

# 基于改进 YOLOv4 的公路路面病害检测算法

罗晖, 贾晨\*, 李健

华东交通大学信息工程学院, 江西 南昌 330013

**摘要** 针对公路路面病害存在的类别多、尺度变化大及样本数据集小导致的病害检测困难等问题, 提出了基于改进 YOLOv4 的公路路面多尺度病害检测算法。首先, 在 CSPDarknet-53 骨干网络中采用深度可分离卷积替代普通卷积, 降低了网络参数计算量; 然后, 基于 Focal loss 改进 YOLOv4 的损失函数, 解决了网络训练过程中正、负样本不平衡而导致的检测精度较低的问题; 最后, 借助迁移学习思想对 YOLOv4 网络进行预训练, 并运用翻转、裁剪、亮度变换、噪声扰动等方法进行数据集扩充, 解决了公路路面病害样本不足导致的网络训练过拟合问题。实验结果表明, 所提基于 YOLOv4+DC+FL 算法对公路路面病害检测的平均精度均值可达到 93.64%, 相较于原始的 YOLOv4 检测网络提高了 3.25%, 检测每张图片平均时间为 35.8 ms, 缩减了 7.9 ms。

**关键词** 图像处理; 公路路面病害; YOLOv4; 深度可分离卷积; Focal loss; 迁移学习; 数据增广

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.1410025

## Road Surface Disease Detection Algorithm Based on Improved YOLOv4

Luo Hui, Jia Chen\*, Li Jian

School of Information Engineering, East China JiaoTong University, Nanchang, Jiangxi 330013, China

**Abstract** In order to solve the problems of multiple types of road surface diseases, large scale changes in the scale, and small sample data sets in road surface disease detection, a road surface multi-scale disease detecting algorithm based on improved YOLOv4 is proposed. First, the depth separable convolution method is used to replace the ordinary convolution method in the CSPDarknet-53 backbone network, which reduces the amount of network parameter calculations. Then, the loss function of YOLOv4 is improved based on the focal loss, which solves the problem of low detection accuracy caused by the imbalance of positive and negative samples in the process of network training. Finally, the YOLOv4 network is pre-trained with the help of transfer learning ideas, and the data set is expanded using methods such as flipping, cropping, brightness conversion, noise disturbance and other methods, so as to solve the problem of over-fitting of network training caused by insufficient samples of road surface disease. Experimental results show that, compared with original YOLOv4 detection network, the mean average precision of road surface disease detection based on YOLOv4 + DC + FL algorithm can reach 93.64%, which increases by 3.25%. The detection time is 35.8 ms per picture, which is reduced by 7.9 ms.

**Key words** image processing; road surface disease; YOLOv4; depth separable convolution; Focal loss; transfer learning; data augmenting

**OCIS codes** 100.2000; 100.3008; 100.4996

## 1 引言

公路运输是陆上运输的基本方式之一, 其灵活

机动的特点使其不仅成为一个独立的运输体系, 也成为港口、车站和机场物资集散的重要手段, 是国家经济发展的重要基础。由于公路在使用过程中, 路

收稿日期: 2020-08-17; 修回日期: 2020-09-09; 录用日期: 2020-09-30

基金项目: 江西省重点研发计划项目(20202BBEL53001)

通信作者: \*jc\_ejtu@163.com

面会出现不同形状、不同尺寸的病害。公路路面病害按路面破损类型分为裂缝类、变形类、表面病害类三类,其中裂缝病害可分为横向裂缝、纵向裂缝和网状裂缝;表面病害类主要是指坑槽等;变形类病害则包括车辙、隆起等<sup>[1]</sup>。常见且难检测的病害主要有裂缝类及坑槽病害。若这些病害产生初期没有得到及时处理,不仅会影响公路的美观、降低行车人员的舒适感、减少公路的承重量和使用寿命,而且会诱发更严重的病害,轻则造成更大的养护资金投入,重则引发交通事故、造成严重的经济损失和人员伤亡。因此,采用有效的目标检测方法对公路路面病害进行准确、快速的检测具有重大的意义<sup>[2-3]</sup>。

近年来,基于深度学习的公路路面病害检测方法可以很好地实现特征学习和特征分类,具有检测精度高、检测速度快等特点。Zhang 等<sup>[4]</sup>基于图像局部块信息,利用卷积神经网络判断单块图像是否属于裂缝;Cha 等<sup>[5]</sup>应用滑动窗口将图像分割成块,然后应用卷积神经网络对裂缝病害特征进行提取并对裂缝进行分类;李刚等<sup>[6-7]</sup>针对传统裂缝图像分割方法不能准确提取目标表面裂缝的难题,提出了一种改进的轻量级全局卷积网络的路面裂缝图像分割模型和交错低秩组卷积混合深度网络,提高了裂缝的检测精度;张宁<sup>[8]</sup>提出一种基于 Faster R-CNN 的公路路面病害检测算法,使用 Soft-NMS 算法替代 NMS,减少了检测框的冗余度与病害的漏检率,提升了裂缝及坑槽两类病害的检测精度;通过上文文献分析,基于深度学习的方法可以解决公路路面病害的检测问题。

