

# 基于校园监控的多帧图像超分辨率重建技术

陈亮, 梁暄浩

(沈阳理工大学 自动化与电气工程学院, 沈阳 110159)

**摘要:** 为解决校园监控视频质量低、细节信息少的问题,在 BasicVSR 网络模型的基础上提出一种增加辅助机制的多帧图像超分辨率重建技术。首先,选择双向循环机制的 BasicVSR 网络以保证信息传递的多样性与连贯性;其次,在图像配准模块增加辅助机制以获得关键帧特征,修正主干网络的特征图,该辅助机制在一定程度上解决了图像长期传递的信息误差,减少了纹理造假现象;最后,使用早期图像融合的方式,将前向分支与后向分支的特征图像融合为新的特征并使用亚像素卷积进行图像重建。在自制校园监控数据集上的对比实验结果表明,改进后模型的峰值信噪比相比于原模型提升了 1.51 dB,提高了监控视频的清晰度,具有较好的应用价值。

**关键词:** 校园监控;超分辨率重建技术;图像配准;图像融合

中图分类号: TP391.41 文献标志码: A DOI:10.3969/j.issn.1003-1251.2024.04.002

## Multi Frame Image Super-resolution Reconstruction Technology Based on Campus Monitoring

CHEN Liang, LIANG Xuanhao

(Shenyang Ligong University, Shenyang 110159, China)

**Abstract:** To solve the problem of low quality and insufficient detail information in campus surveillance videos, a multi frame image super-resolution reconstruction technique with added auxiliary mechanisms is proposed based on the BasicVSR network model. Firstly, the BasicVSR network with a bidirectional loop mechanism is chosen to ensure the diversity and coherence of information transmission. Secondly, an auxiliary mechanism is added to the image registration module to obtain key-frame features and correct the feature maps of the backbone network. This auxiliary mechanism to some extent solves the information error in long-term image transmission and reduces texture falsification. Finally, using early image fusion methods, the feature images of the forward and backward branches are fused into new features and reconstructed using sub-pixel convolution. The comparative experimental results of a self-made campus monitoring dataset show that the improved model has a 1.51 dB increase in peak signal-to-noise ratio compared to the original model, which improves the clarity of monitoring videos and has good application value.

**Key words:** campus monitoring; super-resolution reconstruction technology; image registration; image fusion

为预防校园暴力行为,解决校园内的安全隐患,如何提升校园监控视频质量成为热点话题。

针对校园监控视频分辨率较低,监控摄像头位置距离人像位置较远的问题,使用多帧图像超分辨率重

建技术将低分辨率图像重建为高分辨率图像<sup>[1]</sup>。

多帧图像超分辨率重建技术旨在通过连续帧图像序列之间的关系<sup>[2]</sup>,利用帧间信息对当前帧进行信息补充,增加更多真实的细节信息。多帧图像超分辨率重建技术的关键点在于图像配准模块<sup>[3]</sup>,若相邻帧之间的信息传递存在误差,会导致重建的图像出现伪影和偏差,降低重建图像质量<sup>[4]</sup>。多帧图像超分辨率重建技术可分为基于单向或双向神经网络传输方式的多帧图像超分辨率重建技术。基于双向神经网络传输方式的多帧图像超分辨率重建技术通过前向分支和后向分支的信息传输,得到相邻帧之间的相关图像信息,用以重建图像。

随着深度学习算法的发展,众多学者将卷积神经网络应用到了图像超分辨率重建技术中<sup>[5]</sup>。从2017年的时空亚像素卷积网络<sup>[6]</sup>开始,建立了完整的多帧图像超分辨率重建流程。2018年,Sajjadi等<sup>[7]</sup>提出了FRVSR网络,该网络使用了帧循环机制解决了网络计算成本的问题。2019年,Wang等<sup>[8]</sup>提出了EDVR网络,使用金字塔卷积进行图像的特征提取,提升了可变形卷积图像配准模块的精准度。2020年,Saadatkah等<sup>[9]</sup>提出了TGAN网络,使用组内分组的机制进行注意力融合,解决了图像信息利用率的问题,但增加了计算量,模型运行时间较长。2021年,超分辨率重建(BasicVSR)网络<sup>[10]</sup>被提出,该网络首次使用双向循环机制进行超分辨率重建,重建后的图像质量得到大幅度提升。

