

夜间动物图像自监督学习增强与检测方法

王驰 沈晨 黄庆 张国峰 卢汉 陈金波

Self-supervised learning enhancement and detection methods for nocturnal animal images

WANG Chi, SHEN Chen, HUANG Qing, ZHANG Guo-feng, LU Han, CHEN Jin-bo

引用本文:

王驰, 沈晨, 黄庆, 张国峰, 卢汉, 陈金波. 夜间动物图像自监督学习增强与检测方法[J]. *中国光学*, 2024, 17(5): 1-11. doi: 10.37188/CO.2024-0011

WANG Chi, SHEN Chen, HUANG Qing, ZHANG Guo-feng, LU Han, CHEN Jin-bo. Self-supervised learning enhancement and detection methods for nocturnal animal images[J]. *Chinese Optics*, 2024, 17(5): 1-11. doi: 10.37188/CO.2024-0011

在线阅读 View online: <https://doi.org/10.37188/CO.2024-0011>

您可能感兴趣的其他文章

Articles you may be interested in

固态微光实时偏振成像集成技术

Real-time polarization imaging integrated technology for solid-state low-light imaging

中国光学 (中英文). 2021, 14(3): 578 <https://doi.org/10.37188/CO.2020-0086>

结合光源分割和线性图像深度估计的夜间图像去雾

Nighttime image dehazing with a new light segmentation method and a linear image depth estimation model

中国光学 (中英文). 2022, 15(1): 34 <https://doi.org/10.37188/CO.2021-0114>

结合空洞卷积和迁移学习改进YOLOv4的X光安检危险品检测

Improved YOLOv4 for dangerous goods detection in X-ray inspection combined with atrous convolution and transfer learning

中国光学 (中英文). 2021, 14(6): 1417 <https://doi.org/10.37188/CO.2021-0078>

利用低信噪比小样本太赫兹光谱实现心肌淀粉样变检测

Detection of myocardial amyloidosis by a small number of terahertz spectra with low signal-to-noise ratio

中国光学 (中英文). 2022, 15(3): 443 <https://doi.org/10.37188/CO.2021-0223>

基于偏振差分图像的海天线检测方法

Sea-sky-line detection method based on polarization difference images

中国光学 (中英文). 2023, 16(3): 596 <https://doi.org/10.37188/CO.2022-0181>

基于机器学习的过焦扫描显微测量方法研究

Through-focus scanning optical microscopy measurement based on machine learning

中国光学 (中英文). 2022, 15(4): 703 <https://doi.org/10.37188/CO.2022-0009>

文章编号 2097-1842(2024)05-0001-11

夜间动物图像自监督学习增强与检测方法

王 驰¹, 沈 晨¹, 黄 庆², 张国峰¹, 卢 汉¹, 陈金波^{1*}

(1. 上海大学 机电工程与自动化学院, 上海 200444;

2. 中国航空工业集团公司洛阳电光设备研究所, 洛阳 471023)

摘要:为了解决动物夜间实时监测所面临的图像曝光度低、对比度低、特征提取困难等问题,通过研究轻量化自监督深度神经网络 Zero-Denoise 和改进型 YOLOv8 模型,来进行夜间动物目标的图像增强与精准识别。首先,通过轻量化的 PDCE-Net 进行第一阶段快速增强。提出了一个新的光照损失函数,并利用参数可调的 Gamma 校正原图与快速增强图,在基于 Retinex 原理和最大熵理论的 PRED-Net 中进行第二阶段的重增强。然后,改进 YOLOv8 模型,并对重增强后的图像进行目标识别。最后,在 LOL(Low-Light dataset)数据集与自建动物数据集进行实验分析,验证 Zero-Denoise 网络和改进型 YOLOv8 模型对于夜间动物目标监测的改善效果。试验结果显示,Zero-Denoise 网络在 LOL 数据集上的 PSNR、SSIM 与 MAE 指标达到 28.53、0.76、26.15,结合改进型 YOLOv8 在自建动物数据集上比 YOLOv8 基线模型的 mAP 值提升 7.1%。使用 Zero-Denoise 和改进型 YOLOv8 能获得良好的夜间动物目标图像。结果表明所提方法可用于进一步研究夜间动物目标精确监测。

关键词:夜间动物监测;低光增强;自监督学习;Retinex;低光去噪

中图分类号:TP394.1;TH691.9

文献标志码:A

doi:10.37188/CO.2024-0011

Self-supervised learning enhancement and detection methods for nocturnal animal images

WANG Chi¹, SHEN Chen¹, HUANG Qing², ZHANG Guo-feng¹, LU Han¹, CHEN Jin-bo^{1*}

(1. School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China;

2. Aviation Industry Corporation of China Luoyang Electro-Optical

Equipment Research Institute, Luoyang 471023, China)

* Corresponding author, E-mail: jbchen@shu.edu.cn

Abstract: In order to solve the problems of low image exposure, low contrast and difficulty of feature extraction in real-time animal monitoring at night, we proposed a lightweight self-supervised deep neural network Zero-Denoise and an improved YOLOv8 model for image enhancement and accurate recognition of nocturnal animal targets. The first stage of rapid enhancement was performed by lightweight PDCE-Net. A new lighting loss function was proposed, and the second stage of re-enhancement was carried out in PRED-Net based on the Retinex principle and the maximum entropy theory, using the original image and fast enhance-

收稿日期:xxxx-xx-xx; 修订日期:xxxx-xx-xx

基金项目:国家自然科学基金项目(No. 62175144);北京市航空智能遥感装备工程技术研究中心开放基金课题(No. AIRSE20233)

Supported by National Natural Science Foundation of China (No.62175144); Beijing Engineering Research Center of Aerial Intelligent Remote Sensing Equipments Open Fund (No. AIRSE20233)

ment image corrected by the parameter adjustable Gamma. Then, the YOLOv8 model was improved to recognize the re-enhanced image. Finally, experimental analysis was conducted on the LOL dataset and the self-built animal dataset to verify the improvement of the Zero-Denoise network and YOLOv8 model for nocturnal animal target monitoring. The experimental results show that the PSNR, SSIM, and MAE indicators of the Zero-Denoise network on the LOL dataset reached 28.53, 0.76, and 26.15, respectively. Combined with the improved YOLOv8, the mAP value of the baseline model on the self-built animal dataset increased by 7.1% compared to YOLOv8. Zero-Denoise and improved YOLOv8 can achieve good quality images of nocturnal animal targets, which can be helpful in further study of accurate methods of monitoring these targets.

