

8-25-2023

Improved Object Detection of YOLOv4 in Foggy Conditions

Shugang Liu

North China Electric Power University, Baoding 071000, China, lsg69@qq.com

Linkun Zhang

North China Electric Power University, Baoding 071000, China, 1021759619@qq.com

Haodong Du

North China Electric Power University, Baoding 071000, China

Hongtao Wang

North China Electric Power University, Baoding 071000, China

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Improved Object Detection of YOLOv4 in Foggy Conditions

Abstract

Abstract: Aiming at the low detection accuracy in foggy weather, a new defogging target detection method based on DeblurGANv2 and YOLOv4 is proposed. *In the method, image enhancement algorithm DeblurGANv2 in the generation countermeasure network is added to the preprocessing module of YOLOv4 to preprocess the foggy image and retain the high-quality texture and color information of the image, lightweight neural network ShuffleNet V2 is used to replace the CSPDarkNet53 network used for backbone feature extraction in YOLOv4 to improve the speed of model mark detection. Attention mechanism is added to the feature extraction network of YOLOv4 to enhance the recognition effect of small targets.* The experimental results show that the proposed method can reduce the large color difference and fog residue, and the mAP value in the rest data set reaches 86.56%. The result of practiced defogging target test is good.

Keywords

generative adversarial nets, DeblurGANv2, Coordinate Attention, ShuffleNet V2, defogging target detection

Recommended Citation

Liu Shugang, Zhang Linkun, Du Haodong, et al. Improved Object Detection of YOLOv4 in Foggy Conditions[J]. Journal of System Simulation, 2023, 35(8): 1681-1691.

雾天条件下改进 YOLOv4 的目标检测

刘书刚, 张林坤*, 杜昊东, 王洪涛

(华北电力大学, 河北 保定 071000)

摘要: 针对雾霾天气下现有的目标检测方法存在检测精度低的问题, 提出了一种基于 DeblurGANv2 与 YOLOv4 的去雾目标检测方法。在 YOLOv4 的预处理模块中加入生成对抗网络中的图像增强算法 DeblurGANv2, 对有雾的图像进行去雾预处理, 保留图像高质量的纹理和色彩信息。用轻量级神经网络 ShuffleNet V2 替代 YOLOv4 中用于主干特征提取的 CSPDarkNet53 网络, 提高模型目标检测速度。在 YOLOv4 的特征融合模块中加入注意力机制, 增强对小目标的识别效果。实验结果表明: 该方法能够减少色差较大和雾残留的问题, 在 RESIDE 数据集中 mAP 值达到了 86.56%, 在实际去雾目标测试中取得较好的效果。

关键词: 生成对抗网络; DeblurGANv2; Coordinate Attention; ShuffleNet V2; 去雾目标检测

中图分类号: TP183 文献标志码: A 文章编号: 1004-731X(2023)08-1681-11

DOI: 10.16182/j.issn1004731x.joss.22-0423

引用格式: 刘书刚, 张林坤, 杜昊东, 等. 雾天条件下改进 YOLOv4 的目标检测[J]. 系统仿真学报, 2023, 35(8): 1681-1691.

Reference format: Liu Shugang, Zhang Linkun, Du Haodong, et al. Improved Object Detection of YOLOv4 in Foggy Conditions[J]. Journal of System Simulation, 2023, 35(8): 1681-1691.

Improved Object Detection of YOLOv4 in Foggy Conditions

Liu Shugang, Zhang Linkun*, Du Haodong, Wang Hongtao

(North China Electric Power University, Baoding 071000, China)

Abstract: Aiming at the low detection accuracy in foggy weather, a new defogging target detection method based on DeblurGANv2 and YOLOv4 is proposed. In the method, image enhancement algorithm DeblurGANv2 in the generation countermeasure network is added to the preprocessing module of YOLOv4 to preprocess the foggy image and retain the high-quality texture and color information of the image, lightweight neural network ShuffleNet V2 is used to replace the CSPDarkNet53 network used for backbone feature extraction in YOLOv4 to improve the speed of model mark detection. Attention mechanism is added to the feature extraction network of YOLOv4 to enhance the recognition effect of small targets. The experimental results show that the proposed method can reduce the large color difference and fog residue, and the mAP value in the rest data set reaches 86.56%. The result of practiced defogging target test is good.

Keywords: generative adversarial nets; DeblurGANv2; Coordinate Attention; ShuffleNet V2; defogging target detection

收稿日期: 2022-04-27 修回日期: 2022-07-18

基金项目: 国家自然科学基金(61802124)

第一作者: 刘书刚(1969-), 男, 副教授, 硕士, 研究方向为人工智能。E-mail: lsg69@qq.com

通讯作者: 张林坤(1996-), 男, 硕士生, 研究方向为计算机视觉。E-mail: 1021759619@qq.com

0 引言

传统的目标检测是在复杂场景下快速准确地识别出目标的类别信息并且能够准确定位出目标在图像中的位置，目标检测是计算机视觉领域中最基本、最常见的问题、随着人工智能和深度学习的快速发展，传统的目标检测也取得了一系列进展。常见的目标检测可以分为单阶段目标检测和两阶段目标检测。常见的单阶段目标检测有 SSD(single shot multibox detector)^[1]、YOLO 系列^[2-5]和 Retina-Net^[6]。常见的两阶段目标检测有特征金字塔(feature pyramid networks, FPN)^[7]和 R-CNN^[8]。

来由于环境污染导致很多城市雾霾天气频发，会对目标检测产生影响，降低目标检测的精度。随着人工智能与机器学习的快速发展，传统的目标检测方法在检测精度和速度方面均取得明显进步，但在雾霾场景下的图片质量，使神经网络对图像特征提取不准确，严重影响目标检测的准确性。

目前，去雾目标检测大体可以分为两大类，一类是基于去雾和检测的无关联模型，即先进行去雾操作后进行目标检测，如 Li 等^[9]提出一种去雾网络 POR-Net 结合 Faster R-CNN^[10]的无关联去雾目标检测方法，检测精度提高了 2%。另一类是将去雾算