而基于深度学习的目标检测算法主要分为基于区域建议和基于无区域建议的两类目标检测算法。基于区域建议的目标检测算法主要有 R-CNN、Fast R-CNN、Faster R-CNN、特征金字塔(FPN)等<sup>[9-12]</sup>,此类算法分为两步,先生成一系列候选框,再用卷积神经网络进行分类和坐标回归,具有较高的检测精度,但检测速度较慢。而基于无区域建议的目标检测算法主要有:2016 年 Redmon 等<sup>[13]</sup>提出 YOLOv1 算法,将生成候选框和分类回归合并成一个网络中,大大降低了网络计算的复杂度,但目标定位精度较低;SSD 算法<sup>[14]</sup>则采用了锚点思想,以提高检测精度和速度;Redmon 等<sup>[15-16]</sup>又先后提出 YOLOv2 和 YOLOv3 算法,从训练数据集、网络结构和锚点处理等多方面对 YOLOv1 进行改进,在保证检测速度的同时提高了目标检测精度;YOLOv4<sup>[17]</sup>算法在

YOLOv3 的基础上增加了空间金字塔池化(SPP)<sup>[18]</sup>、跨阶段部分连接(CSP)<sup>[19]</sup>等模块,平均精度均值(mAP)和速度分别提升了 10% 和 12%;此类算法虽相较于基于区域提名的目标检测算法而言,对多类、多尺度目标检测精确度略低,但实时性较好。

基于以上分析,为了保证公路路面病害检测实时性,同时进一步提高检测精度,本文提出一种基于改进 YOLOv4 的公路路面病害检测算法。

## 2 基于 YOLOv4 的目标检测算法

在 YOLO 系列算法中,YOLOv4 算法模型由 CSPDarknet 主干网络、空间金字塔池化、PANet 多尺度特征提取模块、YOLO Head 输出层四部分组成,整体结构如图 1 所示。其中,CSPDarknet53<sup>[19]</sup>结合残差网络 ResNet 中的跳跃连接作为特征提取网络(图 2),增强了卷积神经网络的学习能力,能够使网络轻量化的同时保证检测精确度;并在批归一化(BN)层后使用 Mish<sup>[20]</sup>提出的 Mish 激活函数,可以更好地提取目标的特征信息,提高目标的检测精度;在 FPN<sup>[12]</sup>模块中使用了 SPP<sup>[18]</sup>和路径汇聚网络(PAN)<sup>[21]</sup>结构框架;其中,SPP 用于对 CSPDarknet53 的最后一个特征层的输出进行四个不同尺度的最大池化处理,卷积核大小分别为  $13 \times 13, 9 \times 9, 5 \times 5, 1 \times 1$ ,不仅能提高图像的尺度不变性,还可以减少过拟合现象;而 PAN 在完成特征金字塔从下到上的特征提取后,还完成了从上到下的特征提取。

在基于 YOLOv4 的目标检测过程中,首先将图像划分为  $S \times S$  个网格,每个网格负责预测中心落入该网格的目标,并计算出 3 个预测框,每个预测框对应  $5 + C$  个值;其中, $C$  表示数据集中的类别总数, $5$  代表预测边界框的中心点坐标  $(x, y)$ 、框的宽高尺寸  $(w, h)$  和置信度等参数信息。然后,求解网格预测的类别置信度,其与目标落入网格的概率  $P(n_{\text{object}})$ 、网格预测第  $i$  类目标的准确率  $P(n_{\text{class}i} | n_{\text{object}})$ 、完全交并比<sup>[22]</sup>(CIoU)有关,求解表达式为  $P(n_{\text{class}i} | n_{\text{object}}) \times P(n_{\text{object}}) \times n_{\text{CIoU}_{\text{pred}}^{\text{truth}}}$ ;若目标中心落入该网格,则  $P(n_{\text{object}}) = 1$ ,否则为 0; $n_{\text{CIoU}_{\text{pred}}^{\text{truth}}}$  是在预测边框与真实框的距离交并比(DIoU)<sup>[22]</sup>上引入两个惩罚因子  $\alpha, v$ 。最后,使用 DIoU NMS<sup>[23]</sup>筛选出置信度值较高的预测框作为目标检测框,输出特征图大小分别为  $13 \times 13, 26 \times 26, 52 \times 52$ ,从而实现目标的定位与分类。