Key words: nocturnal animal detection; low-light enhancement; self-supervised learning; Retinex; low-light denoising

1 引言

夜间低光照度下动物目标的被动式快速监测一直是领域课题和关键难题重要研究^[1]。夜间动物图像监测虽然可以通过在相机上加探照功能,或者使用主动红外相机等实现,但这会在一定程度上对动物造成惊扰而破坏动物行为^[2];同时受到背景光、低光干扰和不均匀照明等影响^[3],导致对动物观测的图像质量大幅度下降。这对于夜间动物观测和其行为轨迹追踪造成了困难^[4]。

近年来,随着深度学习的蓬勃发展,针对夜间条件下拍摄动物图像总是存在低对比度、低亮度和特征提取困难等问题^[5],相关人员提出许多方法,其中,低光增强方面的研究主要有卷积神经网络(CNN)^[6]方法和生成对抗网络(GAN)^[7]方法。Li 等人^[8]通过 LightenNet 预测光照成分并基于 Retinex 模型获得增强图像。Ding 等人^[9]构建了一个端到端的 FCN(全卷积神经网络),并在自建的 SID 数据集上实现了有效的噪声抑制和色彩转换。上述基于有监督训练的解决方案依赖于数量有限的配对型数据,其模型泛化能力较差。针对此问题,Guo 等人^[10]提出了 Zero-DCE 网络,通过零参考曲线估计策略,将低光增强转化为曲线拟合问题。Jiang^[11]等人提出了 Enlighten-GAN 可在无配对图像情况下的训练模型。Fu 等人^[12]提出的 LE-GAN 网络,设计了照明感知注意力模块,以解决噪声和颜色偏差问题,并研究了一种恒等不变损失来解决过度曝光问题,自适应增强低光图像。Wang 等人^[13]提出了利用混合注意力引导生成对抗网络,并用以降噪与图像增强。此外,还有利用物理模型与深度学习相结合的方法,如

Monakhova 等人^[14]在极低光照(0.6m 流明)环境下,通过对相机噪声模型进行建模以克服极低光照环境下视频成像的限制。

上述研究虽均取得一定成效,但为获得高质量图像牺牲了运算速度,同时在进行亮度和对比度增强的情况下会产生颜色失衡和噪点放大等问题,而且低光数据集的缺失使其获取成对数据集极为困难。针对上述问题,本文提出一种自监督低光图像增强模型,在缺少配对数据集的条件下,通过自监督学习以实现去噪和图像质量增强,研究轻量化零参考深度曲线估计增强网络(PDCE-Net)和重增强去噪网络(PRED-Net)模型,或称之为 Zero-Denoise 网络模型。此外,根据夜间图像特征难以提取的情况对 YOLOv8 进行有针对性地改进,以进一步提升夜间动物目标的识别精度。

2 夜间图像增强的自监督学习模型

为了解决动物夜间实时监测中面临的低曝光度、低对比度及特征提取困难等问题,本文提出了一种两阶段轻量化自监督学习模型——Zero-Denoise。该模型能根据不同图像拍摄设备和野外环境进行自适应训练。第一阶段,通过构建部分卷积深度曲线估计网络(Partial Convolution Depth Curve Estimation Network, PDCE-Net),旨在提高图像亮度的同时避免颜色失真,输出阶段性增强图像。PDCE-Net 网络是基于 DCE^[10]网络进行改进的,其通过学习输入图像与其最佳拟合曲线参数图之间的映射进行增强,本文将深度可分离卷积替换为 Partial Convolution (PConv)^[15],以进一步提升运算速度。PConv 卷积原理如图 1(彩图见期刊电子版)所示。PConv 还可以更好地

利用硬件设备上的计算能力以减少计算冗余和内存访问数量,也可有效提取空间特征,Pconv 还可以利用特征图的冗余度进一步优化成本^[15]。

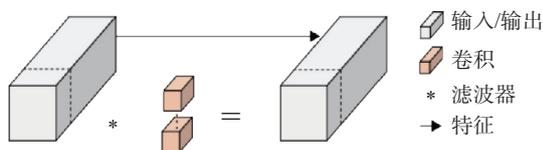


图1 部分卷积 PConv 原理图

Fig. 1 Convolution principle of PConv

本文在 Zero-DCE++ 中引入 PConv 算子以提升运算速度,同时将卷积算子后的 ReLU 激活函数替换为 LeakyRelu。因为 x 小于 0 时 ReLU 会将特征屏蔽,当屏蔽特征太多时,则会导致模型无

法学习到有效特征,本文以此来防止负梯度不被激活而存在神经元坏死现象。

Zero-Denoise 网络整体结构如图 2(彩图见期刊电子版)所示。第一阶段模型继承了原网络的零参考损失,通过简洁的卷积层来估计光增强曲线映射表达式,如式 (1) 所示^[10]。4 组精心设计的损失函数表达式如式 (2) 所示。其能够对给定图像进行精准动态调整,最终得到更加轻量化的 PDCE-Net,其整体结构如图 2 中 Zero-Denoise 网络整体结构中 PDCE-Net 部分所示。该网络通过光增强曲线 (LE-curve) 的 8 次迭代,输出具有较好光照信息、颜色和对比度图像,但可能仍含有一定的噪声。

