

基于深度学习多帧海面图像超分辨率算法研究

张玉爽, 韩文波*, 黄丹飞**, 赵丽颖, 钟艾琦

长春理工大学光电工程学院, 吉林 长春 130022

摘要 近年来,深度学习在图像超分辨重建方面表现出色。然而,由于复杂的海洋环境,传统的图像超分辨算法存在着调节参数困难等问题,单帧图像超分辨率算法亦存在病态的恢复,生成的像素点具有不确定性。提出了一种多帧图像的超分辨重建算法以用于海面图像重建研究,利用深度学习中的卷积神经网络来学习多帧低分辨率图像与高分辨率图像之间的映射关系,从而实现超分辨率重建。同时,由于海洋监测成像系统需要更多的高频信息来判别目标及锁定轮廓,提出应用残差网络框架来改善网络重建图像的质量,恢复更多的高频信息,丰富图像的细节。实验结果表明,所提算法具有较好的图像重建能力,与其他方法相比具有较好的主客观评价结果。

关键词 图像处理;深度学习;超分辨率;多帧图像

中图分类号 TP751.1

文献标志码 A

doi: 10.3788/LOP202158.1610005

Research on Multi-Frame Sea Surface Image Super Resolution Algorithm Based on Deep Learning

Zhang Yushuang, Han Wenbo*, Huang Danfei**, Zhao Liying, Zhong Aiqi

College of Optoelectronic Engineering, Changchun University of Science and Technology,

Changchun, Jilin 130022, China

Abstract In recent years, deep learning has made a great achievement in image super-resolution reconstruction. Due to the complex ocean environment, the traditional image super-resolution algorithm has some problems, such as the difficulty in adjusting parameters. In addition, the single frame image super-resolution algorithm has an ill conditioned recovery and the generated pixels are uncertain. In this paper, a super-resolution reconstruction algorithm of multi-frame images is proposed for the study of sea surface image reconstruction. The convolution neural network in deep learning is used to learn the mapping relationship between multi-frame low-resolution images and high-resolution images, so as to realize super-resolution reconstruction. At the same time, because the ocean monitoring imaging system needs more high-frequency information to identify the target and lock contour, the residual network framework is proposed to improve the quality of network reconstruction images, recover more high-frequency information and enrich the image details. The experimental results show that the proposed algorithm has a better image reconstruction ability and better subjective and objective evaluation results compared with other methods.

Key words image processing; deep learning; super resolution; multi-frame image

OCIS codes 100.4996; 100.6640; 100.3010

1 引言

图像超分辨重建是一类经典的计算机视觉问

题,旨在从一个或多个低分辨率图像中恢复出高分辨率图像,其应用领域及其广泛,在遥感图像处理^[1]、军事侦察、医学图像处理^[2]、天文图像处理以

收稿日期: 2020-11-06; 修回日期: 2020-11-26; 录用日期: 2020-12-08

基金项目: 国家自然科学基金(61890963)

