

DOI: 10.13382/j.jemi.B2407719

基于深度学习的雾天非法采砂船只辨识方法*

张慧¹ 王新猛¹ 程占军² 杨一涛¹

(1. 南京警察学院信息技术学院 南京 210023; 2. 南京警察学院警犬研究所 南京 210012)

摘要:为解决长江流域非法采砂船只识别困难、监测效率低、精度差等问题,提出一种基于深度学习的雾天非法采砂船只辨识方法。首先,提出改进的生成对抗网络对采集图像进行去雾处理,得到清晰水域图像。其中,利用融合特征注意力机制的生成器提取雾天环境下船只的复杂纹理特征;在判别器中加入谱归一化操作,解决网络训练梯度消失的问题;改进损失函数,引入循环一致性损失保证生成图像与原始图像间的结构一致性。其次,提出融合注意力机制的YOLOv8算法,有效增强网络对图像重要特征的提取能力,实现对清晰水域图像中非法采砂船只的精准定位与辨识。实验结果表明,本研究提出的改进生成对抗网络对图像的去雾效果较好,峰值信噪比(PSNR)与结构相似性(SSIM)分别为31.86和0.64,较Cycle GAN和GC-GAN算法分别提升了3.6%~13.1%、4.9%~56.1%。去雾后的图像经过融合注意力机制的YOLOv8算法处理,可实现对非法采砂船只的准确识别与准确定位,其mAP@0.5:0.95和帧率可达到89.6%、36 FPS,满足公安实战对精度与速度的要求,可有效提高长江流域非法采砂信息化、智能化监管与执法水平。

关键词:非法采砂;图像去雾;生成对抗网络;目标检测;YOLOv8算法。

中图分类号: TP391.4; TN29 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Identification method of illegal sand mining vessels in foggy conditions based on deep learning

Zhang Hui¹ Wang Xinmeng¹ Cheng Zhanjun² Yang Yitao¹

(1. School of Information Technology, Nanjing Police University, Nanjing 210023, China;

2. Police Dog Technology Institute, Nanjing Police University, Nanjing 210012, China)

Abstract: To address the challenges of low monitoring efficiency and poor accuracy in identifying illegal sand mining vessels in foggy conditions, this study proposes a deep learning-based identification method. First, an improved generative adversarial network is employed to dehaze collected images, producing clear water area images. The generator integrates the feature attention mechanism to extract complex texture features of vessels in foggy environments, while spectral normalization is added to the discriminator to prevent gradient vanishing during training. Cycle consistency loss is introduced to ensure structural consistency between generated and original images. The CBAM attention mechanism is integrated into YOLOv8 algorithm to improve feature extraction, enabling precise localization and identification of illegal sand mining vessels in dehazed images. The improved GAN achieves superior dehazing performance, with PSNR and SSIM values of 31.86 and 0.64, representing 3.6%~13.1% and 4.9%~56.1% improvements over Cycle GAN and GC-GAN, respectively. The YOLOv8 enhanced by CBAM achieves mAP@0.5:0.95 of 89.6% and FPS of 36 on dehazed images, meeting the accuracy and speed requirements for practical law enforcement. The proposed method effectively enhances the informatization and intelligence of illegal sand mining supervision and enforcement in the Yangtze River Basin.

Keywords: illegal sand mining; image dehazing; generative adversarial network; object detection; YOLOv8 algorithm

0 引言

随着城镇化的推进,砂石作为建筑物主要原材料需求量剧增。非法采砂成本低,利润空间大,非法采砂现象日益猖獗。计算机视觉领域的快速发展给非法采砂行为治理带来了新思路。通过对监控视频进行目标船只检测,确定船只位置并对其进行有效识别,可以对非法采砂行为进行精准打击^[1-2]。然而,长江流域雾霾天气频发^[3],非法采砂船只识别困难。因此,有必要先有针对性地去除雾天环境下采集图像中的雾气,进而对非法采砂船只进行实时、高效辨识。

目前的图像去雾方法主要有3类:图像增强去雾算法、图像复原去雾算法、基于深度学习的图像去雾算法^[4]。图像增强主要通过通过对图像边缘、轮廓、对比度等退化特征进行处理,消除图像非重点区域的杂质,突出图像细节,提高图像清晰度,主要包括直方图均衡化、小波变换等算法。其中,直方图均衡化算法通过调整图像的直方图分布来改变图像各像素灰度,从而增强图像对比度,达到图像去雾的目的^[5]。小波变换算法则是通过将图像分解为不同频率域的组合,提取图像多尺度结构信息并进行重构来突出图像细节完成去雾操作^[6]。然而此类方法仅靠提高全局或局部对比度解决图像模糊问题,去雾效果较差,无法应对较为复杂的图像环境。图像复原是基于图像退化过程的先验知识重建退化的图像,以最大限度恢复图像原貌的过程。基于先验知识或数据假设的图像复原技术,如暗通道先验^[7]、颜色衰减先验^[8]、雾霾线先验^[9]等去雾算法,在一定程度上解决了难以从场景中提取光照强度、散射系数等问题,使得处理后的图像对比度、细节等信息得到了改善。但此类方法在低对比度的图像中处理效果较差,如针对天空、水面此类易产生块状效应的图像,其去雾效果较差。