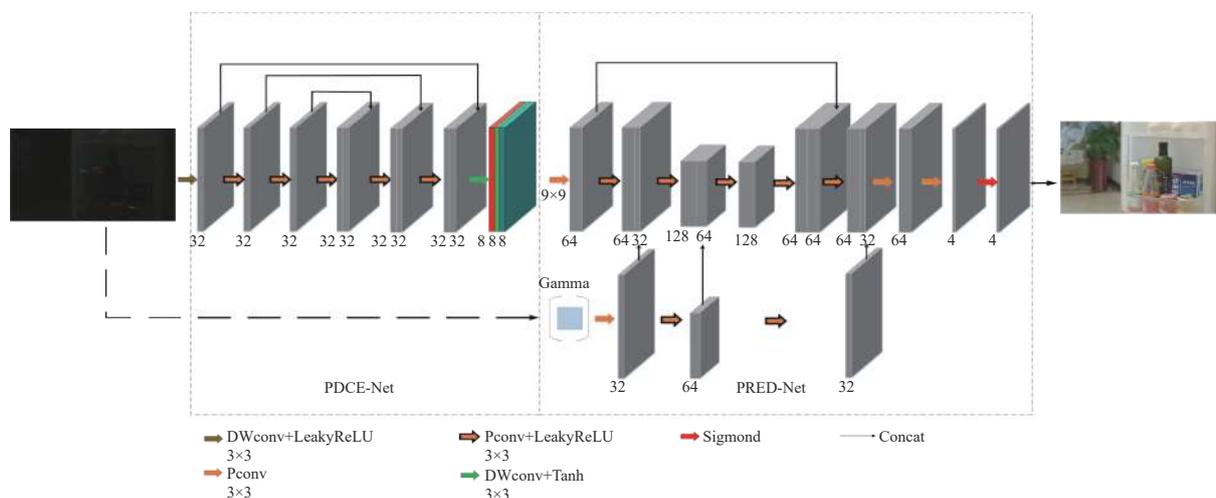


图2 Zero-Denoise 网络整体结构

Fig. 2 Overall structure of Zero-Denoise network

第二阶段网络模型利用重增强思想,将经可调参数的 Gamma 校正后的低光图像与第一阶段输出的图像进行比较,以减少噪声的影响。这一过程通过自监督网络 PRED-Net 实现,最终输出图像不仅亮度有所提升、颜色均衡。对比度有所增强,且噪声较低,与人类视觉感知高度一致。PRED-Net 网络结构在原 RED-Net^[17] 的基础上使用 Pconv 卷积算子进行轻量化,并对经 Gamma 校正后的图像的最大熵通道进行特征复用,其在原模型的基础上增加了 3 层特征提取,加强了对最大熵通道的特征复用,提出了新的光照损失函数。

Zero-Denoise 网络的损失函数分为两部分:亮度增强网络(PDCE-Net)损失函数 L_{PDCE} 与重增强网络(PRED-Net)损失函数 L_{PRED} 。

整个亮度增强网络(PDCE-Net)的图像映射表达式如下:

$$LE_n(x) = LE_{n-1}(x) + \mathcal{A}(x)LE_{n-1}(x)(1 - LE_{n-1}(x)). \quad (1)$$

PDCE-Net 网络损失函数由空间一致性损失 L_{spa} 、曝光控制损失 L_{exp} 、色彩恒常性损失 L_{col} 和光照平滑损失 L_{lva} 等 4 部分组成。PDCE-Net 的损失 L_{PDCE} 公式如下^[10]:

$$L_{PDCE} = L_{spa} + L_{exp} + W_{col}L_{col} + W_{lva}L_{lva}, \quad (2)$$

其中: W_{col} 和 W_{lva} 是权值和损失值,分别为 0.4 和 0.6。

L_{PRED} 损失函数是基于 Retinex 理论的神经网络损失函数^[16]。根据泊松噪声分布式 (3) 中的各项损失均被修改为去噪并保留更多细节的结果;依据最大熵理论,以增强 V 通道(enhanced V channel)为条件,对低照度图像进行重新增强,表示为^[17]:

$$L_{PRED} = \min I'_{rcon} + \lambda_1 I'_R + \lambda_2 I'_I, \quad (3)$$

其中: l'_{con} 为重构损失, l_r 为反射率损失, l_l 为光照损失, λ_1 和 λ_2 分别为其对应系数, 用来平衡反射率一致性和光照平滑度。

本文引入了一个参数可调的 Gamma 校正图作为重增强的基准图, 对此设计了一个新的光照损失函数, 并引入了 Gamma 校正系数, 新的光照损失函数公式如下:

$$l'_1 = \|\|N(\|\nabla I'\|) \cdot \exp(-\lambda_3 N(\|\nabla I'\|)) \cdot \exp(-\lambda_3 W_G \cdot N(\|\nabla R'\|) \cdot \mu^\gamma)\|_1, \quad (4)$$

其中: $\nabla R'$ 和 $\nabla I'$ 分别表示反射图与照度图的梯度算子, μ 表示一个保持非负数的小于 1 的数字, γ 表示 Gamma 校正的指数值, $N(x)$ 表示 x 的局部归一化; W_G 表示为权重, $G(x)$ 表示 x 上的平均滤波器。 W_G 的计算公式为:

$$W_G = N(|G(\nabla R')|) \quad (5)$$

综上所述, 本文首先进行了轻量化改进, 将普通卷积修改为 Pconv 部分卷积模型, 以提升运算

速度, 并在改进 Zero-DCE 快速得到运算结果的基础上, 利用重增强概念, 即以增强 V 通道为条件, 对低照度图像进行重新增强, 以泊松噪声分布作为真实噪声分布, 更加贴近低光照噪声的实际情况。此外, 将经参数可调的 Gamma 校正后的图像作为基准图, 设计了新的光照损失函数, 并改进了原来的提取网络以进行特征复用。这样可以进行有效的重增强和去噪。

3 改进 YOLOv8 目标检测算法

为了使 YOLOv8 能够专注于夜间动物图像中的检测敏感区域, 从而提取出更丰富、更鲁棒和更具判别力的特征, 在主干网络中引入 ShuffleAttention(简称 SA 注意力机制)^[18]。SA 注意力机制通过分组并行化的方式, 能够快速有效地实现通道和空间注意力。这样在保留模型效果的同时大幅降低计算成本。其结构如图 3(彩图见期刊电子版)所示。

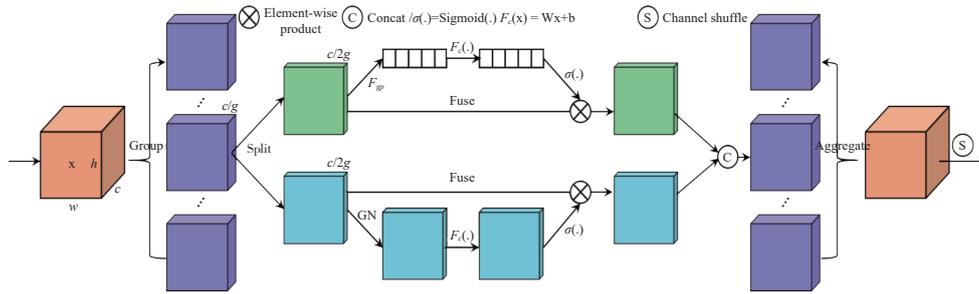


图 3 SA 注意力机制原理图

Fig. 3 Schematic diagram of Shuffle Attention mechanism

SA 模块首先将输入特征图 X 划分为 G 组子特征图: $X = \{X_1, X_2, \dots, X_G\}$, 每一个子特征图均会被分解为两条支路 X_{k1} 和 X_{k2} , 对它们分别进行通道注意力和空间注意力, 公式如下:

$$X'_{k1} = \sigma(F_c(s)) \cdot X_{k1} = \sigma(W_1 s + b_1) \cdot X_{k1}, \quad (6)$$

其中: W_1 为通道注意力机制中的权重矩阵, b_1 是权重向量; σ 表示通过 sigmoid 激活; s 是通道注意力支路中通过全局平均池化 GAP 后的均值信息, 通过空间维度 $H \times W$ 来均值化 X_{k1} 后的值。

$$s = F_{gp}(X_{k1}) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{k1}(i, j), \quad (7)$$