通信作者: *hanwenbo@cust.edu.cn; **huangd_f@163.com

及视频监控等方面都有着重要的应用前景。文献[3]使用插值算法对图像进行分辨率提升,随后这一算法被逐渐借鉴到多帧图像及视频超分辨率算法当中。超分辨算法大多是基于学习的方法,通过学习低分辨率图像到高分辨率图像的映射关系来重建超分图像。作为将卷积神经网络作为图像超分辨的先驱,图像超分辨率的深度卷积学习(SRCNN)在2014年由Dong等^[4]提出,此算法可用端到端的训练方式来学习映射关系,实验结果表明,该方法的效果及性能更优于非深度学习的传统算法。文献[5]将生成对抗网络(GAN)首次应用到图像超分辨重建上,SRGAN使用GAN来生成更多的细节,使图像更加逼真。文献[6]提出一种可用于改进图像超分辨率重建质量的双判别器超分辨率重建网络(DDSRN),该网络在生成对抗网络的基础上增加一个判别器,能够有效避免网络的崩溃。文献[7]提出了基于卷积神经网络(CNN)的单帧图像超分辨VDSR算法,其网络有20个卷积层,引入了全局残差学习,并以较高的学习率训练模型,其超分性能得到较大幅度提升。文献[8]提出一种新型网络结构,图像重建由最后的反卷积上采样来实现。文献[9]将卷积神经网络应用于深度图像超分辨重建中,将深度图进行层次化的抽象处理以自主提取特征。文献[10]中提出一种去卷积的快速图像超分辨率方法,在获得较高质量图像的同时提升了处理速度。虽然单帧图像的超分辨重建算法目前已经达到了较高水平,但在实际情况中,采集到的图像多为具有时空关联性的多帧图像序列,多帧图像的超分辨率重建更具有实际意义。文献[11]提出一种基于动态滤波器的多帧图像超分辨的VSR-DUF算法。文

献[12]提出一种亚像素运动补偿,可有效融合多帧图像的细节信息。文献[13]提出一种时空亚像素卷积网络,可以有效地利用时间冗余提高重建精度并保持实时速度。文献[14]提出一种任务导向型光流法,可用于多帧图像的增强处理。

受到以上图像超分辨重建算法的启发,针对海上目标“看不清、辨不明”的问题,本文利用多帧图像序列具有时空相关性的特点,采用多帧图像的超分辨算法,旨在应用深度学习对多帧海面图像进行超分辨重建。首先对多帧图像进行光流估计,将多帧光流图与目标帧进行空间变换对齐,提高目标帧图像分辨率。在重建图像的同时,应用残差网络学习图像高频部分,恢复更多图像细节,为目标识别及监测打下良好的数据基础。最后对比实验表明,本文算法具有良好的图像超分辨能力。

2 超分辨算法介绍

本文将卷积神经网络应用于运动估计与多帧图像超分辨重建的算法中。首先利用插值法将多帧图像放大至目标倍数,然后将其输入至卷积神经网络中,采用金字塔光流法^[15]对多帧图像进行运动估计,目的是充分利用相邻帧之间的信息,对海上运动物体进行多尺度的运动估计,以恢复出高分辨率图像。利用空间变换网络^[16]对多帧图像进行对齐操作,结合相邻帧与光流估计的结果,完成图像重建。由于海上图像特征信息不易获取,本文采用残差网络^[17]学习目标帧的高频部分,目的是恢复更多的图像高频信息即细节信息,实现最终的多帧图像超分辨率重建。图1所示为本文网络结构示意图,其中 t 表示第 t 帧图像, N 表示时间半径,LR表示低分