近年来,随着计算机视觉的快速发展,深度学习技术被广泛应用在图像去雾中^[10]。基于深度学习的图像去雾算法本质上也是图像复原,利用卷积神经网络学习模糊图像的深度特征,寻找模糊图像与清晰图像间的映射关系,从而实现图像去雾。利用神经网络进行图像去雾主要从两种角度出发:1)利用CNN算法估计大气散射模型的参数,再根据大气散射模型来恢复无雾图像,如DehazeNet^[11]、MSCNN^[12]算法;2)直接利用CNN算法根据模糊图像生成无雾的清晰图像,如基于生成对抗网络(generative adversarial network, GAN)的HardGAN算法^[13]和基于Cycle-GAN网络^[14]的Cycle-Dehaze算法^[15]。GAN网络可以自动学习原始样本的数据分布,并通过噪声生成模拟样本,泛化性能较好,广泛应用于图像去雾领域。然而,此类方法在解决水域去雾问题时主要

面临两类困难。首先,GAN网络需要成对的含有雾的图像以及其对应的不含雾的图像,然而真实水域下,难以获取成对的有雾及对应的无雾非法采砂船只图像;其次,大多数GAN网络缺少对图像局部特征的关注,且损失函数的计算难以恢复图像的复杂纹理特征,无法对水域图像达到令人满意的去雾效果。

因此,针对此问题,本研究提出改进的生成对抗网络。首先,利用两组生成器与判别器分别学习有雾图像与无雾图像特征,解决网络训练缺乏成对的非法采砂船只图像的问题;其次,通过对生成器增加特征注意力模块提取雾天环境下船只的复杂纹理特征,对判别器增加谱归一化模块解决网络训练过程中训练梯度消失的问题;最后,在损失函数中引入循环一致性损失保证生成图像与原始图像间的结构一致性,从而确保算法对水域图像的良好去雾效果。

针对非法采砂船只辨识,诸多研究采用基于深度学习的目标检测算法。目前,目标检测算法主要有两个分支。1)两阶段目标检测算法,首先对目标进行定位,再对区域进行分类,如RCNN^[16]、Fast RCNN^[17]、Mask RCNN^[18-19]等算法。此类算法检测精度高,但速度较慢,无法满足实时检测的要求。2)一阶段目标检测算法,直接生成目标位置坐标与类别概率,具有较高的检测效率,代表算法为YOLO^[20]系列算法,可以很好地平衡检测精度与速度^[21-23]。

本研究针对雾天非法采砂船只识别困难、识别精度低等难题,提出一种基于深度学习的雾天非法采砂船只辨识方法。在图像预处理阶段,采用改进的生成对抗网络的图像去雾算法,挖掘有雾图像特征,降低图像噪声并实现图像去雾。针对去雾后的图像,利用融合注意力机制的YOLOv8算法对非法采砂船只进行精确识别、定位,实现对长江流域非法采砂船只的实时、精准监测,最大限度挤压违法犯罪空间,遏制非法采砂行为。

1 算法总体框架

针对雾天严重影响非法采砂船只实时监测性能的问题,本研究提出一种基于深度学习的雾天非法采砂船只辨识方法,包含两个模块。

1)图像去雾网络。提出一种基于改进GAN的图像去雾网络,多尺度捕捉水域有雾图像特征,对图像进行去雾处理,在图像预处理阶段有效降低现场图片噪声,生成清晰水域图像。

2)非法采砂船只辨识网络。将融合注意力机制的YOLOv8算法与提出的图像去雾网络相结合,对经过去雾处理的清晰水域图像进行检测,提高雾气干扰下非法采砂船只的识别精度。

算法总体框架如图 1 所示。首先,原始图像通过所提出的去雾网络进行处理,生成新的不存在雾气干扰的清晰水域图像。接下来,将处理后的图像数据输入到非法采砂船只辨识网络进行检测。最终,输出船只的检测结果,实现雾天非法采砂船只的实时、精准辨识。

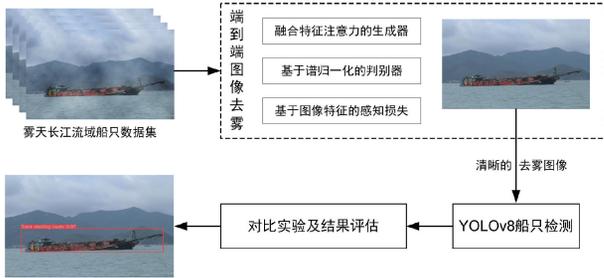


图 1 算法流程

Fig. 1 Flow chart of algorithm

2 雾天非法采砂船只辨识算法设计

2.1 改进的图像去雾算法

为解决雾天采集图像不清晰的问题,本研究对 GAN 网络进行 3 方面的改进,以提高网络对于模糊纹理特征的提取能力。首先,提出融合特征注意力机制的生成器,将特征注意力机制集成到生成器,以便网络能够提取雾天背景下船只的纹理特征。其次,在判别器中加入谱归一化操作,解决网络训练中梯度消失的问题,增强网络训练稳定性。最后,在预训练的 VGG19 网络的基础上,提出感知循环一致性损失函数,提高网络区分图像有雾与无雾特征的能力。