其中: H 和 W 代表 5×5 的特征图尺寸。

$$X'_{k2} = \sigma(W_2 \cdot GN(X_{k2}) + b_2) \cdot X_{k2}, \quad (8)$$

其中: GN 表示使用 Group Norm 来获得空间上的数据, W_2 为空间注意力机制中的权重矩阵, b_2 是权重向量。将两个支路合并, 聚合所有的子特征进行一个通道混洗, 使得跨组信息能够在组间传递。本文选择在骨干网络最后一个 C2f 后添加 SA 注意力机制。

YOLOv8 的颈部部分采用 PANet 网络结构。该网络结构只考虑了相邻层级间的同一尺度和同一层级间的相邻尺度的影响, 在特征融合过程中出现较多的信息丢失。夜间环境下的图像往往存在着特征不清晰、不明显的情况, 原 neck 部分相会丢失较多特征信息, 从而使目标识别困难。RepGFPN^[19] 是在 GFPN 的基础上提出的轻量化金字塔 neck, 展现出了更强的鉴别能力来充分交换高级语义信息和低级空间信息。RepG-

FPN 将 GFPN 与重参数化技术相结合, 显著增强了特征提取能力, 而且没有引入额外的巨大计算负担。此外, 利用多尺度特征融合发生在前一层和当前层的不同尺度特征中这一特性, $\log_2(n)$ 的跨层可以扩展到更深的网络并提供更有效的信息交换。RepGFPN 的结构如图 4(彩图见期刊电子版)所示, 本文将 RepGFPN 神经网络 neck 结构融入到 YOLOv8 网络中, 构建了一个高精度的特征融合网络。

由于在自建数据集中, 低质量样本的存在是

不可避免的, 样本可能会在几何度量(如距离、纵横比)方面表现出较大的差异。这种差异会导致模型过度关注于这些低质量样本, 从而使模型的泛化性能降低。为了提高模型的泛化能力, 在锚框和目标框重合度较高时, 需要适度降低几何度量的梯度惩罚。这样可以降低低质量样本的梯度惩罚, 使模型的目标更加聚焦于普通样本。Wise-IOU V3^[20] 利用离群值 β 来评估锚框的质量, 并通过动态非单调调频的方式, 有效抑制了低质量样本在整个样本集中的权重, 公式如下:

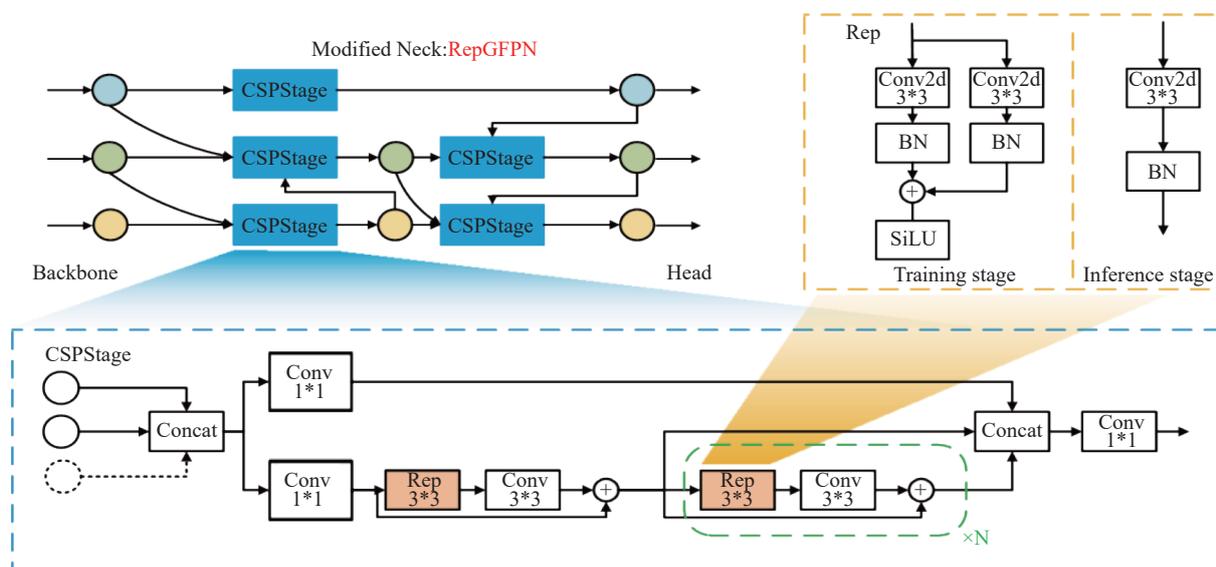


图 4 RepGFPN 整体结构

Fig. 4 Overall structure of RepGFPN

$$L_{WIoUV3} = rL_{WIoUV1}, r = \frac{\beta}{\delta\alpha^{\beta-\delta}}, \quad (9)$$

其中: α 和 δ 代表可调超参数, L_{WIoUV1} 是具有两层注意力机制的距离注意力公式。

至此, 本文根据夜间环境中动物特征难提取的问题改进了 YOLOv8 的骨干网络、颈部网络以及损失函数, 改进后的整体模型网络结构如图 5 所示。

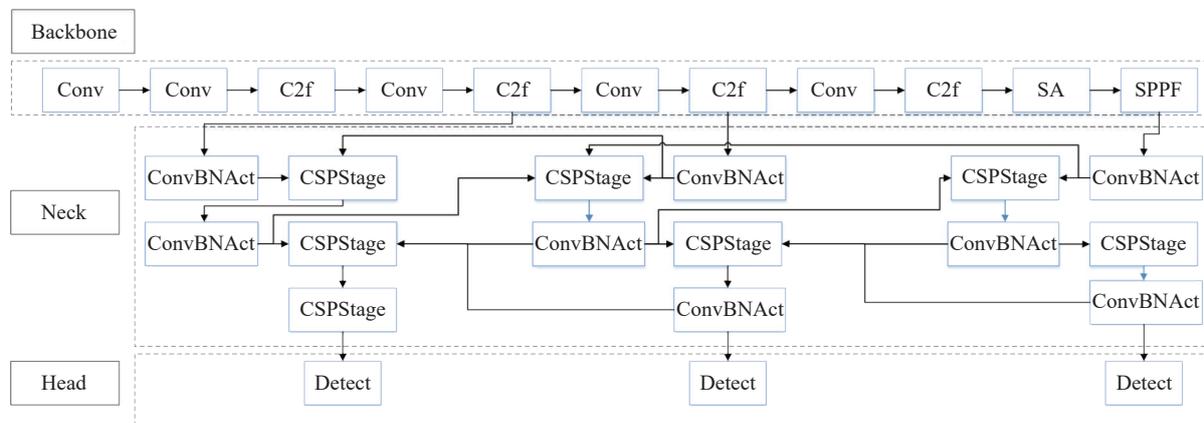


图 5 改进 YOLOv8 网络整体结构

Fig. 5 Overall structure of improved YOLOv8

综上所述,本文在骨干网络中引入了 SA 注意力机制,通过融合 RepGFPN 能够充分交换高级语义信息和低级空间信息,加强 YOLOv8 的特征提取能力,最后通过使用 Wise-IOU v3 损失函数降低低质量样本的不良影响,提升模型的整体性能。