法和目标检测算法进行联合优化，同时进行去雾和检测，如 Li 等^[11]提出一种基于 AODNet 和 Faster R-CNN 的端到端去雾目标检测方法，提升了雾天目标检测的精度，但是去雾效果不理想，去雾图片会出现伪影等。此后，此团队还进行一系列研究^[12-14]，并提出有雾数据集 RESIDE^[12]。

为了提高雾天目标检测的精度和速度，本文提出一种基于 DeblurGANv2^[15]和 YOLOv4^[5]的雾天目标检测方法，首先将质量不佳的有雾图片通过 DeblurGANv2 模块进行图片预处理，得到清晰的去雾图片，将处理好的图片作为 YOLOv4 网络的输入。本文对 YOLOv4 算法进行一些改进提升。首先，用轻量级神经网络 ShuffleNet V2^[16]替代 YOLOv4 中用于主干特征提取的 CSPDarkNet53 网络，提高模型检测速度。其次，在 YOLOv4 的特征融合模块加入注意力机制，增强对小目标的识别效果。

1 改进之后的 YOLOv4

改进的方法由 5 部分构成，分别是去雾预处理模块 DeblurGANv2、主干特征提取网络 ShuffleNet V2、PANet 特征融合模块、SPP 模块和 YOLO Head，如图 1 所示。

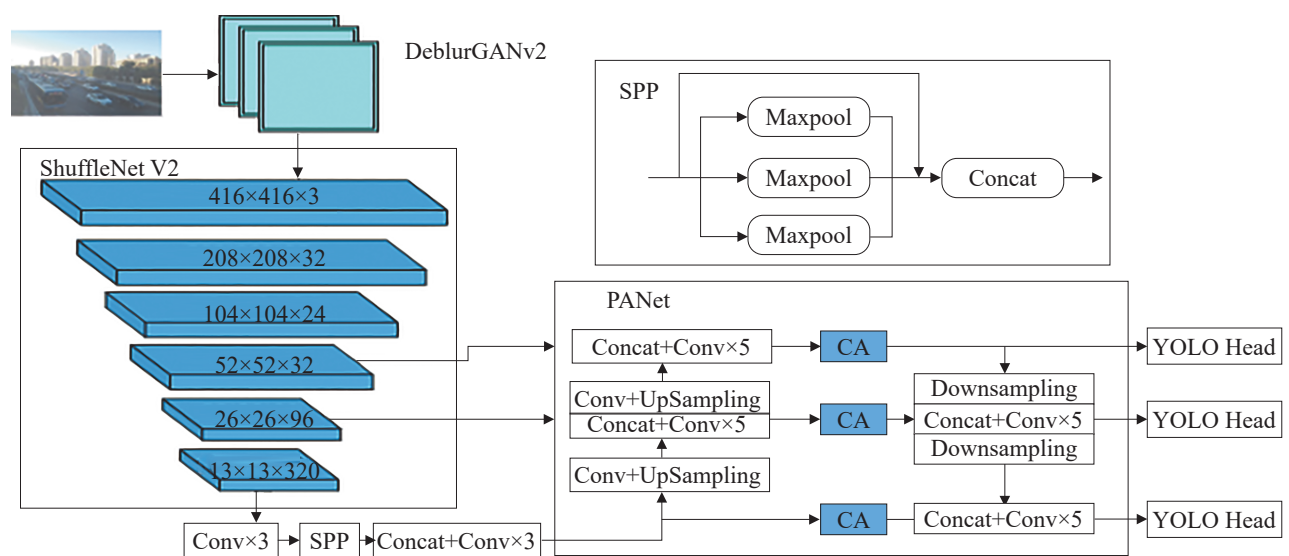


图 1 改进后的 YOLOv4 网络
Fig. 1 Improved YOLOv4 network

<http://www.china-simulation.com>

• 1682 •

1.1 DeblurGANv2 去雾预处理

为了实现去雾功能, 本文采取两阶段目标检测的方法, 即图像先去雾再进行目标检测。在 YOLOv4 网络的预处理模块中使用 DeblurGANv2 将有雾的图片处理, 最大程度上保留原始图片中的色彩信息和纹理信息, 并生成接近于真实场景的去雾图像。当网络接收到大小为 416 像素×416 像素的图像输入时, 首先 DeblurGANv2 网络进行去雾预处理, 在完成去雾过程后, 通过反卷积网络层将图像大小重置为 416 像素×416 像素尺寸, 经过图像去雾算法处理后的图片, 比之前增加了更多的细节信息, 提高了 YOLOv4 对图像的认识准确率。

DeblurGANv2 的网络结构如图 2 所示。与 DeblurGANv1^[17] 不同, DeblurGANv2 是基于 FPN 结构进行特征融合的。FPN 不仅可以融合多个不同的尺度信息, 并且在速度和精度之间也取得了较好的均衡。FPN 结构包括自底向上和自顶向下路径。自底向上路径是一个用于特征提取的卷积网络。DeblurGANv2 的主体包括 2 个尺寸不同的鉴别器和用于图像恢复的生成器。鉴别器在考虑了全局尺寸和局部尺寸的差异之后, 将 FullGAN 和 PatchGAN 鉴别器进行结合, 用以处理复杂的模糊问题。生成器中使用 FPN 框架来提取不同的

特征映射层, 其中, 自下而上的过程为卷积神经网络提取特征, 在提取特征过程中图像进行下采样。将下采样结束后得到的采样结果与自上而下的上采样结果进行横向链接, 获得更多的特征信息, 最终得到 5 种不同比例尺的输出特征图。将这 5 种特征图下采样到输入图片大小的 1/4, 并将 5 种采样结果连接到一个张量, 这个张量包含不同层次的语义信息。在 DeblurGANv2 网络中加入上采样层和卷积层来恢复清晰图像和去除伪影。