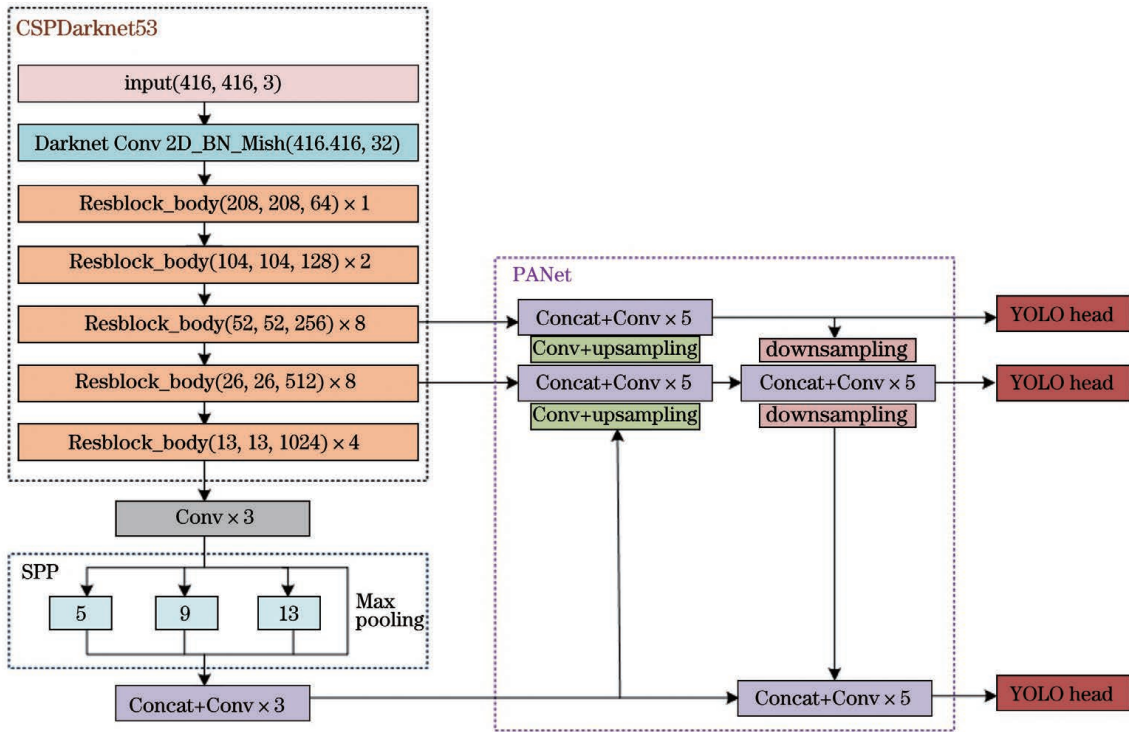


图 1 YOLOv4 网络结构图<sup>[17]</sup>

Fig. 1 Network structure of YOLOv4

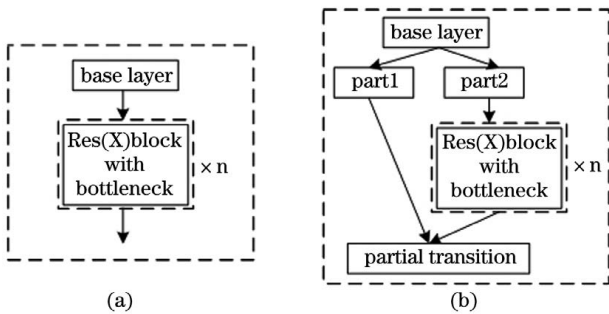


图 2 CSPResNet 示意图<sup>[19]</sup>。(a)残差网络主体；

(b)CSPResNet(X)t

Fig. 2 Schematic diagram of CSPResNet. (a) Resblock\_ body; (b) CSPResNet(X)t

### 3 基于改进 YOLOv4 的公路路面病害检测算法

针对公路路面病害存在多类、多尺度特点导致检测精度低、速度慢等问题,本文对 YOLOv4 的 CSPDarknet-53 网络框架及损失函数提出两点改进:一是采用深度可分离卷积替代普通卷积的方法,减少网络参数量,降低计算复杂度,提高检测速度;二是引入 Focal loss,解决网络训练过程中通过正、负样本不平衡而导致的检测精度较低的问题,在保证检测速度的同时提高公路路面病害的检测精度。

#### 3.1 卷积方式的改进

在标准卷积的运算过程中(图 3),若输入特征图尺寸大小为  $H \times W \times M$ ,采用  $N$  个尺寸大小为  $K \times K$  的标准卷积核,输出特征图的大小为  $H' \times W' \times N$ ,则标准卷积的参数计算量  $C_1$  的公式为

$$C_1 = H \times W \times M \times N \times K \times K, \quad (1)$$

式中: $M$  为输入特征图的通道数; $N$  为输出通道数。

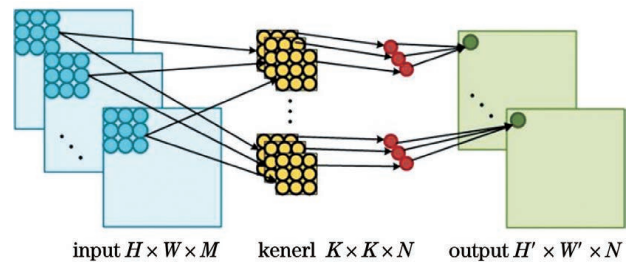


图 3 标准卷积过程

Fig. 3 Standard convolution process

本文在 CSPDarknet 的特征提取网络中加入 MobileNet<sup>[24]</sup> 提出的深度可分离卷积方法,该方法将标准的卷积运算分解为深度卷积和逐点卷积两个过程,具体过程如图 4 所示。首先,在深度卷积过程中,利用  $M$  个尺寸大小为  $K \times K \times 1$  的卷积核对  $H \times W \times M$  的输入图像进行单通道卷积,得到  $H' \times W' \times M$  维度的输出。该过程中深度卷积的参

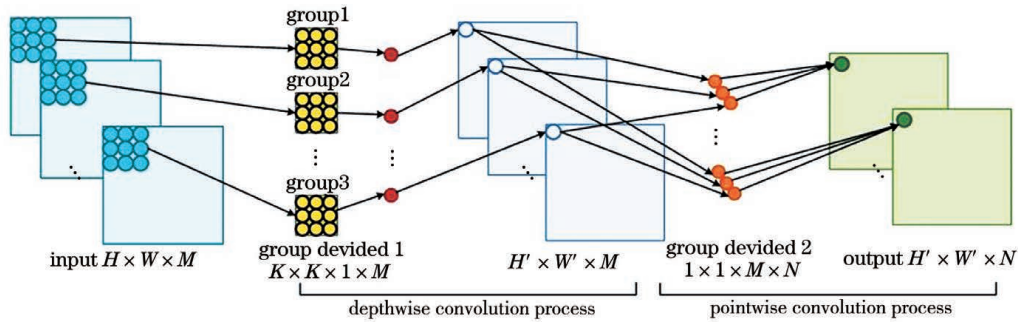


图 4 深度可分离卷积过程

Fig. 4 Depth separable convolution process

数计算量  $C_d$  的计算公式为

$$C_d = H \times W \times M \times K \times K. \quad (2)$$

然后,在逐点卷积过程中,利用  $N$  个尺寸大小为  $1 \times 1 \times M$  的卷积核,对上一步得到的  $H' \times W' \times M$  维度的输出再次进行卷积运算,输出特征图为  $H' \times W' \times N$ ,该过程中深度卷积的参数计算量  $C_p$  的计算公式为