上述技术虽提升了重建图像的整体质量,但是图像细节信息的重建并不理想。本文结合超分辨率重建技术与实际应用背景,提出一种校园监控视频下的多帧图像超分辨率重建技术。在BasicVSR网络的超分辨重建网络中,增加辅助机制进行主干网络的特征矫正。改进后的网络可解决图像信息在长时间传输过程中的计算误差和信息重复利用导致的偏差问题,减少重建图像的纹理造假现象。

## 1 BasicVSR 网络算法

### 1.1 BasicVSR 网络

强化基础组件的 BasicVSR 网络结构如图 1 所示。BasicVSR 网络提出了一个双向视频流的循环结构配准方法,强化基础组件(传播、对齐、聚合、上采样)通过重用和重新设计基础组件,减少

复杂组件的加入。

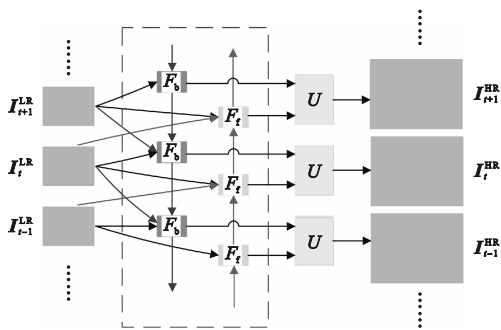


图 1 BasicVSR 网络结构

Fig.1 BasicVSR network structure

给定输入帧序列图像,  $I_t^{LR}$  为当前帧的低分辨率图像,  $I_{t+1}^{LR}$ 、 $I_{t-1}^{LR}$  分别为当前帧的两个相邻帧低分辨率图像。输入图像在前向分支与后向分支中进行图像配准,具体表达式为

$$h_t^b = F_b(I_t^{LR}, I_{t+1}^{LR}, h_{t+1}^b) \quad (1)$$

$$h_t^f = F_f(I_t^{LR}, I_{t-1}^{LR}, h_{t-1}^f) \quad (2)$$

式中:算子  $F_b(\cdot)$ 、 $F_f(\cdot)$  分别为后向分支与前向分支;  $h_t^b$ 、 $h_t^f$  分别表示  $t$  时刻的后向分支与前向分支配准后的特征图,分别作为下一时刻支持帧的后向隐藏状态和前向隐藏状态,同时也作为融合模块的基础。融合模块使用早期融合的方法,其数学表达式为

$$F_{\text{fusion}} = [I_t^{LR}, h_t^f, h_t^b] \quad (3)$$

式中:  $F_{\text{fusion}}$  为融合结果;  $[\cdot]$  表示拼接,输入是当前帧的低分辨率图像与各个分支配准之后的特征图。上采样部分使用了亚像素卷积的方式,其数学表达式为

$$I_t^{HR} = U(F_{\text{fusion}}) \quad (4)$$

式中:算子  $U(\cdot)$  表示上采样模块,包括多个卷积特征提取和一个亚像素卷积上采样,输入是融合模块融合之后的结果;  $I_t^{HR}$  是输出的当前帧高分辨率图像。

### 1.2 图像配准网络

为了计算帧间信息,使重建的图像可以获得相邻帧之间准确的信息补偿,研究人员提出了光流法<sup>[11]</sup>,用于场景中物体的运动估计<sup>[12]</sup>,作为图像配准的算法基础。可变形卷积网络(DCN)<sup>[13]</sup>提供了可变形池化模型,基于增强空间采样的思想,计算模块中具有额外偏移量的位置,并从目标任务中学习偏移量。可变形卷积建模能力强、使用方便、感受野可扩大,适用于多帧图像超分辨率重建技术中。