当 DeblurGANv2 使用不同的主干网进行训练时, 可以获得不同的图像处理效果。为了获得更高效的模糊图像处理方法, 本文选择 Inception 作为 DeblurGANv2 的主干网络。

1.2 轻量级 ShuffleNet V2 替代 CSPDarkNet53

为了有效减少算法中模型的参数, 减小模型大小, 提高算法的目标检测速度, 本文用 ShuffleNet V2 作为 YOLOv4 的主干特征提取网络。

Ma 等在 ShuffleNet V2 中提出 4 条高效且轻量化的网络设计准则, 并遵循这 4 条原则对 ShuffleNet V1^[18] 重新设计得到 ShuffleNet V2。ShuffleNet V2 网络结构如图 3 所示, ShuffleNet V2 网络模型的基本组成单元大致可分为 2 种, 如图 4 所示。

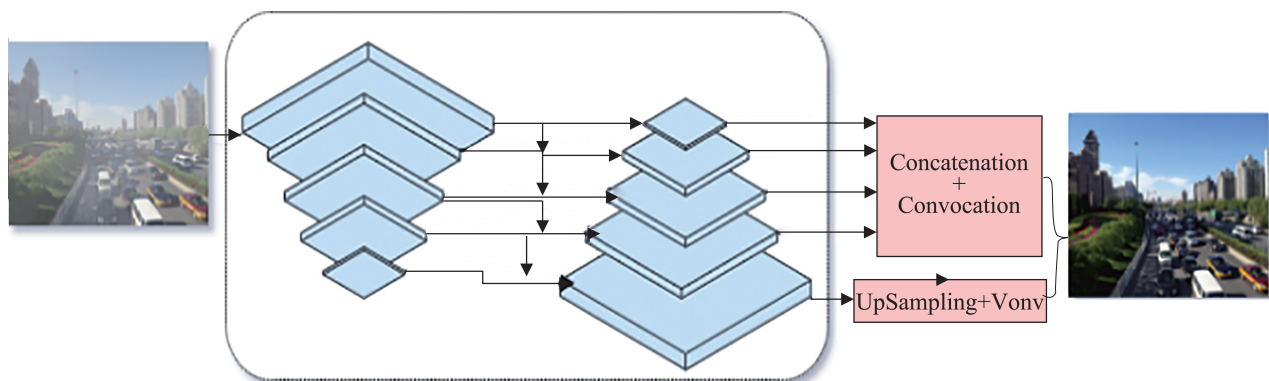


图2 DeblurGANv2
Fig. 2 DeblurGANv2

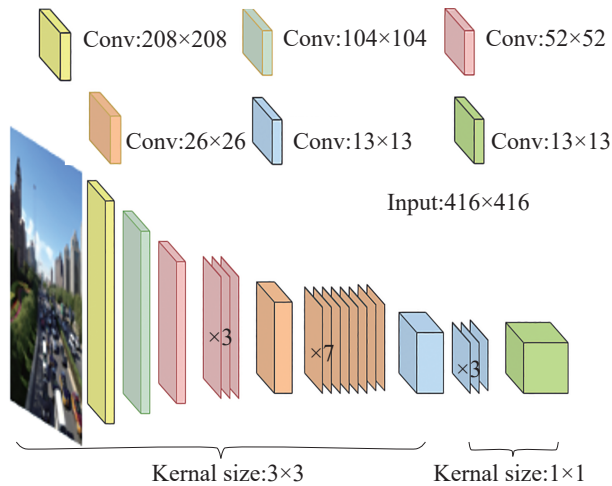


图3 ShuffleNet V2网络
Fig. 3 ShuffleNet V2

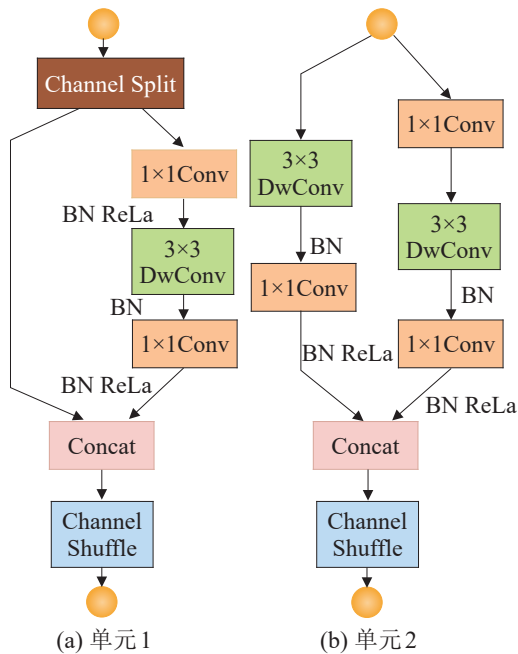


图4 ShuffleNet V2的2个模块
Fig. 4 ShuffleNet V2 two modules

在 ShuffleNet V2 单元 1 中, 先对输入的特征图进行通道划分(Channel Split), 划分成 2 个分支, 通道数各占 1/2。左边分支恒定不变。右分支将使用相同数量的输入和输出通道数, 需要经过 3 个步长为 1 的卷积。其中, 1×1 卷积表示普通卷积, 3×3 卷积表示深度卷积(DwConv)。当 3 个卷积完成后, 2 个分支会进行 Concat 操作, 通道数相加的同时特征进行融合, 最后进行通道随机混合操

作(Channel Shuffle), 使不同的通道充分融合。与单元 1 不同, ShuffleNet V2 单元 2 直接将特征图输入到 2 个分支。2 个分支都使用步长为 2 的 3×3 深度卷积, 对特征图的长和宽进行降维, 起到减少网络计算量的作用。2 个分支输出后进行 Concat 操作, 通道数相加后是原始输入的 2 倍, 增加了网络的宽度, 可在不显著增加 FLOPs 的情况下增加通道的数量, 使网络提取特征能力更强。最后, 同样进行通道混洗, 实现不同通道之间的信息交流。