$$C_p = H' \times W' \times M \times N. \quad (3)$$

综上所述,深度可分离卷积的参数计算量  $C_2$  的计算公式为

$$C_2 = C_d + C_p = H \times W \times M \times K \times K + H' \times W' \times M \times N. \quad (4)$$

深度可分离卷积与标准卷积的参数计算量之比表示为

$$\frac{C_2}{C_1} = \frac{1}{N} + \frac{1}{K^2}. \quad (5)$$

若卷积核大小为  $3 \times 3$ ,步长为 1,则采用深度可分离卷积方式的参数计算量约为采用标准卷积方式的  $1/9$ ,相较于标准卷积运算,减少了参数计算量,可提高目标检测速度.改进后的卷积运算模块

如图 5(b)所示。

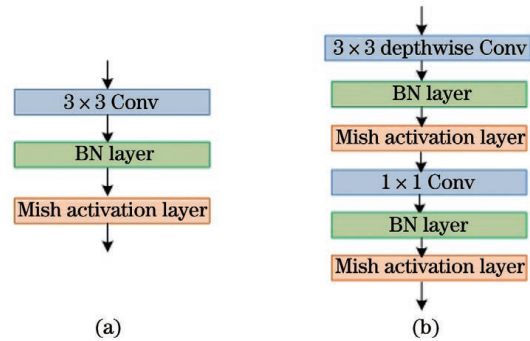


图 5 基于 YOLOv4 网络结构的标准卷积与深度可分离卷积模块。(a)标准卷积模型;(b)深度可分离卷积模型

Fig. 5 Standard convolution and depth separable convolution model based on YOLOv4 network structure. (a) Standard convolution model; (b) depth separate convolution model

### 3.2 损失函数的改进

本文针对正负样本数不均衡问题,在 YOLOv4 的损失函数中引入 Focal loss<sup>[25]</sup>。Focal loss 是在标准交叉熵损失基础上改进的一种损失函数,表示为

$$F = \begin{cases} (1 - y')^\gamma [-y \log y' - (1 - y) \log(1 - y')] & , y = 1 \\ y'^\gamma [-y \log y' - (1 - y) \log(1 - y')] & , y = 0 \end{cases}, \quad (6)$$

式中: $y$  为真实标签; $y'$  为标签的预测值, $y' \in [0, 1]$ ; $\gamma$  为调节参数,用于调节简单样本权重,以提高网络训练速度,使得模型可以更好地实现样本的特征学习,并解决标准交叉熵损失中存在的正样本

输出概率越大损失越小、负样本输出概率越小损失越小的问题,从而提高目标检测精度。

因此,引入 Focal loss 后的 YOLOv4 损失函数  $L'$ ,表示为

$$\begin{aligned} L' = & \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} [(x_i^j - \hat{x}_i^j)^2 + (y_i^j - \hat{y}_i^j)^2] + \lambda_{\text{coord}} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} (2 - w_i^j \times h_i^j) [(w_i^j - \hat{w}_i^j)^2 + (h_i^j - \hat{h}_i^j)^2] + \\ & \lambda_{\text{obj}} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{obj}} [-C_i^j \log(\hat{C}_i^j) - (1 - C_i^j) \log(1 - \hat{C}_i^j)] + \lambda_{\text{noobj}} \sum_{i=0}^{s^2} \sum_{j=0}^B \mathbf{1}_{ij}^{\text{noobj}} [-C_i^j (1 - \hat{C}_i^j)^\gamma \log(\hat{C}_i^j) - \\ & (1 - C_i^j) (\hat{C}_i^j)^\gamma \log(1 - \hat{C}_i^j)] + \lambda_{\text{class}} \sum_{i=0}^{s^2} \sum_{c \in C_{\text{class}}} \mathbf{1}_{ij}^{\text{noobj}} [-P_i^j \log(\hat{P}_i^j) + (1 - \hat{P}_i^j) \log(1 - \hat{P}_i^j)], \quad (7) \end{aligned}$$

式中:输入图像单元格的个数为  $s^2$ ;  $B$  为每个单元格所预测的边界框个数;  $\mathbf{1}_{ij}^{\text{obj}}$  的取值为 1 或 0, 即第  $i$  个单元格的第  $j$  个边界框中是否存在检测目标, 存在则为 1, 否则为 0;  $(x_i^j, y_i^j, w_i^j, h_i^j)$  分别是预测边界的中心坐标、高度、宽度;  $(\hat{x}_i^j, \hat{y}_i^j, \hat{w}_i^j, \hat{h}_i^j)$  指实际边界框的中心坐标、高度、宽度;  $\lambda_{\text{obj}}$  和  $\lambda_{\text{noobj}}$  是指交并比误差的权重;  $C_i^j$  为预测置信度,  $\hat{C}_i^j$  为实际置信度;  $P_i^j$  是单元格中目标所属类别的实际概率,  $\hat{P}_i^j$  为预测概率。

$$\begin{cases} C = P(n_{\text{object}}) \times n_{\text{CloU pred}}^{\text{truth}} \\ n_{\text{CloU}} = \frac{\rho^2(b, b^{\text{gt}})}{C^2} + \alpha v \end{cases}, \quad (8)$$