BasicVSR 网络的图像配准模块基于光流法

与可变形卷积的算法,提出了光流指导可变形卷积网络。该网络通过光流法计算图像之间的偏移量,使用可变形卷积对偏移量和特征图进行图像配准。

### 1.3 图像融合网络

图像融合将相邻帧图像的相关信息融合到一张图像,使重建帧图像得到更多的细节信息和真实信息。图像融合的基本方法可分为三类:像素融合、特征融合以及决策融合<sup>[14]</sup>。在多帧图像超分辨率重建技术中,常用的方法是在时空亚像素卷积网络中提出的早期融合、缓慢融合、3D卷积融合方式<sup>[15]</sup>。

BasicVSR网络使用早期融合的方式进行图像融合,如图2所示。将输入层的时间深度 $d_0$ 与帧数 $D_0$ 相匹配,其中 $D_0=5$ ,对不同维度进行了颜色编码,并省略了到2D空间的输出映射, $n_1$ 、 $n_2$ 、 $r^2$ 代表了不同的维度。早期融合在卷积之前拼接不同的时序信息,也就是多通道卷积。

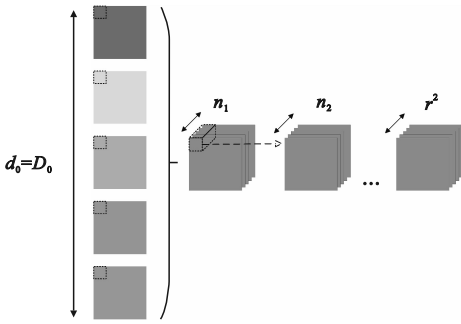


图2 早期融合网络结构

Fig. 2 Early fusion network structure

## 2 改进算法

### 2.1 图像配准网络改进

本文针对网络的长距离循环传播机制所产生的误差累积,提出一种辅助机制减少因计算误差导致的一些细节损失,同时通过该机制增加重建帧的真实信息。增加辅助机制后的图像配准网络模型如图3所示。图中: $s_t^f$ 为经过光流估计后特征图; $\bar{h}_t^f$ 为光流融合模块输出的特征图。

以前向分支为例,输入为 $I_t^{LR}$ 、 $I_{t-1}^{LR}$ 、 $h_{t-1}^f$ ,输出的数学计算表达式为

$$s_t^f = S(I_t^{LR}, I_{t-1}^{LR}) \quad (5)$$

$$\bar{h}_t^f = W(h_{t-1}^f, s_t^f) \quad (6)$$

$$\hat{h}_t^f = C\{\bar{h}_t^f, T(I_{t-1}^{LR}, I_t^{LR}, I_{t+1}^{LR})\} \quad (7)$$

$$h_t^f = R(I_t^{LR}, \bar{h}_t^f, \hat{h}_t^f) \quad (8)$$

式中:算子 $S(\cdot)$ 、 $W(\cdot)$ 、 $T(\cdot)$ 、 $C(\cdot)$ 、 $R(\cdot)$ 分别表示光流估计模块、光流融合模块、辅助机制、卷积运算、特征矫正模块; $h_t^f$ 表示前向分支的输出,为增加辅助机制后的特征图;后向分支的算法流程与前向分支相同, $h_t^b$ 表示后向分支的输出。

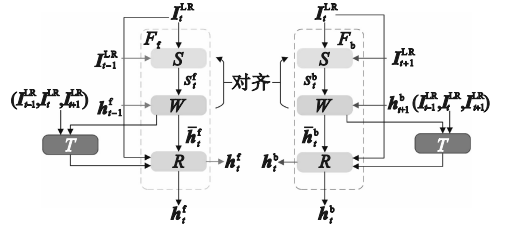


图3 图像配准网络结构

Fig. 3 Image registration network structure

### 2.2 辅助机制网络结构

在图像配准模块中,原特征校正模块之前引入一个额外的辅助机制。如果当前参考帧在关键帧集合内,那么由关键帧及其相邻支持帧组成该模块的输入,经过辅助机制的特征提取,得到可用于矫正主干网络的特征图像。该辅助机制模块结构如图4所示。