1.3 注意力机制

目标像素值的大小可以通过卷积神经网络中周围像素的大小来计算, 但是, 局部信息计算的结果往往会导致全局信息的丢失和偏差。为了缓解信息偏差, 不仅可以使使用更大的卷积滤波器, 还可以构建更深层的卷积神经网络, 但计算量上升较快, 且影响不显著。因此, 本文通过加入注意力机制来提高目标检测的精度。

不同于 SE(squeeze-and-excitation)^[19] 和 CBAM (convolutional block attention module)^[20], 注意力机制(CA)^[21] 是一种新型的轻量级注意力方法。CA 通过精确的位置信息对长期依赖关系和通道关系进行编码, 获得较为全面的空间信息。CA 工作流程可分为 Coordinate 信息嵌入和 Coordinate Attention 生成 2 个步骤。网络结构如图 5 所示。

Coordinate 信息嵌入阶段使用一种特殊的池化方式, 使用大小为 $(H, 1)$ 和 $(1, W)$ 的池化核, 对输入 X 的水平和垂直方向的各个通道进行编码, 在第 c 通道中高度为 h 的输出可表示为

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (1)$$

第 c 通道中宽度为 w 的输出可表示为

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (2)$$

经过上述 2 种转换, 注意力模块可以沿一个空间方向捕获长期依赖关系, 并且获得沿着空间方向的准确位置信息, 帮助检测网络更准确地识别出感兴趣的目标。

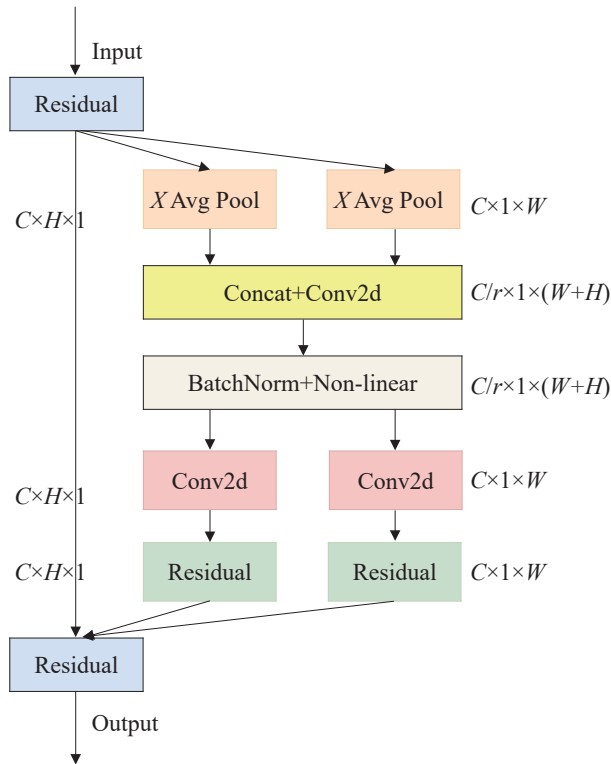


图5 CA网络
Fig. 5 CA network

经过第1阶段 Coordinate 信息嵌入后, 在第2阶段将第1阶段完成的2种变换进行 Concatenate 操作, 然后使用 1×1 卷积进行变换操作。

$$f = \delta(F_1([z^h, z^w])) \quad (3)$$

式中: δ 为非线性激活函数; $[\]$ 为 Concatenate 操作; f 为对空间位置信息分别沿垂直方向和水平方向进行编码并得到的中间特征映射。将 f 沿着空间维数分解成2个不同的张量 $f^w \in \mathbb{R}^{C/r \times W}$ 和 $f^h \in \mathbb{R}^{C/r \times H}$, 最后再用2个 1×1 卷积把2个张量转换成具有相同通道数的张量:

$$g^h = \delta(F_h(f^h)) \quad (4)$$

$$g^w = \delta(F_w(f^w)) \quad (5)$$

式中: δ 为非线性激活函数。为了进一步降低算法计算开销, 一般会选择一个缩减比 γ 对通道数进行缩减, 然后对 g^w 和 g^h 分别进行扩展, 得到最终的注意力权重。Coordinate Attention Block 的最后输出为

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(i) \quad (6)$$

CA 模块可以在目标检测中取得理想的效果。因此, 为提高 YOLOv4 的检测精度, 将 CA 模块插入 CSPNet 结构的3个输出通道后对 YOLOv4 进行修改, 可以获得更有针对性的网络结构学习效果。

1.4 损失函数

YOLOv4 目标检测算法的损失函数可以由3部分来表示:

$$l = l_{\text{location}} + l_{\text{confidence}} + l_{\text{class}} \quad (7)$$

式中: l_{location} 为位置损失; $l_{\text{confidence}}$ 为置信度损失; l_{class} 为类别损失。

为了避免使用 IoU 函数造成训练过程中的发散和振荡问题, YOLOv4 中的 l_{location} 采用了 CIoU 进行损失函数的计算, CIoU 考虑到了边界框回归的重叠率。距离和纵横比:

$$CIoU = IoU - \frac{\rho^2(b, b^{g'})}{c^2} - av \quad (8)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{g'}}{h^{g'}} - \arctan \frac{w}{h} \right)^2 \quad (9)$$

$$a = v / (1 - IoU + v) \quad (10)$$

式中: IoU 为真实框和预测框的交集和并集的比值; b 、 $b^{g'}$ 为预测框和真实框的中心点; $\rho^2(b, b^{g'})$ 为 b 和 $b^{g'}$ 之间的欧式距离; c 为能够同时包含预测框和真实框的最小闭合区域的对角线距离。

YOLOv4 各部分损失计算式为

$$l_{\text{location}} = 1 - CIoU \quad (11)$$

$$l_{\text{class}} = \sum_{i=0}^{S^2} \prod_{i \text{ obj}} \sum_{c \in \text{class}} (p_i(c) - \hat{p}_i(c))^2 \quad (12)$$

$$l_{\text{confidence}} = \sum_{i=0}^{S^2} \sum_{j=0}^B \prod_{ij \text{ obj}} (C_i - \hat{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B \prod_{ij} (C_i - \hat{C}_i)^2 \quad (13)$$