式中:  $\alpha$  为权重参数;  $v$  是用来衡量长宽比一致性的参数。具体计算公式为

$$\begin{cases} \alpha = \frac{v}{(1 - n_{\text{IoU}}) + v} \\ v = \frac{4}{\pi^2} \left( \arctan \frac{w^{\text{gt}}}{h^{\text{gt}}} - \arctan \frac{w}{h} \right)^2 \end{cases} \quad (9)$$

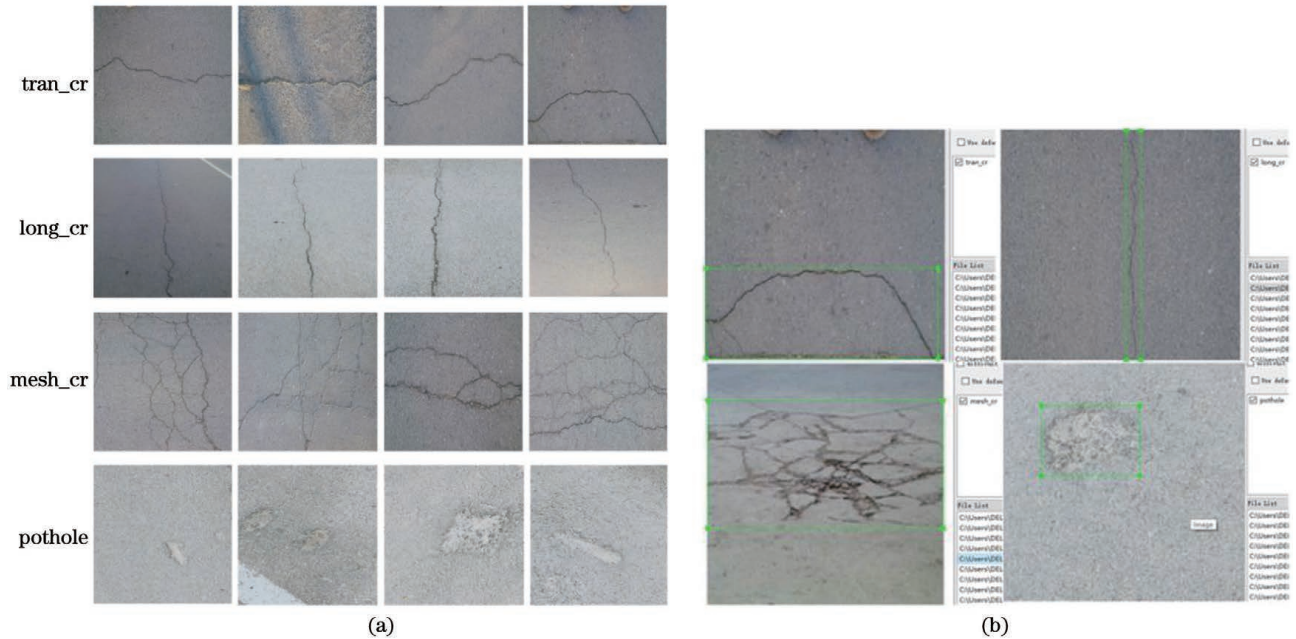


图 6 公路路面病害数据集及其标签示例。(a)公路路面病害数据集示例;(b)公路路面病害数据集标签示例  
Fig. 6 Examples of data set and label for road surface disease. (a) Examples of data set for road surface disease;  
(b) examples of label for road surface disease

针对公路路面病害数据集中多类公路路面病害样本数量较少而导致网络模型出现过拟合的问题, 对已标记后的自建公路路面病害样本进行翻转、裁剪、亮度变换、噪声扰动等数据集增广操作, 以扩充多类公路路面病害样本数量, 从而减少过拟合现象,

## 4 实 验

在实验过程中, 首先建立公路路面病害数据集; 其次, 借助迁移学习<sup>[26]</sup>的思想, 采用 VOC2007 数据集对公路路面病害检测网络进行预训练, 得到网络预训练权重参数; 然后, 选用所建立的公路路面病害数据集对预训练好的网络结构进行参数微调; 最后, 完成公路路面病害的检测。

### 4.1 数据集的建立

公路路面病害原始数据集共有 1309 张样本图像, 其中包含公共公路路面数据集 Crack500<sup>[23]</sup> (共 500 张样本图像) 和自建公路路面病害数据集 (共 809 张样本图像)。其中, 自建公路路面病害数据集包括横向裂缝 (记为 tran\_cr, 共 305 张样本图像)、纵向裂缝 (记为 long\_cr, 共 285 张样本图像)、网状裂缝 (记为 mesh\_cr, 共 115 张样本图像)、坑槽 (记为 pothole, 共 104 张样本图像), 其样本数据示例如图 6(a) 所示。在公路路面病害数据集标注过程中, 采用开源标注工具 Labelimg, 生成对应的 XML 格式的标记文件, 其标签示例如图 6(b) 所示。

具体如表 1 所示。表中“-”代表未采用此方法, “√”代表采用此方法。

其增广效果如图 7 所示。

数据增广后建立本文所用的公路路面病害数据集, 包含 5774 张样本图像。在实验过程当中, 采用

表 1 不同公路路面病害类型的数据增广方式  
Table 1 Data augmenting methods of different road surface disease types