本研究提出的图像去雾网络包括两个自编码器生成器 G 和 F ,分别将有雾图像转换为无雾图像,以及将无雾图像转换为有雾图像。两个判别器 D_a 和 D_b ,分别用于判别由生成器 F 生成的有雾图像是真实图像还是生成图像,以及由生成器 G 生成的无雾图像是否是真实图像。生成器与判别器间的交互过程如图 2 所示。

首先,将有雾图像输入生成器 G ,由 G 生成相应的无雾图像,该无雾图像传入判别器 D_b ,以训练 D_b 区分真假无雾图像。随后,将无雾图像传入生成器 F 生成有雾图像,并利用 VGG19 网络提取的真实图像与生成图像间特征计算循环损失与感知损失,进一步从纹理细节与结构特征方面提高生成器的生成质量,保持循环一致性。此外,循环以相反的方向进行,训练 D_a 区分真假有雾图像,并保持原始无雾图像与生成图像间的循环一致性。

1) 融合特征注意力机制的生成器

为了在去雾的同时保留图像纹理细节,对图像去雾网络的生成器进行结构改进,如图 3 所示。该生成器由

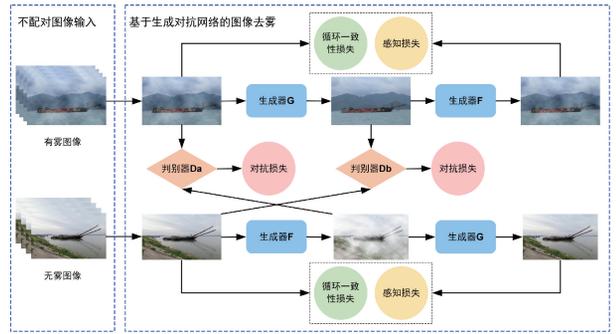


图 2 图像去雾网络框架

Fig. 2 Framework of image dehaze network

编码器、特征注意力模块、解码器组成。编码器首先使用的是卷积核 7×7 ,步长为 1 的卷积,在卷积之后引入 LeakyReLU 的修正线性单元以防止梯度消失。随后,使用卷积核 3×3 ,步长为 2 的卷积,并加入平均池化层,在保证信息丢失最小化的同时实现有效的下采样,保护图像的整体特征。之后再引入 LeakyReLU,后接卷积核 3×3 ,步长为 2 的卷积。整个编码器确保了网络捕捉复杂图像特征的能力。解码器负责将变换后的特征进行 4 倍上采样,从而重构出清晰的无雾图像。

由于图像中雾气分布不均会导致船只信息扭曲,在编码器与解码器之间引入来自 FFA-net 模型的特征注意力 (feature attention, FA) 模块,包含通道注意力和像素注意力子模块。前者通过对不同通道特征信息进行加权处理,解决雾气导致船只信息扭曲问题,后者使网络更关注雾气集中区域,解决雾气分布不均的问题。

通道注意力子模块主要关注不同频道特征的不同加权信息。首先使用全局平均池化将通道维度的全局空间信息封装成一个通道描述符,如式(1)所示。

$$f_c = H_p(F_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_c(i, j) \quad (1)$$

式中: $X_c(i, j)$ 表示第 c 通道在 (i, j) 位置的值; H_p 为全局池化函数。特征图大小从 $C \times H \times W$ 到 $C \times 1 \times 1$ 。

随后,通道特征通过两个卷积层处理,并经过 ReLU 激活层后连接 Sigmoid 函数。通道注意力计算公式如式(2)所示。

$$CA = \sigma(\text{Conv}(\delta(\text{Conv}(f_c)))) \quad (2)$$

式中: σ 是 Sigmoid 函数; δ 是 ReLU 函数。

将输入 F_c 和通道注意力权重进行逐元素相乘,则通道注意力子模块输出为 F_c^* 计算如式(3)所示。

$$F_c^* = CA \otimes F_c \quad (3)$$

像素注意力子模块将通道注意力子模块输出送入两个卷积层,并利用 ReLU 和 Sigmoid 激活函数,将特征大小从 $C \times H \times W$ 到 $1 \times H \times W$ 。