4 实验结果与分析

4.1 实验条件及参数配置

为了验证 Zero-Denoise 算法的有效性,本文选取 LOL^[16](Low-Light dataset)数据集进行训练和测试,其包含 500 个低/正常光图像对。其中 485 个用于训练,图像大小为 400×600,该数据集

中通过改变曝光时间和 ISO 感光度来收集低光图像。

改进型 YOLOv8 算法能有效提取夜间环境下的动物特征,本文建立一个夜间动物低光照数据集,以验证 Zero-Denoise 算法的低光照增强效果。采用微光夜视相机拍摄建立数据集,相机型号为 PCO 4.2 edge,输出图像大小为 1024 pixel×1024 pixel,色彩为黑白,总数量为 3736 张,其中包含的实例为 13476 个,每张图中包含 1~10 个实例不等,训练集、验证集与测试集比例为 8:1:1。本文建立的动物低光照数据集如图 6 所示,包含 5 类在低光照环境下拍摄的动物,其中包括鸽子、老鼠、鸡、鸭子以及兔子。

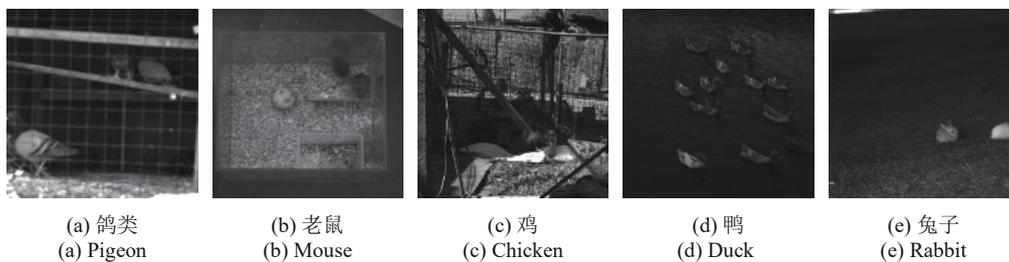


图 6 数据集各类动物种类

Fig. 6 Various animal species in the dataset

本文研究的模型使用 Pytorch 库在 Python 上实现,并在有 NVIDIA 2080ti GPU 的服务器上进行训练,批处理尺寸设置为 16。Zero-Denoise 网络模型实验设置如下:使用 Adam 优化器,学习率为 0.0002,训练轮次为 3000,在 LOL 数据集和自建动物数据集上进行测试实验。改进 YOLOv8 (Im-YOLOv8) 实验设置如下:使用 SGD 优化器,学习率为 0.01,训练轮次为 300,在自建动物数据集上进行测试实验。

4.2 基准评估

与目前先进的算法 Zero-DCE++^[10]、Retinexnet^[16]、URetinex Net^[21]、StableLLVE^[22]、SCI^[23]、MBLLEN^[24] 进行定性与定量对比。图 7(彩图见期刊电子版)从主观角度展现了各类低光照增强算法的优劣,其中左上角为局部放大图。图 7(h)是 LOL 数据集中的一幅低照度下未增强的图像,图 7(i)是通过长曝光拍摄的图像,其作为正确图像。可以看到几乎所有模型都对该图像产生了一定增强效果,但是由 Zero-DCE++、SCI、Retinexnet 这几种模型生成的图像视觉效果不佳,图

像整体出现噪点放大现象。相比较而言,本文所提的 Zero-Denoise 和 URetinex-Net、StableLLVE、MBLLEN 4 种模型无论是在亮度增强还是图像质量恢复方面,都要明显优于其他方法。但是 URetinex-Net、StableLLVE 使图像整体产生了较为明显的过曝,整体呈现雾蒙蒙的效果。

本文 Zero-Denois 增强的图像,从整体上看色彩更加饱满。这是因为本文所设计的网络能够有效去除噪点,同时减少图像中原有细节信息的损失,此外,通过最大熵理论方法有效保留了原图对比度和色彩。从图像整体质量、全局光照信息、局部细节以及对比度提升效果上来看,Zero-Denoise 对于低光图像的增强效果最佳。

接下来,将通过客观评价指标 PSNR(峰值信噪比)、SSIM(结构相似性)、MAE(平均绝对误差)来判断算法的性能。其中,PSNR、SSIM 指标数值越高说明图像质量越好,MAE 数值越低说明图像质量越好。不同算法评价指标结果如表 1 所示,评价最优的用下划线标明。