式中: $\prod_{i \text{ obj}}$ 为第 i 个单元格中是否含有需要被检测的目标, 如果有为1, 没有则为0; S 为将输入的图片分割成 S 个单元格; \prod_{ij} 为单元格 i 中第 j 个预

测框是否负责此次检测； $p_i(c)$ 为目标 c 属于某类别的概率； $\hat{p}_i(c)$ 为预测目标属于某一类别的真实值； λ_{noobj} 为单元格内无目标的权重系数； C_i 为置信度分数； \hat{C}_i 为真实框与预测框相交的部分。

2 实验环境及方法

2.1 实验数据处理

本文采用的训练数据集为 RESIDE 中的 OTS, RESIDE 中的 OTS 包含 8 971 张清晰图片, 每一张清晰图片生成 10 张含雾数据集。本文选取其中包含人、汽车、卡车、自行车、交通信号灯 5 类物体的清晰图片和与之对应生成的模糊图片各 1 张。图 6 为 RESIDE 含雾部分图片。

最终选取清晰数据集和含雾数据集各 3 995 张, 用于训练 DeblurGANv2, 进而用于去雾预处理。

理。利用 Labelimg 软件对上述 2 个数据集进行标注, 用于之后的目标检测训练和测试。标记完所有数据集后, .xml 文件被转换成 .txt 文件, 其中, 包含图像路径、目标类别和目标手势的位置信息。为了解决锚框尺寸差异问题, 使用 K-means 算法对本文使用的物体类别进行聚类分析, 得到新的锚框尺寸。按照 9:1 的比例划分训练集和测试集, 实验中不同类别标签的实际数量如表 1 所示。

2.2 实验平台

实验环境详细信息为 Windows10 系统, CPU 为 Inter(R)Xeon(R)Gold5218CPU@2.30 GHz, GPU 为 NVIDIA GeForce RTX 2080 Ti, 编程语言为 python3.8, 深度学习框架为 pytorch1.6, 目标检测实验初始学习率设定为 0.001, batch size 大小为 15。整个训练过程中总计有 300 个 epoch。

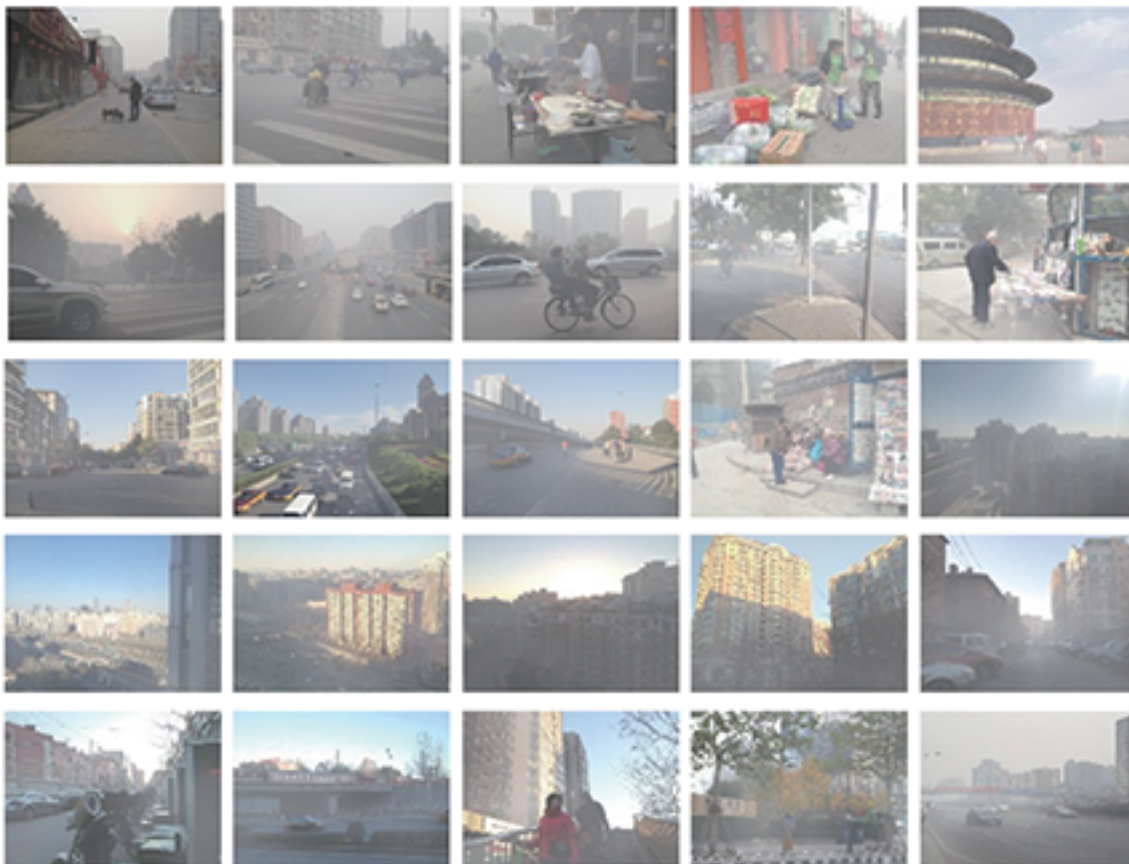


图 6 RESIDE 数据集
Fig. 6 RESIDE dataset

表1 不同标签在数据集中的数量
Table 1 Number of different labels in dataset

数据集	交通信号灯	卡车	人	汽车	自行车
训练集	7 586	6 458	9 632	12 561	263
测试集	836	665	1 193	1 552	35
合计	8 422	7 123	10 825	14 413	298

2.3 实验评价标准

本文对上述数据集进行相关实验。对于去雾效果使用峰值信噪比 PSNR 和结构相似性 SSIM 对图像质量进行评价。对于目标检测采用平均精度 (AP)、平均精度均值(mAP)、ROC 曲线、P-R 曲线和 FPS 对相关算法的精度进行评估, 分析算法的先进性和有效性。