像素注意力计算如式(4)所示。

$$PA = \sigma(\text{Conv}(\delta(\text{Conv}(F_c^*))) \quad (4)$$

将输入 F_c^* 和 PA 进行逐元素相乘,得到像素注意力

子模块的输出 \tilde{F} ,如式(5)所示。

$$\tilde{F} = F_c^* \otimes PA \quad (5)$$

2) 基于谱归一化的判别器

判别器主要从图像整体水平判断并预测该图像为真实图像还是生成图像。为了使判别器满足利普西茨连续性,在判别器中每一层卷积的参数矩阵进行谱归一化,以保证网络训练的稳定性,避免梯度消失^[24]。Lipschitz 约束条件如式(6)所示。

$$\|D(x) - D(y)\| \leq K \cdot \|x - y\| \quad (6)$$

式中: D 为判别器; x, y 为图像空间中任意两点; $\|*\|$ 为2范数; K 为Lipschitz 常数。Lipschitz 条件限定在某局部

区域,函数的变化值不会大于 Lipschitz 常数。对于神经网络,其函数满足 Lipschitz 条件,则对应的梯度变化则不会超过对应的 Lipschitz 常数,也就不会出现梯度爆炸。因此,对判别器的每一层神经网络的参数矩阵 W 做奇异值分解,使其满足 1-Lipschitz,并在权重的每一次更新时都除以 W 的最大奇异值,则可以保证网络训练的稳定性和易收敛性。

$$\overline{W}_{sn}(W) = \frac{W}{\sigma(W)} \quad (7)$$

整体判别器结构如图4所示。网络由4个卷积核 4×4 ,步长为2的卷积进行下采样,并对每个卷积层进行谱归一化操作,后接 LeakyReLU 激活函数。最后,经过全局平均池化层和谱归一化操作,后接 Sigmoid 激活函数,输出 0~1 的判别概率。

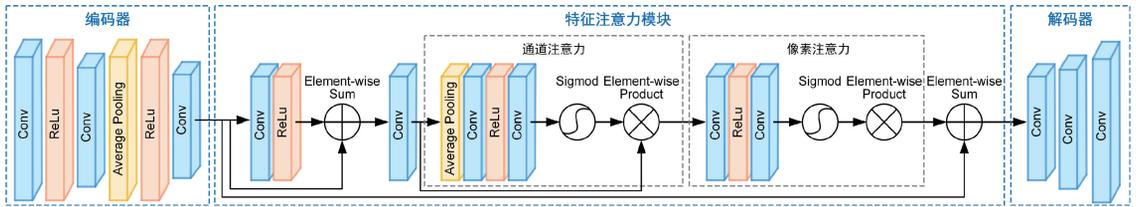


图3 生成器结构

Fig. 3 Structure of generator

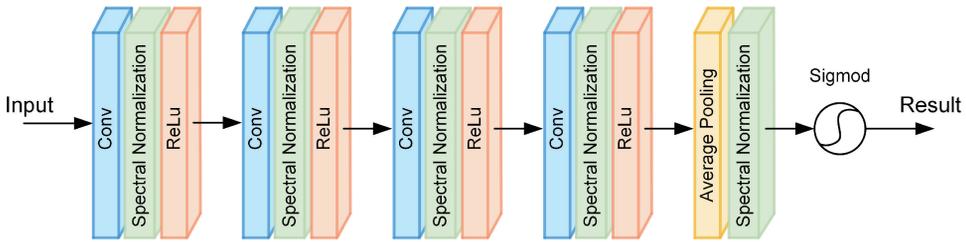


图4 判别器结构

Fig. 4 Structure of discriminator

3) 基于图像特征的感知损失

原始 GAN 利用两种损失函数,即对抗损失和循环一致性损失^[14]。对抗损失用于表示对抗学习的过程,使生成的图片在分布上更接近目标图片。对于映射函数 $G: A \rightarrow B$ 及判别器 D_b ,其目标函数为:

$$L_{adv}(G, D_b, A, B) = E_{b \sim p(b)} [\log D_b(b)] + E_{a \sim p(a)} [1 - \log D_b(G(a))] \quad (8)$$

式中: A 和 B 分别代表有雾和无雾图像数据集; a 和 b 为集合 A, B 中的图像; $a \sim p(a), b \sim p(b)$ 表示数据分布; E 为损失函数期望值。 G 试图生成与 B 数据集相似的无雾图像 $G(a)$,来最小化对抗损失。而 D_b 旨在区分生成

样本 $G(a)$ 和真实样本 b ,试图最大化对抗损失。网络学习、优化的过程即是寻找二者间的纳什均衡。同理,映射函数 $F: B \rightarrow A$,及判别器 D_a ,其目标函数如式(9)所示。

$$L_{adv}(F, D_a, B, A) = E_{a \sim p(a)} [\log D_a(a)] + E_{b \sim p(b)} [1 - \log D_a(F(b))] \quad (9)$$

循环一致性损失为不对的图像到图像的转换任务计算原始图像和循环图像间的 L1 范数,如式(10)所示。

$$L_{cycle}(G, F) = \|\hat{a} - a\|_{L1} \quad (10)$$

由于长沙流域有雾图像通常存在严重的模糊现象,仅依赖原始图像和循环重建图像间的损失值不足以恢复

全面的纹理信息。循环一致性感知损失通过利用从 VGG19 架构的第 2 和第 5 池化层中提取的高级和低级特征的组合来保持原始图像的结构完整性^[15], 如式 (11) 所示。

$$L_{per}(G, F) = \|\phi(a) - \phi(F(G(a)))\|_2 + \|\phi(b) - \phi(G(F(b)))\|_2 \quad (11)$$