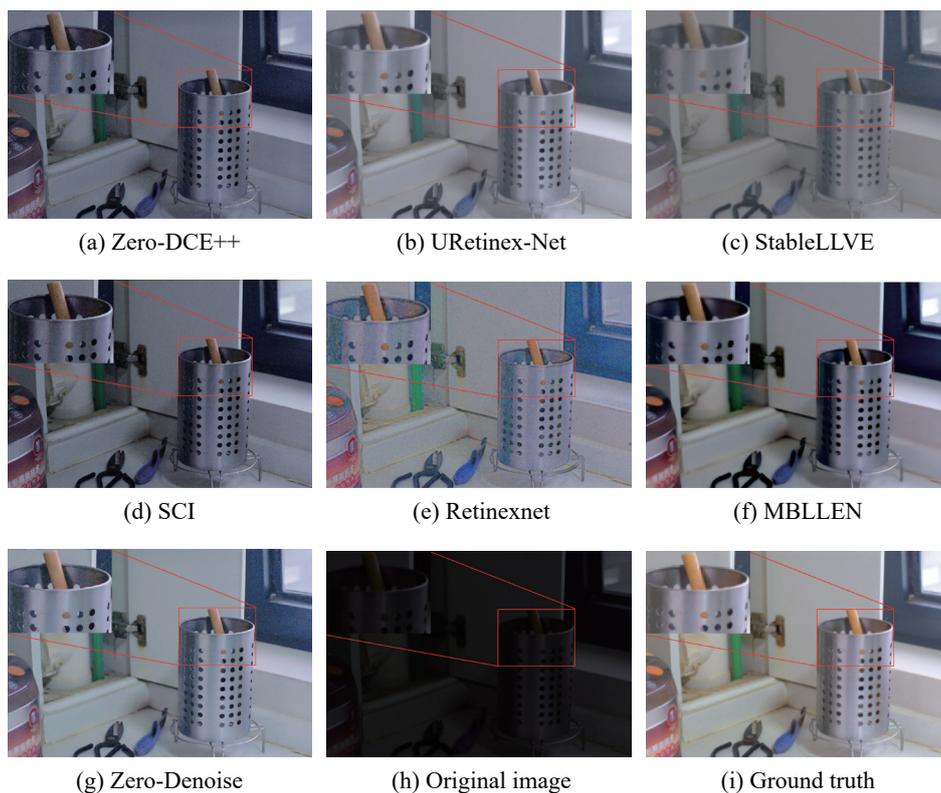


图 7 各类先进图像增强算法对比

Fig. 7 Performance comparison of various advanced image enhancement algorithms

表 1 不同算法评价结果

Tab. 1 Evaluation results of different algorithms

	算法	PSNR	SSIM	MAE	Runtime(s)
无监督 自监督	Zero-DCE ++	27.93	0.5725	44.9393	0.0008
	SCI	27.90	0.5254	48.7550	<u>0.0006</u>
	Zero-Denoise(ours)	<u>28.53</u>	<u>0.7665</u>	<u>26.1516</u>	0.0181
有监督	StableLLVE	27.92	0.7373	32.4789	0.5900
	URetinex-Net	28.45	<u>0.8332</u>	<u>21.1562</u>	0.0367
	Retinexnet	28.06	0.4250	32.0174	0.1200
	MBLEN	28.04	0.7247	31.2498	8.5633
	Zero-Denoise(ours)	<u>28.53</u>	0.7665	26.1516	<u>0.0181</u>

由表 1 中数据可知: 和有监督算法相比, 本文算法 PSNR 指标优于其他算法, SSIM 和 MAE 指标表现良好, 运行速度达到 0.0181s 优于其他算法, 仅为其他算法的 50%。说明 Zero-Denoise 模型轻量化水平有显著提升, 模型简单且复杂度低, 兼顾优秀的增强效果与实时性。同时, 虽然 URetinex-Net 在 LOL 数据集上的评价指标较高, 但作为有监督算法的 URetinex-Net 没有在与相应的配对数据集进行训练, 所以在不同环境下的目标

检测性能不一定好, 鲁棒性不强。而本文算法可以根据自监督进行训练, 在实际运用中能够根据现有数据集进行训练, 对目标检测算法的提升效果较为明显。对目标检测算法的精度提升结果见夜间动物数据集中的实验测试部分。

4.3 消融实验

为了验证本文提出的改进型 YOLOv8(Im-YOLO v8) 的各项改进措施对夜间动物图像目标检测算法的提升效果, 以 YOLOv8 为基准算

法,对相应的改进措施进行了一系列消融实验。为了能够有效直观地展示 YOLOv8 的改进效果,采用查准率(Precision, P)、召回率(Recall, R)、平均精度均值(Mean Average Precision, mAP)等参量进行定量评价。mAP@0.5 表示 IoU 取值为 0.5 时的总类别平均精度, mAP@0.5:0.95 表示不同 IoU 取值(从 0.5 到 0.95,步长 0.05)的平均精度, mAP 值越高说明算法目标识别能力更强。实验结果如表 2 所示。表 2 中, A 表示在主干网络中引入 SA 注意力机制; B 表示将 RepGFPN 结构应用于 YOLOv8n 的颈部模型; C 表示用

WIoUV3 作为回归损失函数; D 表示同时引入 SA 注意力机制和改进颈部网络为 RepGFPN; Ours 表示对 YOLOv8 添加本文所有的改进措施。可见,相较于基准算法,每一项改进对应的各项指标都有所提升,证明改进有效。本文方法查准率提升了 5.6%,召回率提升了 5.6%,平均精度均值 mAP@0.5 提升了 1.5%, mAP@0.5:0.95 提升了 2.0%,融合所有改进方法的各项指标最优。由此可知,本文的 Im-YOLOv8 能改善模型的特征提取能力,提升对夜间动物数据集的检测精度。

表 2 改进 YOLOv8 消融实验结果

Tab. 2 Experimental results of improved YOLOv8 ablation

算法	改进主干	改进颈部	改进损失	P	R	mAP@0.5	mAP@0.5:0.95
YOLOv8				72.3	70.8	80.7	49.7
A	√			73.3(+1.0%)	72.3(+1.5%)	81.2(+0.5%)	50.3(+0.6%)
B		√		75.0(+2.7%)	74.6(+3.8%)	81.4(+0.7%)	50.4(+0.7%)
C			√	73.3	72.3(+1.5%)	80.9(+0.2%)	50.1(+0.4%)
D	√	√		76.0(+3.7%)	74.9(+4.1%)	81.9(+1.2%)	51.3(+1.4%)
Ours	√	√	√	<u>77.9(+5.6%)</u>	75.2(+5.6%)	82.2(+1.5%)	51.7(+2.0%)

4.4 Zero-Denoise 对夜间动物数据集的提升效果验证

由于夜间可见度低,且目标移动速度较快,故只能采用低曝光来拍摄动物图像,但低曝光情况下的特征更不清晰,导致夜间动物检测难度更大。直接使用现有目标检测算法处理夜视相机采集的夜间动物图像时,会导致目标误检率、漏检率等均较高,并且难以依靠人眼判断检测结果是否可靠。为了验证本文提出的方法 Zero-Denoise 在动物目标检测任务中的有效性,将 Zero-

Denoise 与不需要配对图像的 SCI 算法在上述自建夜间动物数据集上进行训练并进行增强预处理。由于其他有监督算法无法直接训练所以使用原权重增强数据集。然后,将所有增强后的数据集与 im-YOLOv8 目标检测模型相结合并测试增强算法对目标检测算法性能的影响,以进一步验证本文算法的鲁棒性以及自监督的优越性。本文采用平均精度均值 (mean Average Precision, mAP) 作为检测指标,将交并比 (Intersection of Union, IoU) 设置为 0.5,实验结果如表 3 所示。

表 3 不同算法对改进 YOLOv8 的效果

Tab. 3 The effect of different algorithms on improved YOLOv8

算法	Rabbit	Bird	Chicken	Mouse	Duck	mAP@0.5
Im-YOLOv8	94.0	86.5	62.2	80.3	63.5	82.2
URetinex-Net+Im-YOLOv8	94.5	88.0	73.5	91.9	66.2	83.5(+1.3%)
StableLLVE+Im-YOLOv8	93.5	90.6	71.6	82.2	64.5	82.5(+0.3%)
SCI+Im-YOLOv8	94.7	91.1	72.0	93.2	69.5	84.7(+2.5%)
Retinexnet+Im-YOLOv8	89.2	76.6	60.9	68.5	58.7	78.5(-3.7%)
MBLLEN+Im-YOLOv8	94.2	90.8	72.5	89.0	65.5	83.1(+1.1%)
Zero-Denoise+Im-YOLOv8	<u>95.0</u>	<u>91.7</u>	<u>74.6</u>	<u>97.4</u>	<u>76.0</u>	<u>87.8(+5.6%)</u>