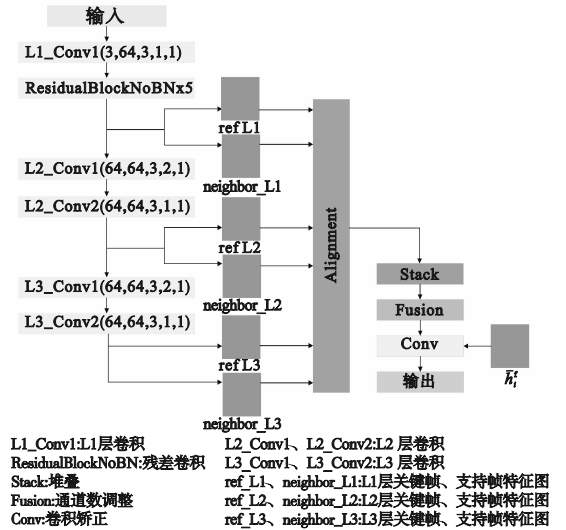


图4 辅助机制网络结构

Fig. 4 Auxiliary mechanism network structure

首先,输入图像经过基础的特征提取与5个残差块得到L1层的特征图,选取关键帧图像与支持帧图像分别作为L1层输入。之后,经过L2卷积层,得到比上一层级尺寸小一倍的特征图,做L2层级的处理。L3层级做相同处理。将L1、L2以及L3层级的特征图一同输入到对齐模块Alignment中做图像配准,得到图像配准后的特征

图。该特征图经过堆叠与通道数调整,与主干网络图像配准后的特征图  $\bar{h}_i^f$  进行卷积,调整特征图误差,得到输出的特征图。

Alignment 首先将同一时间下的关键帧和支持帧的特征图作为输入。从第 L3 层开始,对该层的关键帧和支持帧的特征图进行融合,通过卷积得到偏移量,利用该偏移量和关键帧的特征图经过 DCN 卷积得到对齐后的特征图。将 L3 层学习得到的特征图输送给 L2 层,L2 层的偏移量由该层的两个特征图之间的偏移量和 L1 层偏移量上采样后共同组成。特征图由该层可变形卷积的输出和上层特征图经过上采样处理之后的特征图组成。重复此过程,得到对齐后的特征图。最后,将 L1 层的特征图与关键帧特征图进行图像融合,得到特征图之间的偏移量,将该偏移量与图像配准后的特征图一起输入到可变形卷积模块中做特征对齐,得到该对齐模块的特征图。

本文提出的辅助机制减少了长时间传播的信息重复利用的误差,增加了重建帧的真实信息,并起到了矫正重建帧的作用。

### 3 实验与分析

#### 3.1 自制数据集

REDS 数据集<sup>[16]</sup>常被用于多帧图像超分辨率重建技术,该数据集包含不同环境下的 54 000 张图像,每一个文件夹包含 100 张连续帧图像。本文参考 REDS 数据集格式,制作校园监控数据集。以学校教室监控视频为基础,人工筛选不同教室的监控视频,选取正在上课且教室光线较为明亮的视频。数据集制作流程如图 5 所示。

首先,将选取的视频分帧为连续帧图像,将 100 张连续帧图像作为一组保存至文件夹内,并重命名为数据集要求的格式。之后,将数据集的 000 文件至 239 文件作为训练集的高分辨率图像,240 文件至 269 文件作为验证集高分辨率图像。最后,对所有图像做 4 倍下采样,得到低分辨率的图像用以训练和测试模型。

#### 3.2 实验结果与分析

本文所有实验基于表 1 环境运行。为验证改进网络的性能,对改进后的重建网络(Ours)和其他多帧图像超分辨率重建算法在自制校园监控数据集上的表现进行对比。实验使用校园监控数据集上的训练集,包含 54 000 张图片,测试数据使用相同的 3 000 张图片,保证对比实验的变量唯