预测为正的样本中实际正样本的比例为准确率 P ; 实际正样本中被预测为正样本的概率为召回率 R ; 预测样本的平均准确率为平均精度 AP ; 多个验证集之间求 AP 的平均值得到平均精度均值 mAP :

$$P = P_t / (P_t + P_f) \quad (14)$$

$$R = P_t / (P_t + N_f) \quad (15)$$

$$AP = \int_0^1 P(R) d(R) \quad (16)$$

$$mAP = \frac{\sum AP}{K} \quad (17)$$

式中: P_t 为被正确分类的正样本; P_f 为被错误分类的正样本; N_f 为被错误分类的负样本。

网络样本分配示意图如图7所示。

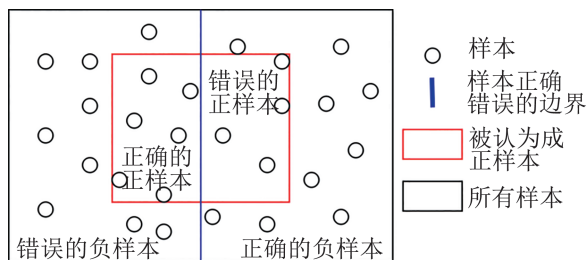


图7 网络样本分配

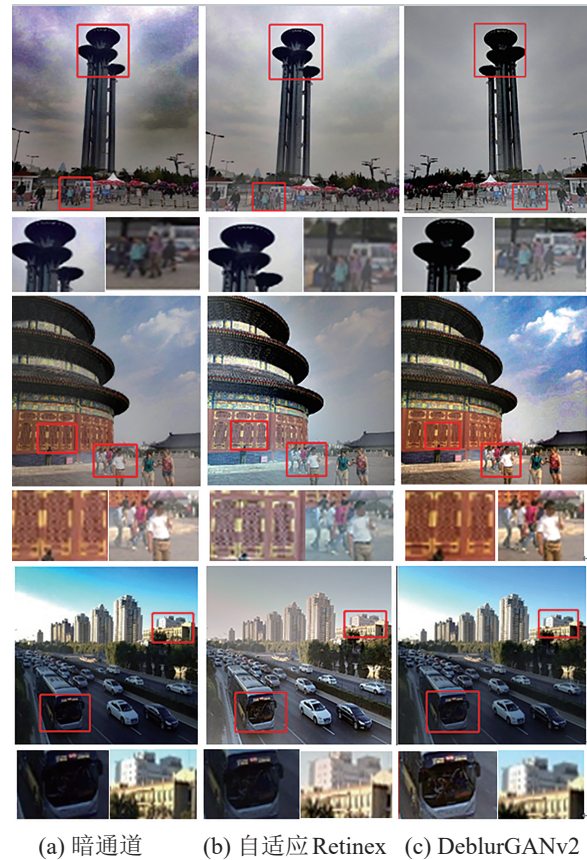
Fig. 7 Network sample allocation

3 实验结果及分析

3.1 DeblurGANv2 去雾效果

本文将准备好的清晰照片和含雾照片进行对

抗训练, 得到去雾模块的权值。使用 DeblurGANv2 网络模块对输入含雾图像进行处理, 实现图像去雾。为了验证 DeblurGANv2 的优越性, 将 DeblurGANv2 与暗通道先验法^[22]和自适应 Retinex^[23]进行对比实验。图8给出了在不同算法下合成雾霾图像的去雾处理结果。可以看出, 暗通道先验算法处理过的图像整体亮度有所降低, 在天空部分出现块效应。自适应 Retinex 算法处理后亮度差异较大, 产生失真。本文所采用的 DeblurGANv2 去雾效果最佳而且保留了与原始图像高度一致的纹理信息, 去雾还原效果比其他2种更优。



(a) 暗通道 (b) 自适应 Retinex (c) DeblurGANv2

图8 去雾效果对比

Fig. 8 Comparison of defogging effect

为了更客观地评价不同去雾算法的去雾效果, 本文使用峰值信噪比 (PSNR) 和结构相似性指数 (SSIM) 对不同去雾算法处理的图片进行客观评价。从表2可以看出, DeblurGANv2 的 PSNR 和 SSIM 值好于暗通道、自适应 Retinex, 其中, DeblurGANv2

在图片类型为住宅的峰值信噪比高达 28.78 dB，可以有效还原出原图片的细节信息，并且与原图片的结构相似度高达 76%，优于暗通道先验法的 69% 和自适应 Retinex 的 74%。DeblurGANv2 在公路和广场这 2 组图片的峰值信噪比相比暗通道先验和自适应 Retinex 分别提高了 4.73 dB，2.14 dB 和 4.44 dB，0.93 dB。

表2 不同模型的去雾效果评价
Table 2 Evaluation of defogging effect of different models

图像	指数	暗通道先验法	自适应 Retinex	DeblurGANv2
住宅	PSNR/dB	23.70	26.68	28.78
	SSIM/%	69	74	76
公路	PSNR/dB	23.50	26.09	28.23
	SSIM/%	68	74	75
广场	PSNR/dB	23.10	26.61	27.54
	SSIM/%	63	70	72

3.2 目标检测效果分析

在上述实验环境下，绘制出本文算法与原始 YOLOv4 的平均损失函数曲线。由图 9 可知，本文方法在前 50 次呈现指数下降趋势，50~100 次下降趋势趋于平缓，100 次以后损失函数值稳定在 0.9 左右。YOLOv4 稳定在 1.1 左右。实验结果表明，本文模型具有较好的训练效果。