Name of sample	Method			
	Flip	Crop	Bright	Add noise
tran_cr	-	✓	✓	✓
long_cr	-	✓	✓	✓
mesh_cr	✓	✓	✓	✓
pothole	✓	✓	✓	✓

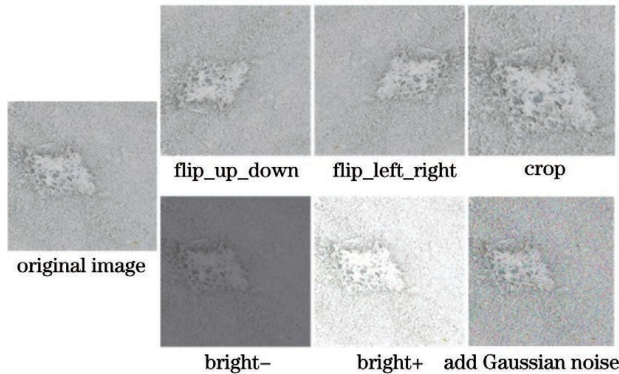


图 7 公路路面病害数据增广示例

Fig. 7 Examples of road surface disease data augmenting  
80% 的样本图像作为训练集、10% 的样本图像作为验证集、10% 的样本图像作为测试集, 具体样本数量分布见表 2。

表 2 公路路面病害数据样本分布

Table 2 Sample distribution of road surface disease data

Name of sample	Traning samples	Validation samples	Test samples	Total
tran_cr	1776	222	222	2220
long_cr	1552	194	194	1940
mesh_cr	792	99	99	990
pothole	500	62	62	624
Total	4620	577	577	<b>5774</b>

表 3 不同网络模型下对公路路面病害检测的结果对比

Table 3 Comparison of detection results of road surface disease based on different network models

Name of model	Evaluation index / %					
	mAP	AP (tran_cr)	AP (long_cr)	AP (mesh_cr)	AP (pothole)	Time /ms
Faster R-CNN	93.55	96.29	96.41	91.54	89.96	105.0
SSD	82.44	86.02	87.83	88.07	67.81	44.3
YOLOv3	84.18	91.44	87.35	88.26	69.69	39.5
YOLOv4	90.39	93.24	93.14	88.49	86.67	43.7
YOLOv4+DC	91.01	93.26	93.65	89.18	87.95	35.6
YOLOv4+FL	92.22	95.96	94.67	90.05	88.18	43.8
YOLOv4+DC+FL	<b>93.64</b>	<b>96.37</b>	<b>96.38</b>	<b>91.77</b>	<b>90.04</b>	<b>35.8</b>

## 4.2 实验结果与分析

### 4.2.1 实验平台及训练过程

本实验采用的硬件环境为: NVIDIA GeForce RTX 2080GPU 显卡、Intel i7-7800X CPU 处理器、内存 32 GB。软件环境为: Windows10 操作系统、Python3.6.5、CUDA10.0、cuDNN7.4.1.5、TensorFlow-gpu1.13.2。

在训练过程中, 采用随机梯度下降法(SGD) 对网络进行训练, 输入图像尺寸大小为  $416 \times 416$ , 初始学习率为 0.001, 学习衰减速率设置为 0.0001, 动量为 0.9, batch size 设置为 8, 迭代次数限制为  $5 \times 10^4$ 。网络损失值的变化曲线如图 8 所示, 训练经过 1000 次迭代后损失值开始大幅度下降, 损失曲线大约在 6000 次迭代后逐渐趋于平稳, 损失值低于 0.05。由此说明, 所提基于 YOLOv4+DC+FL 的网络各项超参数设置合理, 学习效果理想。

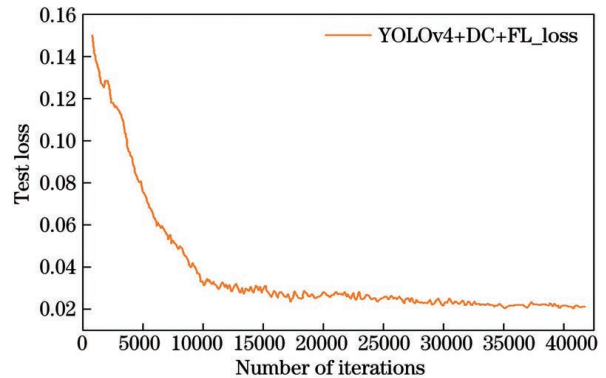


图 8 基于 YOLOv4+DC+FL 网络的训练损失值变化

Fig. 8 Changes of training losses based on improved YOLOv4

### 4.2.2 评价指标

本文在公路路面病害检测过程中, 采用平均精度(AP)、mAP、每张图片的平均检测时间作为算法的主要评价指标, 对改进的 YOLOv4 网络与其他网络对公路路面病害测试集的检测效果进行对比, 如表 3 所示。

