共有 7 个卷积层,每个卷积层的卷积核大小为 7*7。原始的判别器模型采用的是 Sigmoid 函数,但是本文判别器是用来判断生成图像分辨率的好坏,不适合用 Sigmoid 函数,因此将 Sigmoid 函数修改为 Wasserstein 距离最小化作为损失函数^[13-14],将网络从二分类问题转成回归问题。

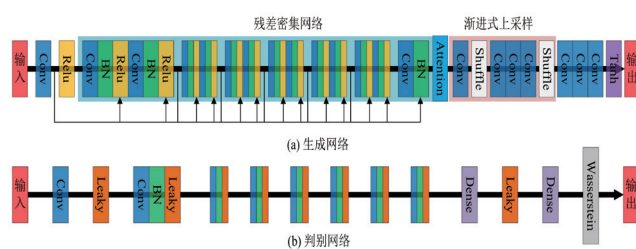


图 3 生成器和判别器的网络结构

2 GoogLeNet 卷积神经网络

与传统特征提取方法相比,卷积神经网络可以通过反向传播自适应学习特征,学习更具表现力的特征^[15],其基本结构如图 4 所示,包括卷积层、池化层、全连接层和输出层。卷积核通过卷积操作提取图像高维度特征,对图像的空间信息进行建模。

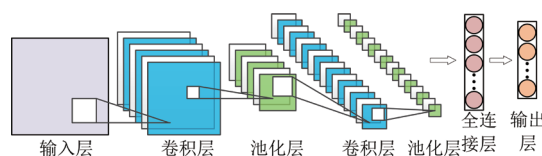


图 4 卷积神经网络的基本结构

输入的原图像用 X 表示,第 i 层的特征图用 F_i 表示。输入图像 X 的尺寸决定了输入层的大小,卷积计算过程可表示为:

$$F_i = f(F_{i-1} * w_i + b_i) \quad (5)$$

式中: f 表示激活函数, w_i 为卷积核的权重值, b_i 为偏置项, $*$ 代表卷积操作。

池化层的主要作用是在保持局部不变性的同时降低维度,针对卷积层的局部区域做下采样操作,既可以提取图像显著特征,又可以在一定程度上缓解过拟合现象。

全连接层中每个输入通过权值连接到输出,再将所有局部特征整合为整体特征,根据学习到的特征对目标判别分类,假设全连接网络为 F_i ,则可将其描述为:

$$F_i = f(w_i F_{i-1} + b_i) \quad (6)$$

全连接层的每个神经元经过激励函数处理后传递给输出层,输出层使用 Softmax 层进行目标分类,而输出值与期望值之间的残差由损失函数计算得到,常见的损失函数有均方根误差、负对数似然等,其公式为:

$$M(w, b) = \frac{1}{|Y|} \sum_{l=1}^{|Y|} (Y(l) - \hat{Y}(l))^2 \quad (7)$$

$$N(w, b) = - \sum_{l=1}^{|Y|} \log Y(l) \quad (8)$$

式中: Y 为基于输入的概率分布。

在损失函数中添加 M_2 范数可以有效抑制过拟合现象, 其定义为:

$$E(w, b) = M(w, b) + \frac{\mu}{2} w^T w \quad (9)$$

式中: μ 为控制拟合的强度。

卷积神经网络训练过程中利用到梯度下降法, 保证残差尽可能小的同时更新权值 w 和偏置项 b :

$$w_i = w_i - \eta \frac{\partial E(w, b)}{\partial w_i} \quad (10)$$

$$b_i = b_i - \eta \frac{\partial E(w, b)}{\partial b_i} \quad (11)$$

式中: η 为学习速率, 用于控制残差反向传播强度。

GoogLeNet 网络是图像分类领域的代表算法之一, 具有分类精度高和运行内存少等优势。GoogLeNet 网络模型是基于图 5 的 Inception 模块堆叠而成。

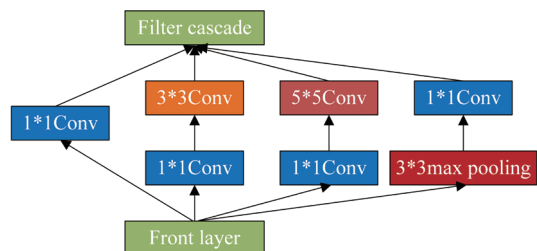


图 5 Inception 模块

Inception 模块在增加了网络深度的同时减轻了网络结构负担。为解决传统 CNN 网络参数过于复杂且泛化能力差等问题, GoogLeNet 网络最后一层采用全局均值池化操作。