在这个框架内, 循环一致性损失有助于确保高峰值信噪比 (peak signal-to-noise ratio, PSNR) 值。同时, 感知损失保持了图像的锐度和清晰度, 确保在重建后更细致、有效地保留图像细节和纹理。

总损失函数为:

$$L_{GAN} = L_{adv}(G, D_b, A, B) + L_{adv}(F, D_a, B, A) + \alpha L_{cycle}(G, F) + \beta L_{per}(G, F) \quad (12)$$

其中, α 与 β 为超参数, 控制循环一致性损失和循环损失的权重。

2.2 融合 CBAM 的 YOLOv8 船只辨识算法

YOLOv8 是基于 YOLOv5 算法的一个更新版本, 主要改进在于将算法的各个部分进行模块化, 以便更好地根据任务需要增添模块进行改进或优化。在网络结构上, YOLOv8 沿用 YOLOv5 的网络结构, 将网络分为特征提取主干网络 (backbone)、特征拼接增强网络 (neck)、检测头网络 (head) 3 部分。为了提高算法辨识精度, 在原始 YOLOv8 算法的 Neck 结构中加入注意力机制模块 (convolutional block attention module, CBAM), 通过学习获取每个特征空间的重要程度, 抑制不重要特征, 从而增强网络对图像重要特征的提取能力^[25], 融合 CBAM 机制的 YOLOv8 算法结构如图 5 所示。

3 实验结果分析

3.1 实验环境及数据集构建

实验采用 Linux 操作系统, 版本为 Ubuntu 20.04, 深度学习框架为 PyTorch2.0.0, CPU 为 15 * Xeon (R) Platinum 8474C, 内存为 80 G, GPU 为 NVIDIA GeForce RTX 4090D, 显存为 24 G。训练图像经过随机裁剪至 256 dpi×256 dpi 以完成数据增强, 并使用 Adam 优化器优化网络参数。

训练的超参数设置如下: epoch 设置为 200, batch size 为 4, 初始学习率为 0.0002, 在 100 个 epoch 之后开始线性衰减。平滑 L1 损失权重初始值设置为 0.5, 感知循环损失权重初始值为 0.001。

利用某局视频监控及网上公开获取的非法采砂船只与普通船只图像, 共 5 280 张图片构建数据集, 其中 3 884 张清晰图像, 真实有雾图像及利用模糊图像生成方法生成的有雾图像共 1 396 张, 共同用于网络训练。

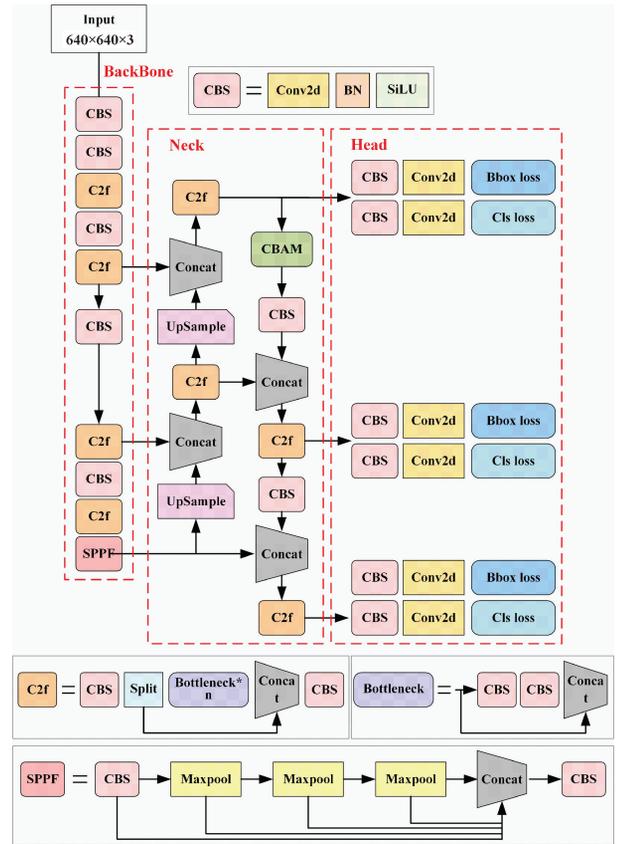


图 5 融合 CBAM 的 YOLOv8 网络结构

Fig. 5 Structure of YOLOv8 algorithm integrating CBAM

3.2 实验评估指标

1) 峰值信噪比

峰值信噪比通过均方差 (MSE) 进行定义, 比较有雾图像与去雾图像间差异来衡量去雾算法性能, 计算公式如式 (13) 所示。

$$PSNR = 10 \times \lg \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (13)$$