表 3 中采用深度可分离卷积结构的 YOLOv4+DC 网络检测效果相较于未改进的 YOLOv4 网络而言,公路路面病害检测的 mAP 值增加了 0.62%,检测时间缩减了 8.1 ms,验证了深度可分离卷积替代普通卷积的方法具有提高公路路面病害的检测精度和检测速度的有效性;而引入 Focal loss 损失的 YOLOv4+FL 网络,相较于未改进的 YOLOv4 网络,mAP 值提高了 1.83%,说明引入 Focal loss 损失的 YOLOv4 网络对公路路面病害检测的 mAP 值有较为明显的提高,有效地解决了网络训练过程中通过正、负样本不平衡而导致的检测精度较低的问题。表 3 中基于 YOLOv4+DC+FL 的网络将深度可分离卷积和 Focal Loss 损失同时加入 YOLOv4 网络,得出公路路面病害的检测 mAP 值、四类病害的 AP 值以及每张图片的平均检测时间。其中公路路面病害的 mAP 值相较于 Faster R-CNN、SSD、YOLOv3 和未改进的 YOLOv4 网络而言,分别提高了 0.09%、11.2%、9.46%、3.25%;四类病害(tran\_cr、long\_cr、mesh\_cr、pothole)的 AP 值相较于未改进的 YOLOv4 网络,分别提高了 3.31%、3.24%、2.59%、3.37%;基于 YOLOv4+DC+FL 网络的每张图片的平均检测时间大约为 35.8 ms,较 Faster R-CNN、SSD 和未改进的 YOLOv4 网络分别缩减了 69.2、8.5、3.7、7.9 ms。综合表 3 中各组实验对比数据,可以验证本文所提

改进方法能有效提高公路路面病害检测精度、检测速度。

基于不同网络的公路路面病害检测精确率-召回率(P-R)曲线图如图 9 所示。

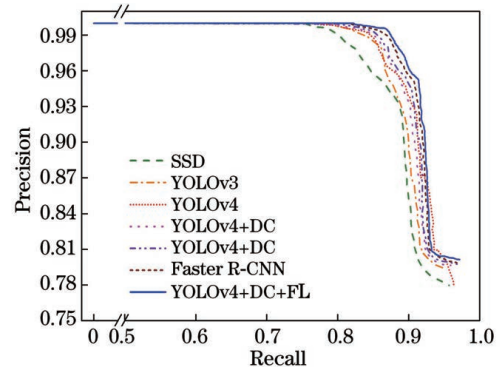


图 9 基于不同网络模型的公路病害检测的 P-R 曲线对比图

Fig. 9 Comparison of P-R curves for road surface disease based on different network models

由图 9 可知,在同一召回率值下,基于 YOLOv4+DC+FL 的公路路面病害检测方法的精确度高于其他组网络模型,说明本文所提改进方法的漏检率和误检率均低于其他组网络模型,具有更好的检测效果。

在实验过程中,基于不同的网络模型对公路路面四类病害(tran\_cr、long\_cr、mesh\_cr、pothole)的检测效果如图 10 所示,其中图 10(a)为输入图像,

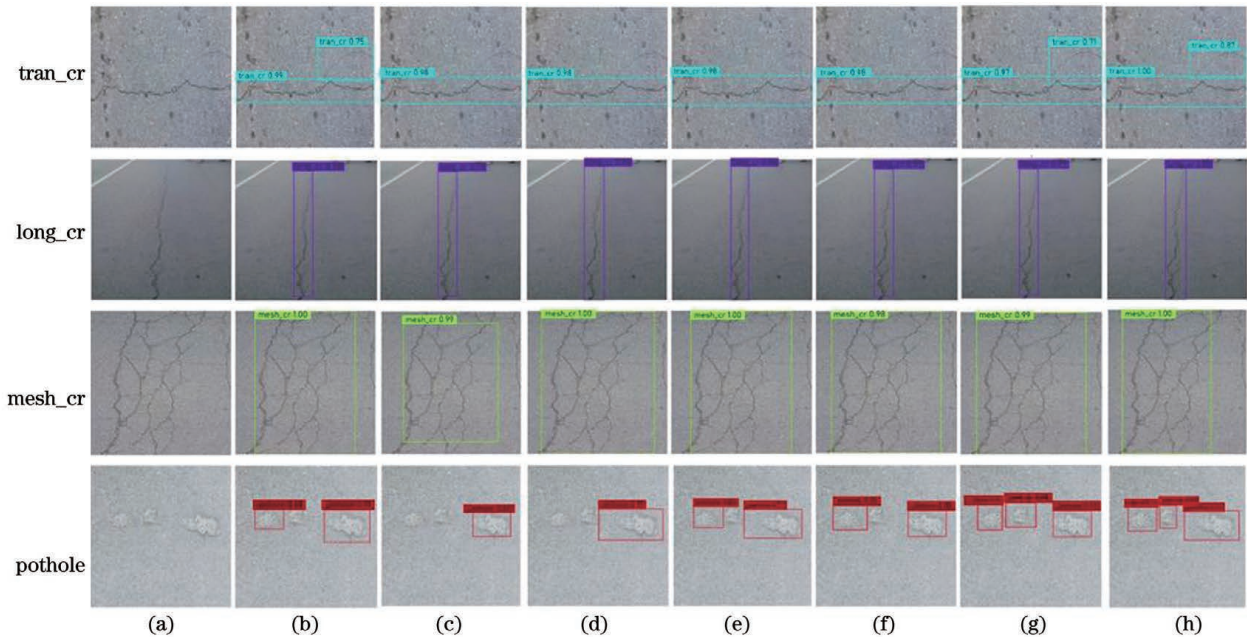


图 10 基于不同网络模型的公路路面病害检测效果。(a)输入图像;(b)Faster R-CNN;(c)SSD;(d)YOLOv3;(e)YOLOv4;(f)YOLOv4+DC;(g)YOLOv4+FL;(h)YOLOv4+DC+FL

Fig. 10 Detection results of road surface disease based on different network models. (a) Input images; (b) Faster R-CNN; (c) SSD; (d) YOLOv3; (e) YOLOv4; (f) YOLOv4+DC; (g) YOLOv4+FL; (h) YOLOv4+DC+FL