GoogLeNet 网络结构如图 6 所示。

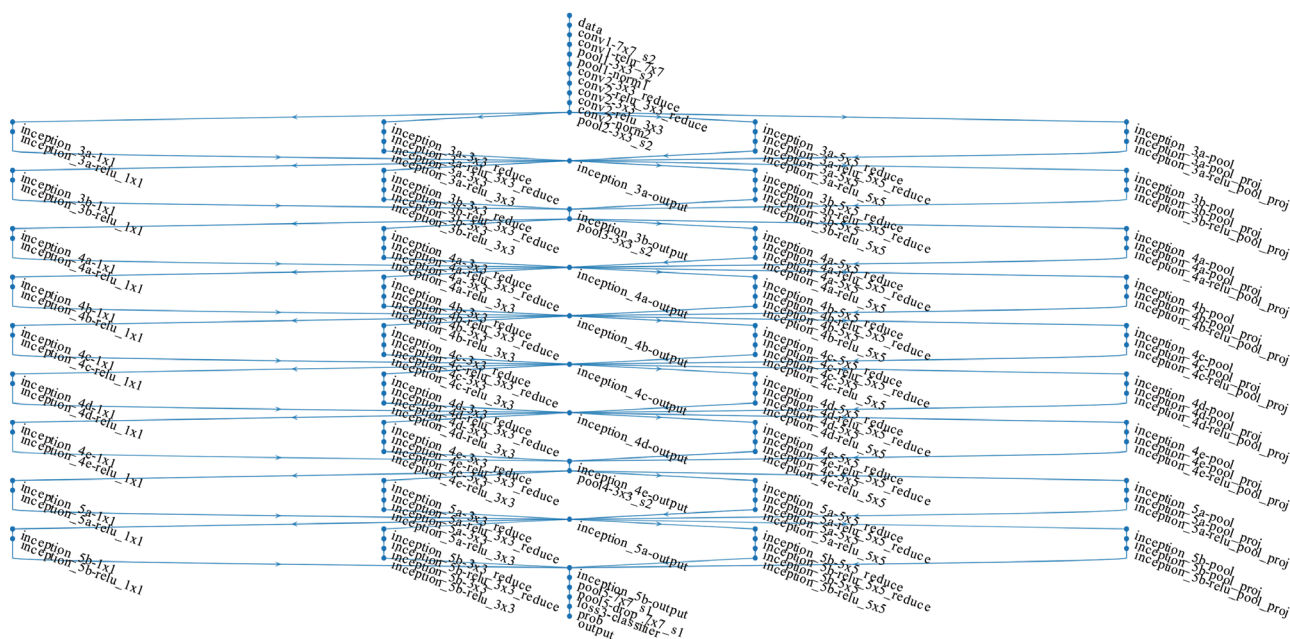


图 6 Inception 模块

为了提高安全帽佩戴安全性检测准确率并实现安全帽的自动标注, 在 Faster RCNN 框架下, 将 RPN 区域建议网络^[16]与 GoogLeNet 网络相结合。利用 GoogLeNet 网络提取建议区域的特征对安全帽安全性佩戴数据集进行检测分类。将 GoogLeNet 网络的“loss3-classifier”替换为“fc”层, 将“prob 层”替换为“soft-max”层, 将“output”层替换为“classoutput 层”, 并将最后一层全连接层的神经元个数设为 2, 实现安全帽的检测, 其具体流程见图 7。首先, 利用 SRGAN 网络对采集到的图像进行复原处理; 其次, 利用 GoogLeNet 网络自主挖掘图像特征; 最后, 在 Faster RCNN 框架下实现安全帽的定位与标注。