式中: n 为像素比特数; MSE 为有雾图像与去雾图像间差距。均方差定义为:

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [X(i, j) - Y(i, j)]^2 \quad (14)$$

式中: $X(i, j)$ 和 $Y(i, j)$ 分别表示有雾、无雾图像在 (i, j) 位置的像素值。PSNR 越大, 则算法去雾效果越好。

2) 结构相似性 (structural similarity index, SSIM)

结构相似性对有雾、无雾图像间的亮度、对比度、结构信息特征进行评估, 衡量两图像间的相似程度, 其计算结果在 $[0, 1]$, 越接近 1 则两图像越相似, 算法去雾效果越好。SSIM 计算公式为:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + a_1)(2\sigma_{xy} + a_2)}{(\mu_x^2 + \mu_y^2 + a_1)(\sigma_x^2 + \sigma_y^2 + a_2)} \quad (15)$$

式中: (x, y) 是两固定窗口; μ, σ 分别为均值和方差; σ_{xy} 为 x, y 间协方差; a_1, a_2 为非零常数。

3) 检测精度与速度评价指标

采用多类别平均精度均值 (mean average precision, mAP)、每秒帧数 (frames per second, FPS) 两个评价指标从精度和速度两方面评估目标检测模型性能。

3.3 去雾效果分析

为验证本研究提出的去雾算法的优越性, 选取两种目前常见的去雾算法 Cycle GAN^[14] 和 GC-GAN^[26] 进行比较。表 1 为 3 种算法在同一数据集上 PSNR 和 SSIM 评价指标结果。相较于 Cycle GAN 和 GC-GAN, 本研究提出的算法在 PSNR 上分别有 13.1% 和 3.6% 的提升, 在 SSIM 上分别有 56.1% 和 4.9% 的提升。从测试集中选取两张水域雾图, 分别利用 Cycle GAN、GC-GAN 和本研究提出的去雾算法进行恢复, 其视觉效果对比如图 6 所示, 可以看出, 本研究提出的算法可以更好地保留船只特征细节, 去雾效果更为彻底。

表 1 各算法去雾效果评价指标

Table 1 Results of image dehazing algorithms

算法	PSNR	SSIM
CycleGAN	28.17	0.41
GCGAN	30.76	0.61
本文	31.86	0.64

3.4 辨识效果分析

为了验证注意力机制对网络提取特征能力的增强效果, 本研究在保持实验条件、参数设置一致的前提下进行消融实验。

实验分为如下 4 组: 组别 1 为原始 YOLOv8 模型检测采集到的原始水域雾气图像; 组别 2 表示融合 CBAM 的 YOLOv8 算法检测原始图像; 组别 3 表示原始 YOLOv8 模型检测经过本研究提出的去雾算法处理后的清晰水域图片; 组别 4 表示融合 CBAM 的 YOLOv8 算法检测清晰水域图片。检测目标为普通船只与非法采砂船只两类。实验以非法采砂船只的检测精确率 (precision)、召回率 (recall)、模型的平均精度 $mAP@0.5:0.95$, 以及各模型的帧率为评价指标, 结果如表 2 所示。

YOLOv8 算法在原始水域图像上对非法采砂船只及普通船只的检测精度仅为 78.3%, 经过本研究提出的去雾算法对图像进行恢复后, 检测精度达到 85.3%。利用融合 CBAM 机制的 YOLOv8 算法检测过去雾处理的清晰图片时, 检测精度提升至 89.6%。部分检测结果如图 7 (组别 1 与 4 的对比图)。

图 7 (a) 与 (b) 相比, 经过图像去雾后 YOLOv8 算法



图 6 各算法去雾效果

Fig. 6 Results of image dehazing algorithms

对非法采砂船只的辨识置信度有所提高; 图 7 (c) 中, 算法无法识别图像右上角的船只, 图 7 (d) 中去雾后算法成功辨识出右上角目标较小的船只; 图 7 (e) 中, 算法识别出现错误, 图 7 (f) 中, 去雾后算法成功辨识出两艘非法采砂船只。因此, 经过本研究提出的算法, 目标检测精度得到显著提升, 平均精度均值由 78.3% 提升至 89.6%, 检测速度达 36 FPS。

表 2 消融实验结果

Table 2 Results of ablation experiments

指标	Precision	Recall	mAP@0.5:	
			0.95	帧率/ FPS
YOLOv8 原始图像	0.826	0.870	0.783	64
YOLOv8+CBAM 原始图像	0.899	0.931	0.829	57
YOLOv8 清晰图像	0.897	0.929	0.853	38
YOLOv8+CBAM 清晰图像	0.902	0.955	0.896	36