图 10 (b)~(h) 分别为基于 Faster R-CNN、SSD、YOLOv3、YOLOv4、YOLOv4+DC、YOLOv4+FL、YOLOv4+DC+FL 网络的检测结果。由图 10 可知,本文所提的基于 YOLOv4+DC+FL 的公路路面病害检测方法相较于 Faster R-CNN、SSD、YOLOv3、未改进的 YOLOv4 网络而言,对检测难度较大的多类、多尺度公路病害均获得了较好的检测效果。

## 5 结 论

为解决公路路面的横向裂缝、纵向裂缝、网状裂缝、坑槽类四种病害的多尺度特征及数据集不充分而导致的检测困难问题,本文在 YOLOv4 算法的基础上对网络卷积结构和损失函数进行了改进;并采用翻转、裁剪、亮度调整、噪声扰动等方法对实验获取的数据集进行增广,建立了公路路面病害数据集;同时,借助迁移学习思想,对公路路面病害检测网络进行预训练,然后对预训练权重进行参数微调,从而训练出基于改进 YOLOv4 的公路路面病害检测网络。与其他算法相比,本文所提算法对公路路面病害检测的 mAP 值较高、检测时间较短,可以有效地实现公路路面病害的精确、快速检测。

## 参 考 文 献

- [1] Bai C L. Types and causes analysis of highway surface diseases[J]. *Science & Technology Vision*, 2015(8): 244-245.  
白存亮. 高速公路路面病害类型及原因分析[J]. *科技视界*, 2015(8): 244-245.
- [2] Li L F, Sun R Y. Bridge crack detection algorithm based on image processing under complex background [J]. *Laser & Optoelectronics Progress*, 2019, 56(6): 061002.  
李良福, 孙瑞赞. 复杂背景下基于图像处理的桥梁裂缝检测算法 [J]. *激光与光电子学进展*, 2019, 56(6): 061002.
- [3] Zhang G S, Ge G Y, Zhu R H, et al. Gear defect detection based on the improved YOLOv3 network [J]. *Laser & Optoelectronics Progress*, 2020, 57(12): 121009.  
张广世, 葛广英, 朱荣华, 等. 基于改进 YOLOv3 网络的齿轮缺陷检测 [J]. *激光与光电子学进展*, 2020, 57(12): 121009.
- [4] Zhang L, Yang F, Zhang Y M, et al. Road crack detection using deep convolutional neural network [C]//2016 IEEE International Conference on Image Processing (ICIP), September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE Press, 2016: 3708-3712.
- [5] Cha Y J, Choi W, Büyüköztürk O. Deep learning-based crack damage detection using convolutional neural networks[J]. *Computer-Aided Civil and Infrastructure Engineering*, 2017, 32(5): 361-378.
- [6] Li G, Gao Z Y, Zhang X C, et al. Improved global convolutional network for pavement crack detection [J]. *Laser & Optoelectronics Progress*, 2020, 57(8): 081011.  
李刚, 高振阳, 张新春, 等. 改进的全局卷积网络在路面裂缝检测中的应用 [J]. *激光与光电子学进展*, 2020, 57(8): 081011.
- [7] Li G, Liu Q W, Wan J, et al. A novel pavement crack detection algorithm using interlaced low-rank group convolution hybrid deep network under a complex background [J]. *Laser & Optoelectronics Progress*, 2020, 57(14): 141031.  
李刚, 刘强伟, 万健, 等. 复杂背景下交错低秩组卷积混合深度网络的路面裂缝检测算法研究 [J]. *激光与光电子学进展*, 2020, 57(14): 141031.
- [8] Zhang N. Study on detection algorithm for road surface disease based on Faster R-CNN[D]. Nanchang: East China Jiaotong University, 2019: 39-44.  
张宁. 基于 Faster R-CNN 的公路路面病害检测算法的研究 [D]. 南昌: 华东交通大学, 2019: 39-44.
- [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 Press, 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 Press, 2015: 1440-1448.
- [11] Ren S, He K, 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] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 936-944.
- [13] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [C] // 2016 IEEE Conference on Computer Vision and



- Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [14] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]. Cham: Springer, 2016: 21-37.
- [15] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [16] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2020-08-17]. <https://arxiv.org/abs/1804.02767>.
- [17] Bochkovskiy A, Wang C Y, Liao M H Y. YOLOv4: optimal speed and accuracy of object detection [EB/OL]. (2020-04-23)[2020-08-17]. <https://arxiv.org/abs/2004.10934>.
- [18] He K M, Zhang X Y, Ren S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [19] Wang C Y, Mark Liao H Y, Wu Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 14-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1571-1580.
- [20] Mishra D. Mish: a self regularized non-monotonic neural activation function [EB/OL]. (2020-08-13)[2020-08-17]. <https://arxiv.org/abs/1908.08681v2>.
- [21] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 8759-8768.
- [22] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. AAAI-20 Technical Tracks, 2020, 34(7): 12993-13000.
- [23] Yang F, Zhang L, Yu S J, et al. Feature pyramid and hierarchical boosting network for pavement crack detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(4): 1525-1535.
- [24] Howard A G, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2020-08-17]. <https://arxiv.org/abs/1704.04861>.
- [25] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[EB/OL]. (2018-02-07)[2020-08-17]. <https://arxiv.org/abs/1708.02002>.
- [26] Tan C Q, Sun F C, Kong T, et al. A survey on deep transfer learning [EB/OL]. (2018-08-06)[2020-08-17]. <https://arxiv.org/abs/1808.01974>.