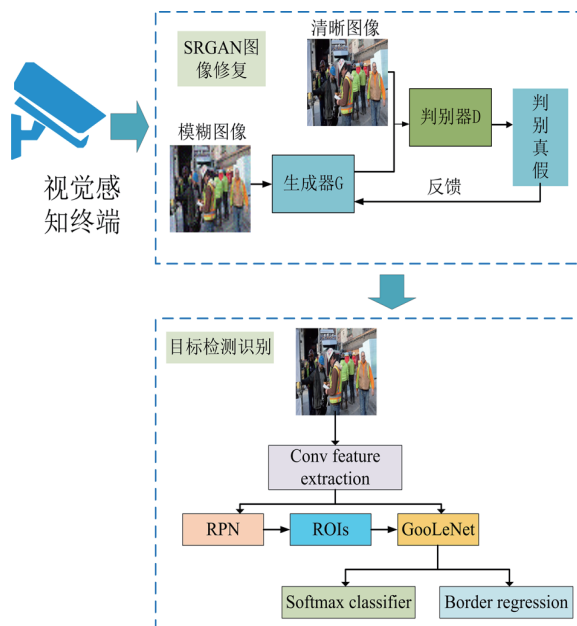


图 7 安全帽检测的整体流程图

3 实验结果与分析

3.1 实验配置

实验的系统配置影响着深度学习的效率，本文选用的硬件配置为：Windows10，CPU Inter Core i7-7800X，GPU NVIDIA GeForce GTX 1080Ti，内存 16 GB。本文采用的数据集包括 8000 张图片，其中训练样本 6000 张，测试样本 2000 张。数据集涵盖了不同角度、不同天气、不同姿势、不同拍摄距离、图像模糊和图像含噪等条件，确保样本库的丰富性，图 8 给出了样本库示例。



图 8 数据集样本实例

3.2 SRGAN 图像复原

图 9 给出了三种方法去模糊的效果图，包括 richardson-lucy (R-Lucy) 算法、盲卷积 (blind convolution, B-conv)、改进的 SRGAN 算法。从图 9 可以看出，改进后的 SRGAN 算法得到的图像主观视觉更好，更加平滑，而且在恢复图像高频特征方面，改进的 SRGAN 算法也是优于其他两种算法的。

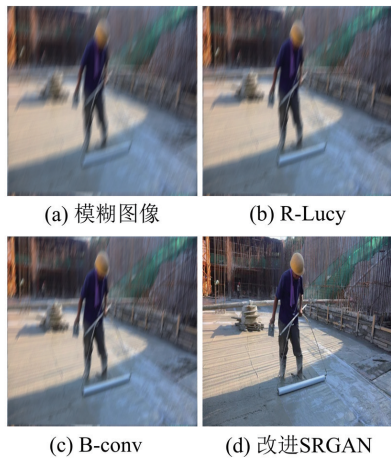


图 9 三种方法的去模糊效果图

表 1 给出了三种算法去模糊效果的客观评价指标。峰值信噪比 (peak signal-to-noise ration, PSNR) 的数值越大，说明图像的失真越少；结构相似度 (structural similarity,

SSIM) 越接近 1，说明与原图像越接近。从表 1 可以看出，改进的 SRGAN 在恢复图像的客观评价指标上优于其他两种算法。

表 1 不同去模糊方法的客观实验结果

方法	PSNR/dB	SSIM
R-Lucy	12.315 2	0.683 1
B-conv	15.783 8	0.702 3
改进 SRGAN	38.193 5	0.986 2

实际拍摄的图像也可能存在噪声，图 10 给出了三种方法的去噪效果图。可以看出，改进 SRGN 去噪效果优于小波阈值去噪和变分模态分解去噪，SRGAN 算法得到图像的噪声更小，得到的图像更加清晰。

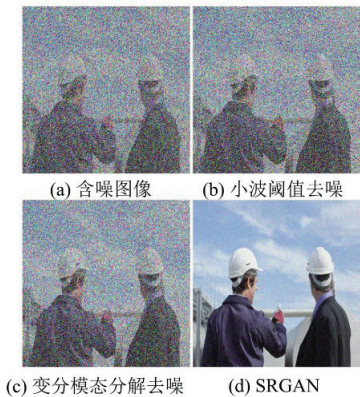


图 10 三种方法的去噪效果图

3.3 基于 GoogLeNet 的安全帽定位结果

通过 SRGAN 算法对图像复原以后，在 Faster RCNN 框架下，应用 GoogLeNet 模型实现安全帽定位，所得结果如图 11 所示，可以看出，在不同拍摄角度、不同拍摄情况下、不同距离下，本文方法均能够准确地定位安全帽的位置。