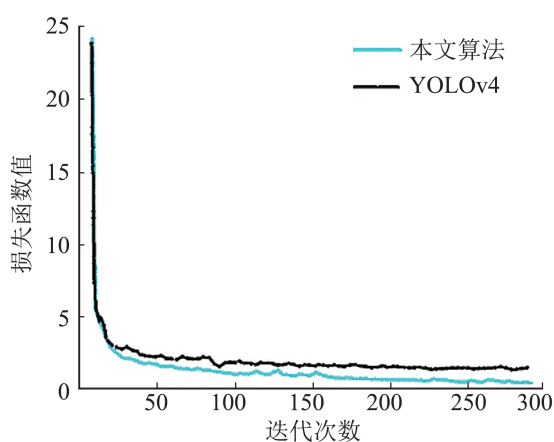


图9 损失函数对比

Fig. 9 Loss function comparison

如图 10 所示，本文模型的整个模型 mAP 值为

86.56%，比 YOLOv4 的 75.63% 提高了 10.93%。所以本文模型具有较好的识别精准度。根据实验可知，本文算法在目标检测精度和训练效果方面都比原 YOLOv4 有所提升。

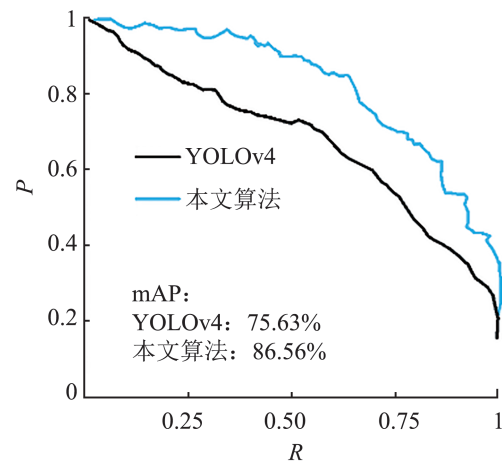


图 10 P-R 曲线

Fig. 10 P-R curve

为了便于对模型的综合能力进行评价，本文绘制出 ROC 曲线。ROC 曲线可以反映网络在不同任务需求下的性能效果。在不同的应用任务中，可以根据需求选择不同的截止点来比较模型的性能。位置越靠前，要求的精度越高。相反，则需要更高的查全率。真阳性率阈值范围在 0~1。因此，曲线越靠近 (0, 1) 点，网络性能越全面。2 种不同网络的 ROC 曲线如图 11 所示。从图 11 中可以看到，本文算法最接近 (0, 1) 点，在任何一个极限值条件下，本文所提模型的效果都要优于 YOLOv4 模型。

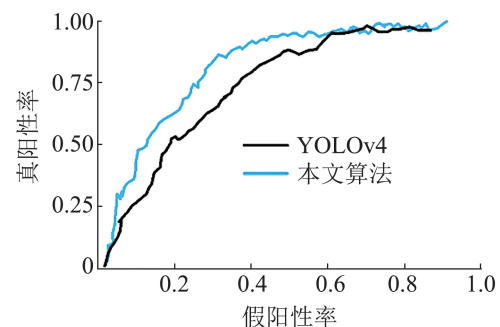


图 11 ROC 曲线

Fig. 11 ROC curve

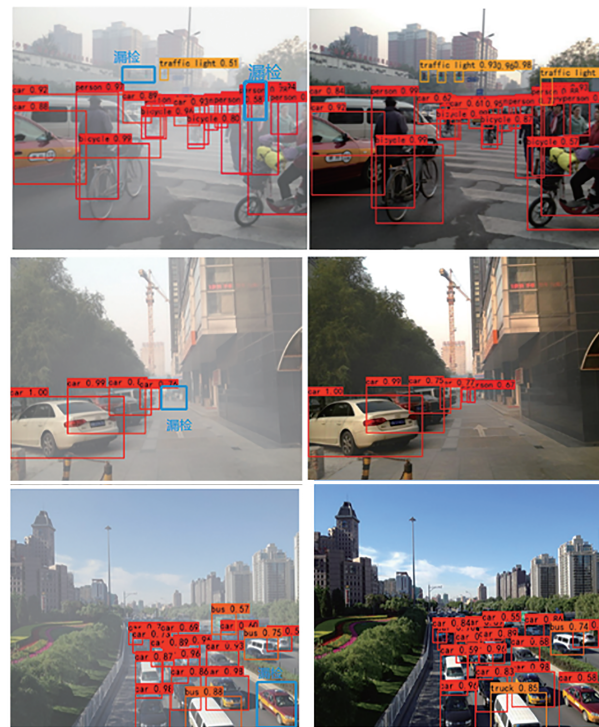
为了更直观地发现更优的检测效果, 将本文算法和YOLOv4的目标检测结果进行对比, 如图12所示。相比于原始的YOLOv4模型, 本文模型在雾霾场景下的识别效果有明显的提升, 有效避免了原始YOLOv4模型在雾霾场景下的漏检状况。如第1张图片中的交通信号灯, 由于YOLOv4模型中图片未过去雾处理, 含有较多的干扰信息, 导致YOLOv4出现漏检的情况, 第2张图片中的人和第3张图片中右下角的车, 都出现漏检的情况。本文算法在目标的预测置信度上也有不同程度的提高, 由此可见本文算法取得了较好的识别效果。

3.3 不同模型雾天目标检测效果对比

为了进一步证明本文算法的有效性和先进性, 将Fast-RCNN与YOLOv4分别与暗通道先验、自适应Retinex、DeblurGANv2算法进行简单结合, 形成YOLOv3-暗通道、YOLOv3-自适应Retinex、YOLOv3-DeblurGANv2、YOLOv4-暗通道、YOLOv4-自适应Retinex、YOLOv4-DeblurGANv2共6种雾天检测模型与本文模型在同一数据集下分别进行实验。实验得到的mAP值, 如表3所示。

由表3可以看出, 本文模型mAP高达86.56%, 比YOLOv3-暗通道提高了10.02%、比YOLOv3-自适应Retinex提高了6.18%、比YOLOv3-DeblurGANv2提高了4.87%、比YOLOv4-暗通道提高了2.11%、比YOLOv4-自适应Retinex提高了1.03%。检测速度为46帧/s, 高