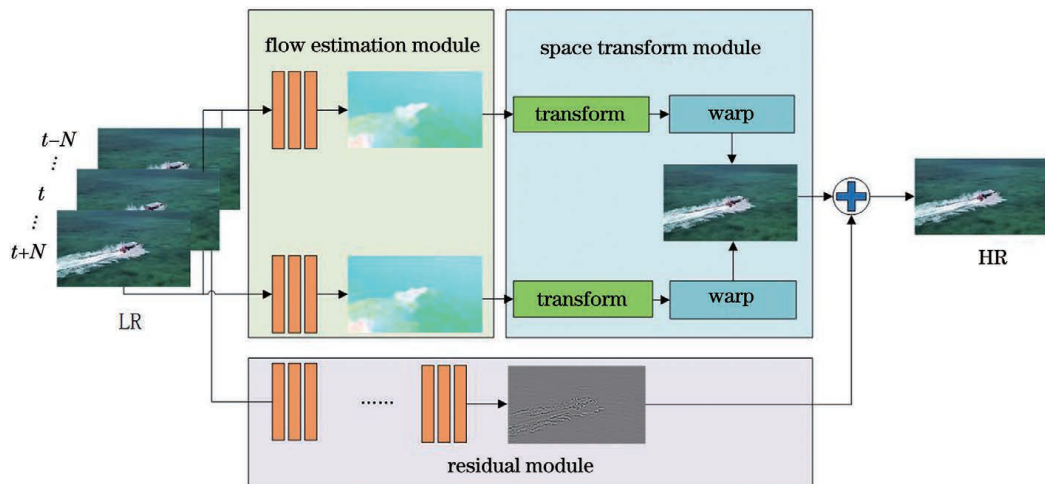


图1 所提网络结构示意图

Fig. 1 Structural diagram of proposed network

分辨率图像,HR 表示高分辨率图像。

2.1 光流估计

光流(optical flow)是指空间运动物体在观察成像平面上的像素的运动瞬时速度。光流法则是利用图像序列中的像素在时间域上的变化以及相邻帧之

间的相关性来找到上一帧与当前帧之间的对应关系,从而计算出相邻帧之间物体的运动信息。如图 2 所示,为了减少计算量并进行多尺度运动估计,本文将采用一种基于卷积神经网络的三层金字塔法进行光流估计。

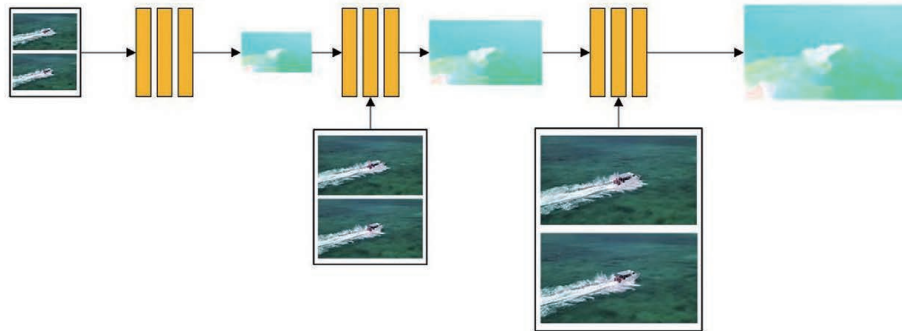


图 2 光流估计示意图

Fig. 2 Schematic of light flow estimation

将两帧图像输入到卷积神经网络中,求解出两帧图像之间的光流场,将光流场上采样图像与输入帧上采样图像进行空间图像扭曲(warp),使其形成新的图像并输入至下一层级。该方法利用每个金字塔层级计算并更新光流,由粗到细地估计物体的运动方向及大小。

如图 3 所示,利用金字塔光流法训练五层神经网络 G_0, \dots, G_4 ,每一层卷积层(Conv)后接 Relu 激活层,训练网络采用 7×7 的卷积核,输出的通道数大小分别为 32,64,32,16,2,最后输出双通道光流场。图 4 所示为测试集中部分海面图像的光流图,为了方便表示,使用不同的颜色和亮度表示光流的大小和方向。