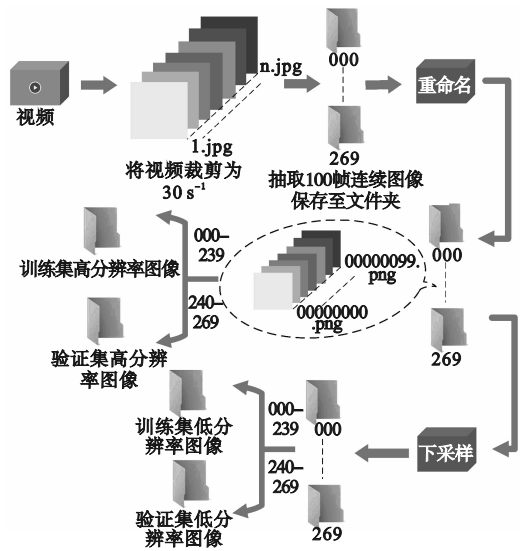


图 5 数据集制作流程图

Fig. 5 Dataset production flow chart

一性。在网络训练过程中,使用相同的网络参数,如表 2 所示。

表 1 实验环境

Table 1 Experimental environment

名称	相关配置
操作系统	Linux
处理器	Tesla V100
内存类型	HBM2
主硬盘	Tesla V100-SXM2-16 GB
软件环境	PyTorch 1.10.0, Python 3.8

表 2 网络实验参数

Table 2 Network experiment parameters

网络参数	数值
单次样本训练次数	30
单次测试样本数	100
迭代次数	300 000
学习率	0.000 1
学习衰减率	0.000 025

本实验采用峰值信噪比(PSNR)和结构相似度(SSIM)<sup>[17]</sup>进行质量评估。峰值信噪比通过简单的均方差计算<sup>[18]</sup>得到 PSNR 值,根据数值的大小判断网络的重建效果。结构相似度可用于计算两张图像之间的比例,用以衡量图像失真程度。SSIM 的计算方式更符合人眼的直观感受,属于感知模型。同时,对网络运行时间和模型参数量进行比较。实验结果如表 3 所示,其中 PSNR、SSIM 取测试结果的平均值。

表 3 与其他方法比较

Table 3 Comparisons with other methods

网络方法	PSNR/ dB	SSIM	运行时间/ ms	参数量/ MB
Bicubic	16.95	0.546 0	—	—
DUF <sup>[19]</sup>	18.37	0.5912	974	5.8
RBPN <sup>[20]</sup>	20.81	0.619 7	1 507	12.2
EDVR	20.92	0.628 1	378	20.6
BasicVSR	22.16	0.826 8	63	6.3
BasicVSR ++ <sup>[21]</sup>	21.91	0.811 9	77	7.3
本文方法	23.67	0.877 1	66	6.5

本文选用多帧超分辨率重建技术领域的经典模型进行对比实验,每种网络使用相同的迭代次数 300 000、单次测试样本数 100 以及相同的测试集与评估指标。从表 3 中可以得出,本文改进的网络相较于其他网络在峰值信噪比和结构相似度上有明显提升,相较于原网络提升了 1.51 dB/0.05,说明增加了辅助机制使重建后的图像具有更多的真实信息。同时,网络仅增加了小部分的运行时间与模型参数量就使得图像重建效果得到了较大提升。

本文算法与其他多帧超分辨率重建算法重建效果如图 6 所示,图中左侧图像是从校园监控数据集中测试集任意挑选的一张图像,图中右方所示为不同重建网络的重建效果,其中 original image 为原图像中红框部分的图像,GT 为真实高分辨率图像。从图中的重建效果来看,本文网络算法重建后的图像与 GT 图像更加相近,与原模型相比减少了图像中不该存在的纹理信息,拥有更真实的细节信息。

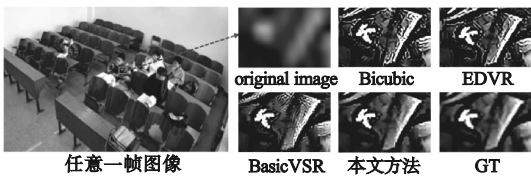


图 6 不同模型重建结果对比

Fig. 6 Comparison of reconstruction results of different models

## 4 结论

本文在 BasicVSR 基础网络的图像配准网络中增加了辅助机制,使用关键帧对主干网络的特征图进行矫正,减少了重建图像的纹理造假现象

与误差。从实验结果中可以看出,改进后的网络相比于原网络在峰值信噪比上提高了 1.51 dB,SSIM 值提升了 0.05。本文模型对低分辨率图像进行重建后,可以获得更多的细节信息,减少了重建图像的纹理造假现象。