于其他模型。虽然YOLOv4-DeblurGANv2模型mAP值比本文模型高了0.96%, 但是本文模型的检测速度比YOLOv4-DeblurGANv2高了6帧/s, 在牺牲较少的精度下获得了较快的检测速度。实验结果表明: 本文模型在一定程度上能够提高对原始模型评价的性能指标, 较其他模型具有优势。



(a) YOLOv4

(b) 本文算法

图12 YOLOv4与本文算法识别对比

Fig. 12 Comparison of proposed method with YOLOv4 recognition

表3 不同模型的mAP
Table 3 mAP value of different models

算法	mAP/%	AP/%					检测速度/(帧/s)
		交通信号灯	卡车	人	汽车	自行车	
YOLOv3-暗通道	76.54	67.21	81.56	82.98	82.50	68.45	35
YOLOv3-自适应Retinex	80.38	69.95	85.23	87.96	88.34	70.45	33
YOLOv3-DeblurGANv2	81.69	71.23	85.27	86.54	84.78	81.63	34
YOLOv4-暗通道	84.45	73.12	90.36	91.71	90.7	76.36	41
YOLOv4-自适应Retinex	85.53	74.95	91.45	90.54	91.29	79.42	39
YOLOv4-DeblurGANv2	87.52	77.86	90.18	92.86	93.05	83.65	40
本文算法	86.56	75.26	90.17	92.08	91.88	83.41	46

4 结论

本文提出了一种基于 DeblurGANv2 与 YOLOv4 的去雾目标检测方法。该方法在 YOLOv4 的预处理阶段使用生成对抗网络的图像增强算法(DeblurGANv2)对图像进行去雾预处理,保留图像中超高质量的纹理信息,同时对 YOLOv4 网络结构进行再设计。

目前,本文所提算法仍有需要改进的地方,如在反光和雾霾严重的情境下出现错检和漏检等问题,这是下一步需要完善的地方。在未来的工作中,本研究会继续改进网络结构,增强算法的泛化能力,以提高去雾目标检测的鲁棒性,同时将算法运用到实际的工程中。

参考文献:

- [1] Liu Wei, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]//Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [2] 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). Piscataway, NJ, USA: IEEE, 2016: 779-788.
- [3] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway, NJ, USA: IEEE, 2017: 6517-6525.
- [4] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement[EB/OL]. (2018-04-08)[2022-05-24]. <https://arxiv.org/abs/1804.02767>.
- [5] Bochkovskiy A, Wang C Y, Liao Hongyuan. YOLOv4: Optimal Speed and Accuracy of Object Detection[EB/OL]. (2020-04-23) [2022-06-04]. <https://arxiv.org/abs/2004.10934>.
- [6] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[C]//2017 IEEE International Conference on Computer Vision(ICCV). Piscataway, NJ, USA: IEEE, 2017: 2999-3007.
- [7] Lin T Y, Piotr Dollár, Girshick R, et al. Feature Pyramid Networks for Object Detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2017: 936-944.
- [8] 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. Piscataway, NJ, USA: IEEE, 2014: 580-587.
- [9] Li Chongyi, Guo Chunle, Guo Jichang, et al. PDR-net: Perception-inspired Single Image Dehazing Network With Refinement[J]. IEEE Transactions on Multimedia, 2020, 22(3): 704-716.
- [10] Ren Shaoqing, He Kaiming, 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.
- [11] Li Boyi, Peng Xiulian, Wang Zhangyang, et al. AOD-net: All-in-one Dehazing Network[C]//2017 IEEE International Conference on Computer Vision(ICCV). Piscataway, NJ, USA: IEEE, 2017: 4780-4788.
- [12] Li Boyi, Ren Wenqi, Fu Dengpan, et al. Benchmarking Single Image Dehazing and Beyond[J]. IEEE Transactions on Image Processing, 2019, 28(1): 492-505.
- [13] Liu Yu, Zhao Guanlong, Gong Boyuan, et al. Improved Techniques for Learning to Dehaze and Beyond: A Collective Study[EB/OL]. (2018-07-30) [2022-06-06]. <https://arxiv.org/abs/1807.00202>.
- [14] Yang Wenhan, Yuan Ye, Ren Wenqi, et al. Advancing Image Understanding in Poor Visibility Environments: A Collective Benchmark Study[J]. IEEE Transactions on Image Processing, 2020, 29: 5737-5752.
- [15] Kupyn O, Martyniuk T, Wu Junru, et al. DeblurGAN-v2: Deblurring (Orders-of-magnitude) Faster and Better[C]//2019 IEEE/CVF International Conference on Computer Vision(ICCV). Piscataway, NJ, USA: IEEE, 2019: 8877-8886.
- [16] Ma Ningning, Zhang Xiangyu, Zheng Haitao, et al. ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design[C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 122-138.
- [17] Kupyn O, Budzan V, Mykhailych M, et al. DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 8183-8192.
- [18] Zhang Xiangyu, Zhou Xinyu, Lin Mengxiao, et al. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 6848-6856.
- [19] Hu Jie, Shen Li, Sun Gang. Squeeze-and-excitation Networks[C]//2018 IEEE/CVF Conference on Computer

- Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 7132-7141.
- [20] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional Block Attention Module[C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [21] Hou Qibin, Zhou Daquan, Feng Jiashi. Coordinate Attention for Efficient Mobile Network Design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Piscataway, NJ, USA: IEEE, 2021: 13708-13717.
- [22] He Kaiming, Sun Jian, Tang Xiaoou. Single Image Haze Removal Using Dark Channel Prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.
- [23] 李梦蕊, 柳晓鸣, 常婧. 基于颜色衰减先验的自适应 Retinex 去雾算法[J]. 计算机仿真, 2021, 38(1): 354-358.
- Li Mengrui, Liu Xiaoming, Chang Jing. An Adaptive Retinex of Image Haze Removal Based on Color Attenuation Prior[J]. Computer Simulation, 2021, 38(1): 354-358.