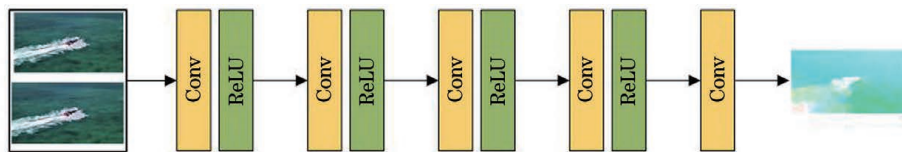


图 3 金字塔层级中的卷积神经网络图

Fig. 3 Map of convolutional neural networks in pyramid hierarchy

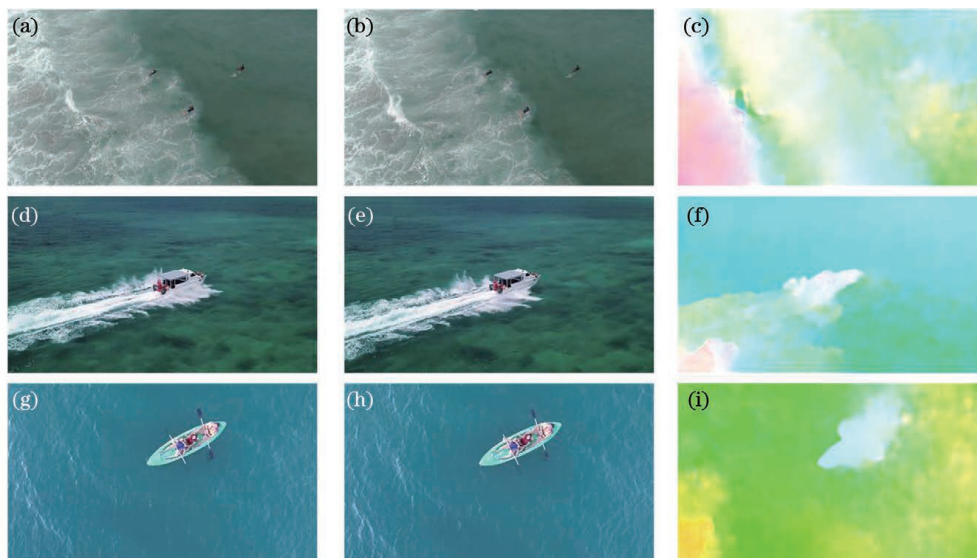


图 4 多帧图像的光流估计。(a)(d)(g)第一帧图像;(b)(e)(h)第二帧图像;(c)(f)(i)光流估计图

Fig. 4 Optical flow estimation of multi-frame images. (a)(d)(g) First image; (b)(e)(h) second image; (c)(f)(i) optical flow estimation map

2.2 多帧图像的空间变换

利用光流估计可以得到多帧图像的预测运动场,图像的空间变换旨在找出输入帧与光流场图像坐标系之间的映射关系,从而获得多个新帧以用于扭曲重建高分辨率图像。由于本文为多帧图像的超分辨重建,输入为 t 帧图像,则需要 $t-1$ 个空间变换模块。图 5 所示为空间变换模块示意图。

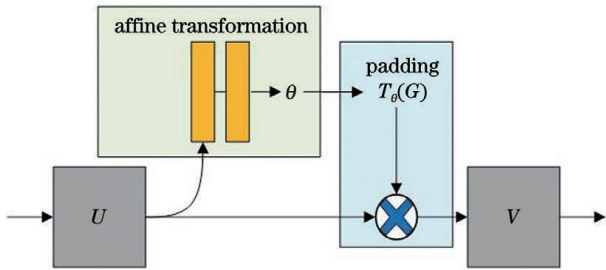


图 5 空间变换模块

Fig. 5 Spatial transformation module

如图 5 所示, U 为输入图像, V 为输出图像, θ 为变化参数, G 为输入至填充模块(padding)的图像。输入与输出图像通过学习得到变换参数 θ , 例如将 U 中的像素 (x, y) 映射到 V 中的 (x', y') , 则仿射变换公式为

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} e \\ f \end{bmatrix}, \quad (1)$$

式中:变换参数 $\theta = a, b, c, d, e, f$, 该 6 个变换参数即为映射输入图与输出图之间坐标点关系的变量。根据预先定义的 V 中的坐标点和变换参数 θ , 计算

出 U 中对应的坐标点, 再将其通过填充规则 $T_\theta(G)$ 函数填充输出图像 V 。

在空间扭曲变换的过程中, 会有一些像素点的坐标不在正整数的网格上, 因此需要通过插值的方法将其变换到网格上。本文采用双线性插值法进行填充。

对于空间变换模块输出的图像, 本文利用 4 个卷积及 ReLU 激活层的网络结构, 输出一幅高分辨率图像。其中网络的输出通道大小分别为 64, 64, 64, 3。