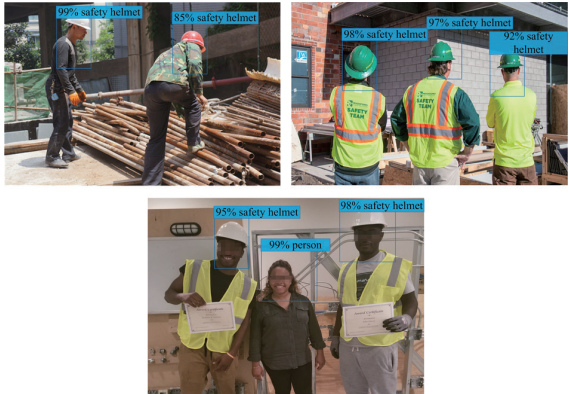


图 11 安全帽定位结果

图 12 给出了不同算法的检测性能对比，可以看出本文方法的检测性能是优于传统算法的，本文方法的安全帽佩戴

安全性智能检测的准确率可以达到 98.10%，召回率可以达到 98.30%，具有较好的智能检测效果。

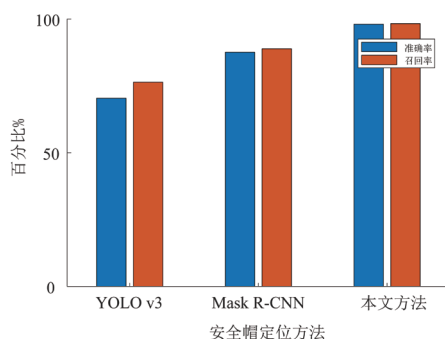


图 12 不同算法检测性能对比

4 结论

利用摄像头检测工作人员是否正确佩戴安全帽时，拍摄的图像存在分辨率低、含噪成分严重和图像模糊等问题，严重影响安全帽佩戴安全性智能检测的效果。本文提出一种结合改进的超分辨率生成式对抗网络（SRGAN）和 GoogLeNet 模型的安全帽佩戴安全性检测方法。SRGAN 可以对模糊、含噪和低分辨率图像进行复原，得到高分辨率图像，主观视觉和客观评价指标均优于传统的算法。在 Faster RCNN 框架下，利用 GoogLeNet 网络可以实现安全帽的定位，智能检测的准确率可以达到 98.10%，召回率可以达到 98.30%。本文方法可为安全帽检测提供一个新的思路，利用该方法可以预警没有戴安全帽的工作人员，提高工作人员的安全意识。

参考文献：

- [1] 薛明华, 艾春美, 律慧瑾, 等. 电厂安全帽佩戴安全性监控的智能图像处理方法 [J/OL]. 中国电机工程学报 :1-9 [2024-01-12]. <https://doi.org/10.13334/j.0258-8013.pcsee.210629>.
- [2] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述 [J]. 计算机学报, 2017,40(6):1229-1251.
- [3] 赵红成, 田秀霞, 杨泽森, 等. YOLO-S: 一种新型轻量的安全帽佩戴检测模型 [J]. 华东师范大学学报(自然科学版), 2021(5): 134-145.
- [4] 张乾, 岳诗琴, 范玉. 增强随机蕨的安全帽佩戴检测 [J]. 计算机仿真, 2021,38(7):429-432.
- [5] 冯国臣, 陈艳艳, 陈宁, 等. 基于机器视觉的安全帽自动识别技术研究 [J]. 机械设计与制造工程, 2015,44(10): 39-42.
- [6] 王雨生, 顾玉宛, 封晓晨, 等. 基于姿态估计的安全帽

佩戴检测方法研究 [J]. 计算机应用研究, 2021,38(3):937-940+945.

- [7] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks [J]. Advances in neural information processing systems, 2014,3:2672-2680.
- [8] 刘云鹏, 许自强, 和家慧, 等. 基于条件式 Wasserstein 生成对抗网络的电力变压器故障样本增强技术 [J]. 电网技术, 2020, 44(4):1505-1513.
- [9] 肖白, 黄钰茹, 姜卓, 等. 数据匮乏场景下采用生成对抗网络的空间负荷预测方法 [J]. 中国电机工程学报, 2020, 40(24): 7990-8001+8236.
- [10] 胡蕾, 王足根, 陈田, 等. 一种改进的 SRGAN 红外图像超分辨率重建算法 [J]. 系统仿真学报, 2021, 33(9): 2109-2118.
- [11] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision, 2018: 3-19.
- [12] 乔昕, 魏延. 一种改进的 SRGAN 图像超分辨重建算法 [J]. 计算机时代, 2021(1): 72-75+79.
- [13] 高晓松, 李更丰, 肖遥, 等. 基于分布鲁棒优化的电-气-热综合能源系统日前经济调度 [J]. 电网技术, 2020, 44(6): 2245-2254.
- [14] ZHU R, WEI H, BAI X, et al. Wasserstein metric based distributionally robust approximate framework for unit commitment [J]. IEEE transactions on power systems, 2019, 34(4): 2991-3001.
- [15] 熊丽婷. 基于 GoogLeNet-GMP 网络的自适应图像水印方法 [J]. 电子测量技术, 2021, 44(14): 128-134.
- [16] 徐一鸣, 张娟, 刘成成, 等. 迁移学习模式下基于 GoogLeNet 网络的风电机组视觉检测 [J]. 计算机科学, 2019, 46(5): 260-265.

【作者简介】

张旺（1989—），女，吉林长春人，硕士，中级工程师，研究方向：机器视觉、故障诊断。

（收稿日期：2024-07-03）