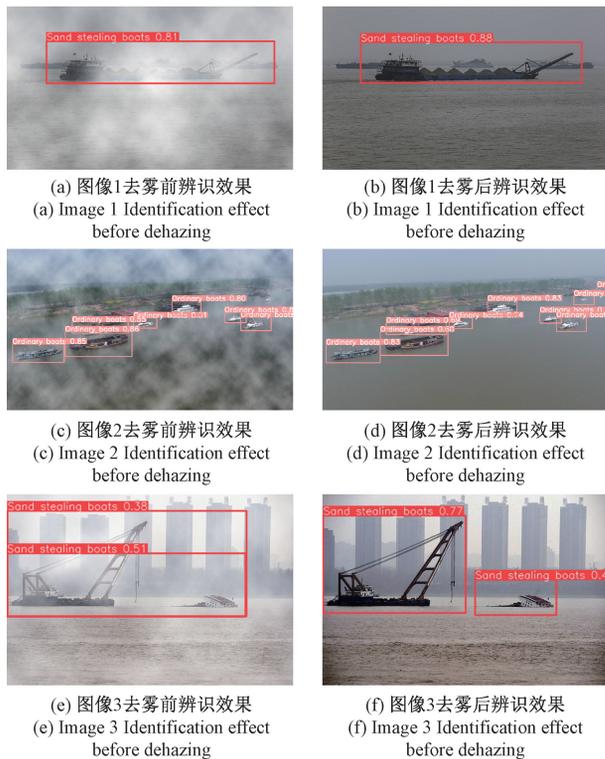


图 7 非法采砂船只检测结果

Fig. 7 Detection results of illegal sand mining vessels

4 结 论

本文针对雾天环境下非法采砂船只识别困难、辨识精度低的关键问题,提出一种基于深度学习的雾天非法采砂船只辨识方法。针对雾天环境下图像雾气干扰问题,提出了一种改进的生成对抗网络,利用两组生成器与判别器分别学习有雾图像与无雾图像特征,解决缺乏对应的非法采砂船只图像的问题,并通过对生成器增加特征注意力模块有效提取雾天环境下船只的复杂纹理特征;对判别器增加谱归一化模块解决网络训练梯度消失的问题;在损失函数中引入循环一致性损失保证生成图像与原始图像间的结构一致性,从而确保算法对水域图像的良好去雾效果。该算法与 Cycle GAN 和 GC-GAN 算法相比具有更好的去雾效果,去雾后的图像中目标船只特征表达更明显,PSNR 与 SSIM 为 31.86 和 0.64,分别提升了 3.6%~13.1%、4.9%~56.1%。随后,利用融合注意力机制的 YOLOv8 网络辨识去雾图像中的船只。结果表明,去雾后的图像在多条船只同时存在的情况下,仍可高效辨识图像中的所有目标,辨识精确率达到 89.6%,检测速度达到 36 FPS,满足公安实战需求。在未来的研究中,可从扩充数据样本、多模态感知等方法入手,优化去雾模型网络,提升对非法采砂船只在不同环境状态下检测识别的准确率与实时性。

参考文献

- [1] 汤文华, 陈灿斌, 向舒华, 等. 基于深度学习的图像识别技术在非法采砂监管中的应用[J]. 中国农村水利水电, 2021(5): 108-112.
TANG W H, CHEN C B, XIANG SH H, et al. Application of image recognition technology based on deep learning in supervision of illegal sand mining[J]. China Rural Water and Hydropower, 2021(5): 108-112.
- [2] 蔺志刚, 刘瑾程, 尤林奇, 等. 基于目标检测的北方河道非法采砂智能监管应用研究[J]. 人民黄河, 2023, 45(1): 135-139.
LIN ZH G, LIU J CH, YOU L Q, et al. Research on intelligent supervision and application of illegal sand mining in northern river channel based on target detection [J]. People's Yellow River, 2023, 45(1): 135-139.
- [3] 孙彧, 牛涛, 马振峰, 等. 最近 40 年中国雾日数和霾日数的气候变化特征[J]. 气候与环境研究, 2013, 18(3): 397-406.
SUN Y, NIU T, MA ZH F, et al. Characteristics of climate change with respect to fog days and haze days in china in the past 40 years [J]. Climatic and Environmental Research, 2013, 18(3): 397-406.
- [4] 蒋华伟, 杨震, 张鑫, 等. 图像去雾算法研究进展[J]. 吉林大学学报(工学版), 2021, 51(4): 1169-1181.
JIANG H W, YANG ZH, ZHANG X, et al. Research progress of image dehazing algorithms [J]. Journal of Jilin University (Engineering and Technology Edition), 2021, 51(4): 1169-1181.
- [5] KIM J Y, KIM L S, HWANG S H. An advanced contrast enhancement using partially overlapped sub-block histogram equalization[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(4): 475-484.
- [6] 韩浩, 何赉泽, 杜旭, 等. 基于偏振信息图像增强的多目标检测[J]. 电子测量与仪器学报, 2023, 37(3): 29-37.
HAN H, HE Y Z, DU X, et al. Multi-object detection based on polarization information image enhancement [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(3): 29-37.
- [7] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 33(12): 2341-2353.
- [8] ZHU Q S, MAI J M, SHAO L. A fast single image haze removal algorithm using color attenuation prior [J]. IEEE Transactions on Image Processing, 2015, 24(11): 3522-3533.