## 参考文献 (References):

- [1] YANG W M, ZHANG X C, TIAN Y P, et al. Deep learning for single image super-resolution; a brief review [J]. IEEE Transactions on Multimedia, 2019, 21 (12): 3106 – 3121.
- [2] 付凌云, 王正勇, 熊淑华, 等. 采用光流特征对齐的压缩视频超分辨率重建 [J]. 通信技术, 2022, 55 (12): 1560 – 1567.
- [3] FU L Y, WANG Z Y, XIONG S H, et al. Super-resolution reconstruction of compressed video using optical flow feature alignment [J]. Communications Technology, 2022, 55 (12): 1560 – 1567. (in Chinese)
- [4] 吴昊, 赖惠成, 钱绪泽, 等. 基于光流残差的视频超分辨率重建算法 [J]. 计算机工程与应用, 2022, 58 (15): 220 – 228.
- [5] WU H, LAI H C, QIAN X Z, et al. Video super-resolution reconstruction algorithm based on optical flow residual [J]. Computer Engineering and Applications, 2022, 58 (15): 220 – 228. (in Chinese)
- [6] 刘村, 李元祥, 周拥军, 等. 基于卷积神经网络的视频图像超分辨率重建方法 [J]. 计算机应用研究, 2019, 36 (4): 1256 – 1260, 1274.
- [7] LIU C, LI Y X, ZHOU Y J, et al. Video image super-resolution reconstruction method based on convolutional neural network [J]. Application Research of Computers, 2019, 36 (4): 1256 – 1260, 1274. (in Chinese)
- [8] LIU H Y, RUAN Z B, ZHAO P, et al. Video super-resolution based on deep learning; a comprehensive survey [J]. Artificial Intelligence Review, 2022, 55 (8): 5981 – 6035.
- [9] 林琦, 陈婧, 曾焕强, 等. 基于多尺度特征残差学习卷积神经网络的视频超分辨率方法 [J]. 信号处理, 2020, 36 (1): 50 – 57.
- [10] LIN Q, CHEN J, ZENG H Q, et al. Video super-resolution method based on multi-scale characteristics residual learning convolutional neural network [J]. Journal of Signal Processing, 2020, 36 (1): 50 – 57. (in Chinese)
- [11] SAJJADI M S M, VEMULAPALLI R, BROWN M. Frame-recurrent video super-resolution [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 6626 – 6634.
- [12] WANG X, CHAN K C K, YU K, et al. Edvr: Video restoration with enhanced deformable convolutional networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019: 1954 – 1963.
- [13] SAADATKHAH N, CARILLO G A, ACKERMANN S, et al. Experimental methods in chemical engineering; thermogravimetric analysis—TGA [J]. The Canadian Journal of Chemical Engineering, 2020, 98 (1): 34 – 43.
- [14] CHAN K C K, WANG X T, YU K, et al. BasicVSR: the search for essential components in video super-resolution and beyond [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN, USA: IEEE, 2021: 4945 – 4954.