由表 3 可知, 在夜间动物数据集上, 除本文算法外的其他增强算法对目标检测的提升效果并不理想。例如 URetinex-Net 增强算法虽然 SSIM 和 MAE 指标最优, 但是对夜间动物数据集的平均精度 mAP 值的提升仅 1.3%; SCI 由于进行了无监督训练, mAP 值提升了 2.5%, 但依旧远小于本文算法的提升幅度。本文的 Zero-Denoise 算法由于对夜间动物数据集的图像进行了自监督训练

故鲁棒性良好, mAP 值提升 5.6%, 远超于其他图像增强模型。

为了更加直观地展现 Zero-Denoise 对夜间动物数据集的提升效果, 将检测精度提升示例进行了部分可视化, 结果如图 8(彩图见期刊电子版)所示, 结论如表 4 所示。可视化结果有效体现了 Zero-Denoise 的检测精度提升效果。

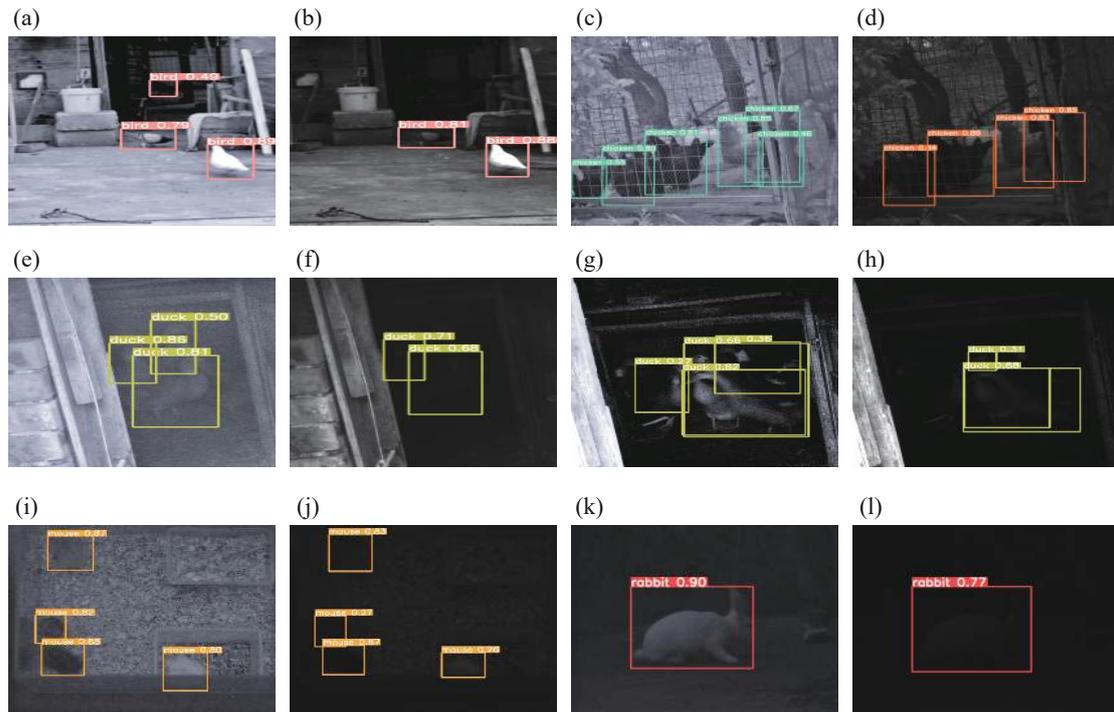


图 8 增强与未增强的目标检测可视化结果

Fig. 8 Visualization results of enhanced and unenhanced target detection

表 4 增强与未增强的目标检测可视化结果对比

Tab. 4 Visualization results comparison of enhanced and unenhanced target detection

组别	应检个数/实检个数	检测精度	结论
a 鸽类增强后检测结果	3/3	0.49/0.79/0.89	未增强图像出现漏检数量1, 阴影处鸽类未检出
b 鸽类未增强检测结果	3/2	0.00/0.81/0.88	
c 鸡类增强后检测结果	6/6	0.55/0.80/0.81/0.85/0.57/0.46	未增强图像出现漏检数量2, 角落处和被遮挡的鸡类未检出
d 鸡类未增强检测结果	6/4	0.00/0.84/0.89/0.83/0.85/0.00	
e 鸭类增强后检测结果	3/3	0.50/0.86/0.81	未增强图像出现漏检数量1, 深处的鸭类未检出
f 鸭类未增强检测结果	3/2	0.00/0.71/0.68	
g 鸭类增强后检测结果	4/4	0.27/0.66/0.82/0.36	未增强图像出现漏检数量2, 左侧和右侧阴影内鸭类未检出
h 鸭类未增强检测结果	4/2	0.00/0.31/0.88/0.00	
i 鼠类增强后检测结果	4/4	0.87/0.82/0.88/0.80	增强图像中鼠类检测精度提升明显, 从0.27、0.26提升至0.82、0.80
j 鼠类未增强检测结果	4/4	0.83/0.27/0.87/0.26	
k 兔类增强后检测结果	1/1	0.90	增强图像中兔类检测精度提升明显, 从0.77提升至0.90
l 兔类未增强检测结果	1/1	0.77	

综上所述,本文中的 Im-YOLOv8 提高了特征提取能力,由表 2 数据可知相较于原始的 YOLOv8, Im-YOLOv8 的平均目标检测精度 mAP 提升了 1.5%。此外,通过采用 Zero-Denoise 方法,在夜间动物数据集上进行训练,并通过预处理优化图像质量。从表 3 和图 8 可以明显看出, Zero-Denoise 在减少漏检以及提升检测精度方面取得了显著成效。特别是与 Im-YOLOv8 结合使用时, Zero-Denoise 进一步使检测精度 mAP@0.5 值提升了 5.6%,与其他低光图像增强方法相比,至少达到了 3.1% 的提升。与仅使用 YOLOv8 相比提升了 7.1%。上述结果表明该方法在获得更高的动物平均检测精度的同时,有效减少了漏检现象,能够提供更为优质的视觉图像质量。