- [9] BERMAN D, AVIDAN S. Non-local image dehazing[C]. Proceedings of the IEEE Conference on Computer Vision And Pattern Recognition, 2016: 1674-1682.
- [10] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [11] CAI B L, XU X M, JIA K, et al. Dehazenet: An end-to-end system for single image haze removal [J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.
- [12] REN W, LIU S, ZHANG H, et al. Single image dehazing via multi-scale convolutional neural networks [C]. Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 – 14, 2016, Proceedings, Part II 14. Springer International Publishing, 2016: 154-169.
- [13] DENG Q L, HUANG Z L, TSAI C C, et al. Hardgan: A haze-aware representation distillation gan for single image dehazing[C]. European conference on computer vision. Cham: Springer International Publishing, 2020: 722-738.
- [14] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2223-2232.
- [15] ENGIN D, GENÇ A, KEMAL EKENEL H. Cycle-dehaze: Enhanced cyclegan for single image dehazing [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2018: 825-833.
- [16] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [17] GIRSHICK R. Fast R-CNN[C]. 2015 IEEE International Conference on Computer Vision (ICCV), 2015.
- [18] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN [C]. Proceedings of the IEEE International Conference on Computer Vision. 2017: 2961-2969.
- [19] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector [C]. Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.
- [20] REDMON J. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- [21] 彭道刚, 潘俊臻, 王丹豪, 等. 基于改进 YOLOv5 的电厂管道油液泄漏检测 [J]. 电子测量与仪器学报, 2022, 36(12): 200-209.
- PENG D G, PAN J ZH, WAMG D H, et. al. Oil leakage detection of power plant pipeline based on improved YOLOv5 [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36 (12): 200-209.
- [22] 闫钧华, 张琨, 施天俊, 等. 融合多层次特征的遥感图像地面弱小目标检测 [J]. 仪器仪表学报, 2022, 43(3): 221-229.
- YAN J H, ZHANG K, SHI T J, et. al. Multi-level feature fusion based dim small ground target detection in remote sensing images[J]. Chinese Journal of Scientific Instrument, 2022, 43(3): 221-229.
- [23] 高嘉, 刘涛, 王显峰, 等. TWRD-Net:一种用于曳引钢丝绳绳缺陷的实时检测网络算法 [J]. 仪器仪表学报, 2023, 44(6): 223-235.
- GAO J, LIU T, WANG X F, et. al. TWRD-Net: A real-time detection network algorithm for traction wire rope defects [J]. Chinese Journal of Scientific Instrument, 2023, 44(6): 223-235.
- [24] ZHOU Z, LIANG J, SONG Y, et al. Lipschitz generative adversarial nets[C]. International Conference on Machine Learning. PMLR, 2019: 7584-7593.
- [25] 吴兴辉, 何赟泽, 周辉, 等. 改进 YOLOv7 算法下的监控水域环境人员识别研究 [J]. 电子测量与仪器学报, 2023, 37(5): 20-27.
- WU X H, HE Y Z, ZHOU H et. al. Research on the personnel recognition in monitored water area based on improved YOLOv7 algorithm [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37 (5): 20-27.
- [26] FU H, GONG M M, WANG CH H, et al. Geometry-consistent generative adversarial networks for one-sided unsupervised domain mapping [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 2427-2436.

作者简介



张慧(通信作者), 2023 年于北京交通大学获得博士学位, 现为南京警察学院信息技术学院讲师, 主要研究方向为图像处理、深度学习、公共安全。

E-mail: huizhang_z@126.com

Zhang Hui (Corresponding author) received her Ph. D. degree in 2023 from Beijing Jiaotong University. Now she is the lecturer of the School of Information Technology, Nanjing Police University. Her main research interests include image processing, deep learning and public safety.



王新猛, 2008 年于南京师范大学获得硕士学位, 现为南京警察学院信息技术学院院长、教授, 主要研究方向为公安技术。

E-mail: wxmjr@sina.com

Wang Xinmeng received his M. Sc. degree in 2008 from Nanjing Normal University. Now he is the Dean and professor of the School of Information Technology, Nanjing Police University. His main research interest includes public security technology.



程占军, 2017 年于南京农业大学获博士学位, 现为南京警察学院警犬研究所繁育研究室主任、副研究员, 主要研究方向为公安技术。

E-mail: Nmgchengzhanjun@163.com

Cheng Zhanjun received his Ph. D.

degree in 2017 from Nanjing Agricultural University. Now he is the Head and associate research fellow of the Canine Breeding Research Department, Police Dog Technology Institute, Nanjing Police University. His main research interest includes public security technology.



杨一涛, 2005 年于华中科技大学获得硕士学位, 现为南京警察学院信息技术学院教授, 主要研究方向为网络安全。

E-mail: youngyt@gmail.com

Yang Yitao received his M. Sc. degree in 2005 from Huazhong University of Science and Technology. Now he is the professor of the School of Information Technology, Nanjing Police University. His main research interest includes cyber security.