- [11] 王剑平,朱芮,张果,等. 帧差与快速密集光流结合的图像法测流研究[J]. 工程科学与技术,2022,54(4):195-207. WANG J P,ZHU R,ZHANG G, et al. Image flow measurement based on the combination of frame difference and fast and dense optical flow[J]. Advanced Engineering Sciences, 2022,54(4):195-207. (in Chinese)
- [12] 郑萌萌,钱慧芳,周璇. 基于监控视频的 Farneback 光流算法的人体异常行为检测[J]. 国外电子测量技术,2021,40(3):16-22. ZHENG M M,QIAN H F,ZHOU X. Human abnormal action detection based on the Farneback optical flow arithmetic of surveillance video[J]. Foreign Electronic Measurement Technology,2021,40(3):16-22. (in Chinese)
- [13] 赵威,宋建辉,刘砚菊,等. 基于改进 TSRN 的图像超分辨率重建算法[J]. 沈阳理工大学学报,2023,42(3):41-45,53. ZHAO W,SONG J H,LIU Y J, et al. Image super-resolution reconstruction algorithm based on improved TSRN[J]. Journal of Shenyang Ligong University,2023,42(3):41-45,53. (in Chinese)
- [14] 曲海成,胡倩倩,张雪聪. 结合信息感知与多尺度特征的红外与可见光图像融合[J]. 红外技术,2023,45(7):685-695. QU H C,HU Q Q,ZHANG X C. Infrared and visible image fusion combining information perception and multiscale features[J]. Infrared Technology,2023,45(7):685-695. (in Chinese)
- [15] 江俊君,程豪,李震宇,等. 深度学习视频超分辨率技术综述[J]. 中国图象图形学报,2023,28(7):1927-1964. JIANG J J,CHENG H,LI Z Y, et al. Deep learning based video-related super-resolution technique;a survey[J]. Journal of Image and Graphics,2023,28(7):1927-1964. (in Chinese)
- [16] 秦昊宇,葛瑶,张力波,等. 基于自注意力机制的视频超分辨率重建[J]. 计算机技术与发展,2022,32(8):42-48. QIN H Y,GE Y,ZHANG L B, et al. Video super-resolution reconstruction based on self attention mechanism[J]. Computer Technology and Development,2022,32(8):42-48. (in Chinese)
- [17] 杨巨成,左美然,魏峰,等. 基于误差反馈和面部后先验信息的人脸超分辨率重建[J]. 天津科技大学学报,2022,37(2):35-42. YANG J C,ZUO M R,WEI F, et al. Face super-resolution using error feedback and facial posterior[J]. Journal of Tianjin University of Science & Technology,2022,37(2):35-42. (in Chinese)
- [18] 张思源,王国胤,刘群,等. 基于多粒度特征融合的边缘一致性图像补全[J]. 控制与决策,2022,37(12):3240-3250. ZHANG S Y,WANG G Y,LIU Q, et al. Edge consistent image completion based on multi-granularity feature fusion[J]. Control and Decision,2022,37(12):3240-3250. (in Chinese)
- [19] JO Y,OH S W,KANG J, et al. Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA:IEEE,2018:3224-3232.
- [20] HARIS M, SHAKHAROVICH G, UKITA N. Recurrent back-projection network for video super-resolution[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA:IEEE,2019:3897-3906.
- [21] CHAN K C K,ZHOU S,XU X, et al. BasicVSR++: Improving video super-resolution with enhanced propagation and alignment[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, LA, USA:IEEE,2022:5972-5981.

(责任编辑:和晓军)

(上接第6页)

- [3] SHI Q,ZHANG C X,CHEN Z, et al. An infrared small target detection method using coordinate attention and feature fusion[J]. Infrared Physics & Technology,2023,131:104614.
- [4] LI R H,SHEN Y. YOLOS-IST:a deep learning method for small target detection in infrared remote sensing images based on super-resolution and YOLO[J]. Signal Processing,2023,208:108962.
- [5] YU C,LIU Y P,WU S H, et al. Infrared small target detection based on multiscale local contrast learning networks[J]. Infrared Physics & Technology,2022,123:104107.
- [6] JU M R,LUO J N,LIU G Q, et al. ISTDet:an efficient end-to-end neural network for infrared small target detection[J]. Infrared Physics & Technology,2021,114:103659.
- [7] WANG K D,LI S Y,NIU S S, et al. Detection of infrared small targets using feature fusion convolutional network[J]. IEEE Access,2019,7:146081-146092.
- [8] HOWARD A,SANDLER M,CHEN B, et al. Searching for MobileNetV3[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South):IEEE,2020:1314-1324.
- [9] ZHENG Z H,WANG P,LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence,2020,34(7):12993-13000.
- [10] YU J H,JIANG Y N,WANG Z Y, et al. UnitBox:an advanced object detection network[C]//Proceedings of the 24th ACM International Conference on Multimedia. New York:ACM,2016:516-520.
- [11] TONG Z J,CHEN Y H,XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. arXiv:2301.10051(2023-01-24)[2023-05-04]. https://arxiv.org/abs/2301.10051.

(责任编辑:和晓军)