5 结 论

针对动物夜间实时监测所面临的图像曝光度

低、对比度低、特征提取困难等难题,研究了轻量化自监督深度神经网络 Zero-Denoise 和改进型 YOLOv8 模型,并通过搭建监测试验系统自建夜间动物数据集进行夜间动物目标的图像增强与精准识别方法研究。得到了如下结论:(1) 本文研究的 Zero-Denoise 在 LOL 数据集上的 PSNR、SSIM 与 MAE 指标分别达到 28.53、0.76、26.15,运行速度达到 55 frames,兼顾低光照增强与实时性;(2) 改进型 YOLOv8 模型相较于原模型在自建数据集上的目标检测 mAP 值提升了 1.5%,提高了夜间动物目标的特征提取能力;(3) Zero-Denoise 与 Im-YOLOv8 的结合后提升了对夜间动物目标图像的增强能力,且能够减少漏检的情况,相较于使用 SCI 等增强算法,其目标检测 mAP 值提升了 3.1% 以上,较仅使用 YOLOv8 mAP 值时提升了 7.1%。由此可知,本文研究方法可以获得优越的动物目标检测率,尤其适合夜间暗环境下动物目标的被动式快速智能监测方法的进一步研究。

参考文献:

- [1] QI Y L, YANG ZH, SUN W H, *et al.*. A comprehensive overview of image enhancement techniques[J]. *Archives of Computational Methods in Engineering*, 2022, 29(1): 583-607.
- [2] 刘彦磊, 李孟喆, 王宣宣. 轻量化 YOLOv5s 车载红外图像目标检测[J]. *中国光学 (中英文)*, 2023, 16(5): 1045-1055.
LIU Y L, LI M ZH, WANG X X. Lightweight YOLOv5s vehicle infrared image target detection[J]. *Chinese Optics*, 2023, 16(5): 1045-1055. (in Chinese).
- [3] HU H F. Illumination invariant face recognition based on dual-tree complex wavelet transform[J]. *IET Computer Vision*, 2015, 9(2): 163-173.
- [4] MUNIAN Y, MARTINEZ-MOLINA A, MISERLIS D, *et al.*. Intelligent system utilizing HOG and CNN for thermal image-based detection of wild animals in nocturnal periods for vehicle safety[J]. *Applied Artificial Intelligence*, 2022, 36(1): 2031825.
- [5] MURUGAN R A, SATHYABAMA B. Object detection for night surveillance using Ssan dataset based modified Yolo algorithm in wireless communication[J]. *Wireless Personal Communications*, 2023, 128(3): 1813-1826.
- [6] BHATT D, PATEL C, TALSANIA H, *et al.*. CNN variants for computer vision: history, architecture, application, challenges and future scope[J]. *Electronics*, 2021, 10(20): 2470.
- [7] 任凤雷, 周海波, 杨璐, 等. 基于双注意力机制的车道线检测[J]. *中国光学 (中英文)*, 2023, 16(3): 645-653.
REN F L, ZHOU H B, YANG L, *et al.*. Lane detection based on dual attention mechanism[J]. *Chinese Optics*, 2023, 16(3): 645-653. (in Chinese).
- [8] LI CH Y, GUO J CH, PORIKLI F, *et al.*. LightenNet: a Convolutional Neural Network for weakly illuminated image enhancement[J]. *Pattern Recognition Letters*, 2018, 104: 15-22.
- [9] DING X, HU R M. Learning to see faces in the dark[C]. *2020 IEEE International Conference on Multimedia and Expo (ICME)*, IEEE, 2020: 1-6.
- [10] LI CH Y, GUO CH L, CHEN CH L. Learning to enhance low-light image via zero-reference deep curve estimation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 40(8): 4225-4238.
- [11] JIANG Y F, GONG X Y, LIU D, *et al.*. EnlightenGAN: deep light enhancement without paired supervision[J]. *IEEE Transactions on Image Processing*, 2021, 30: 2340-2349.
- [12] FU Y, HONG Y, CHEN L W, *et al.*. LE-GAN: Unsupervised low-light image enhancement network using attention

- module and identity invariant loss[J]. *Knowledge-Based Systems*, 2022, 240: 108010.
- [13] WANG R J, JIANG B, YANG CH, *et al.*. MAGAN: Unsupervised low-light image enhancement guided by mixed-attention[J]. *Big Data Mining and Analytics*, 2022, 5(2): 110-119.
- [14] MONAKHOVA K, RICHTER S R, WALLER L, *et al.*. Dancing under the stars: video denoising in starlight[C]. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2022: 16220-16230.
- [15] CHEN J R, KAO S H, HE H, *et al.*. Run, Don't walk: chasing higher FLOPS for faster neural networks[C]. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2023: 12021-12031.
- [16] WEI CH, WANG W J, YANG W H, *et al.*. Deep retinex decomposition for low-light enhancement[C]. *British Machine Vision Conference 2018*, BMVA Press, 2018: 155.
- [17] ZHANG Y, DI X G, ZHANG B, *et al.*. Self-supervised low light image enhancement and denoising[J]. arXiv preprint arXiv: 2103.00832, 2021.
- [18] ZHANG Q L, YANG Y B. SA-Net: Shuffle attention for deep convolutional neural networks[C]. *2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2021: 2235-2239.
- [19] XU X ZH, JIANG Y Q, CHEN W H, *et al.*. DAMO-YOLO: a report on real-time object detection design[J]. arXiv preprint arXiv: 2211.15444, 2022.
- [20] TONG Z J, CHEN Y H, XU Z W, *et al.*. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism[J]. arXiv preprint arXiv: 2301.10051, 2023.
- [21] WU W H, WENG J, ZHANG P P, *et al.*. URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement[C]. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2022: 5891-5900.
- [22] ZHANG F, LI Y, YOU SH D, *et al.*. Learning temporal consistency for low light video enhancement from single images[C]. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2021: 4965-4974.
- [23] MA L, MA T Y, LIU R SH, *et al.*. Toward fast, flexible, and robust low-light image enhancement[C]. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2022: 5627-5636.
- [24] LV F F, LU F, WU J H, *et al.*. MBLEN: Low-light image/video enhancement using CNNs[C]. *British Machine Vision Conference 2018*, BMVA Press, 2018: 220.

作者简介:



王 驰(1982—), 男, 河南周口人, 博士(后), 教授, 2009年于天津大学获得博士学位, 现为上海大学机电工程与自动化学院教师, 主要从事精密测试技术及仪器等方面的研究。E-mail: wangchi@shu.edu.cn



陈金波(1980—), 男, 内蒙古呼和浩特人, 博士, 工程师, 2002年于上海大学获得硕士学位, 2014年于上海大学获得博士学位, 现为上海大学机电工程与自动化学院教师, 主要从事机器视觉与检测方面的研究。E-mail: jbchen@shu.edu.cn