2.3 图像高频分量的学习

由于本文是对海面图像进行超分辨率重建, 海面环境气候多变, 图像信息存在噪声干扰等问题, 因此本文采取残差模块对多帧图像中目标帧的高频分量进行学习, 用来增强图像细节, 减少噪声干扰, 恢复出更高质量图像。

$$I_{HR} = Q + R_t, \quad (2)$$

式中: I_{HR} 表示输出的高分辨率图像; Q 表示多帧低分辨率图像经过光流估计、空间变换及扭曲后得到的图像; R_t 表示残差图像。

如图 6 所示, X_t 表示多帧图像中的低分辨率目标帧图像, F_0 表示 X_t 经过第一个卷积层后得到的浅层特征图, F_1, F_2, \dots, F_k 表示经过 1, 2, \dots, k 层残差模块后输出的特征结果, 本文采用 18 层残差模块, 3×3 卷积层, 激活函数采用 ReLU 函数。最后经过一个 1×1 卷积层后输出残差图像。

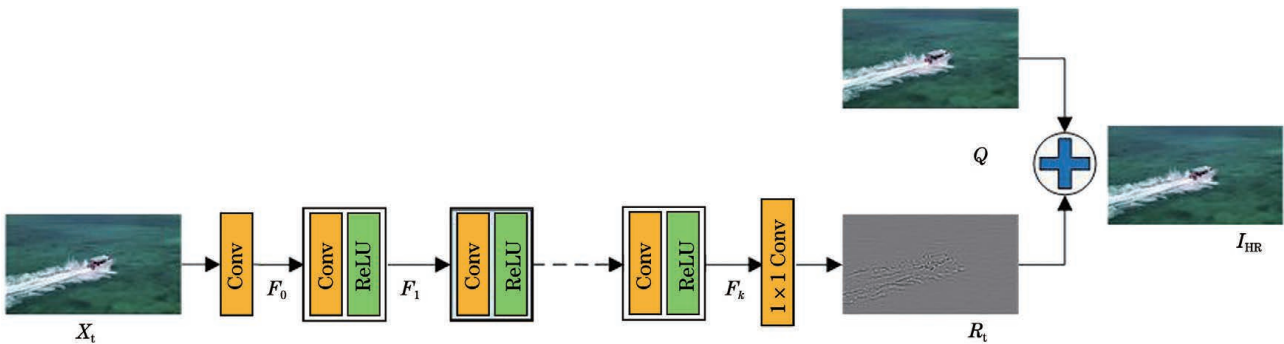


图 6 残差网络结构图

Fig. 6 Structural diagram of residual network

3 实验结果与分析

3.1 数据集

由于现有的数据集不包含海洋类场景的多帧图像, 本文数据集采取 Pexels Videos、Vimeo、Mixkit 网站上下载的免费海洋、沙滩等场景的高

清视频, 经过打散、裁剪等预处理和双三次插值(Bicubic), 多帧图像被放大至目标倍数, 形成用于训练的多帧连续的高分辨率与低分辨率图像对。在数据集中, 训练集为 5000 组多帧图像序列, 测试集为 500 组多帧图像序列, 图 7、8 所示为其中一组数据集图像。



图 7 高分辨率图像序列及其目标帧图像

Fig. 7 High-resolution image sequence and its target frame image



图 8 仿真的低分辨率图像序列及其目标帧图像

Fig. 8 Simulated low-resolution image sequence and its target frame image

3.2 实验结果

本文网络采用 Adam 优化器优化, 设置训练周期(epoch)为 15, 图像大小为 448 pixel×256 pixel。训练所使用的计算机图像处理器 (Graphics Processing Unit, GPU) 型号为英伟达 1080ti, 中

央处理器 (central processing unit, CPU) 为 i5 七代处理器, 所用深度学习框架为 pytorch。对本文所提算法、Bicubic 算法、SRCNN 算法、VDSR 算法及 VSR-DUF 算法进行对比实验, 如图 9~12 所示。

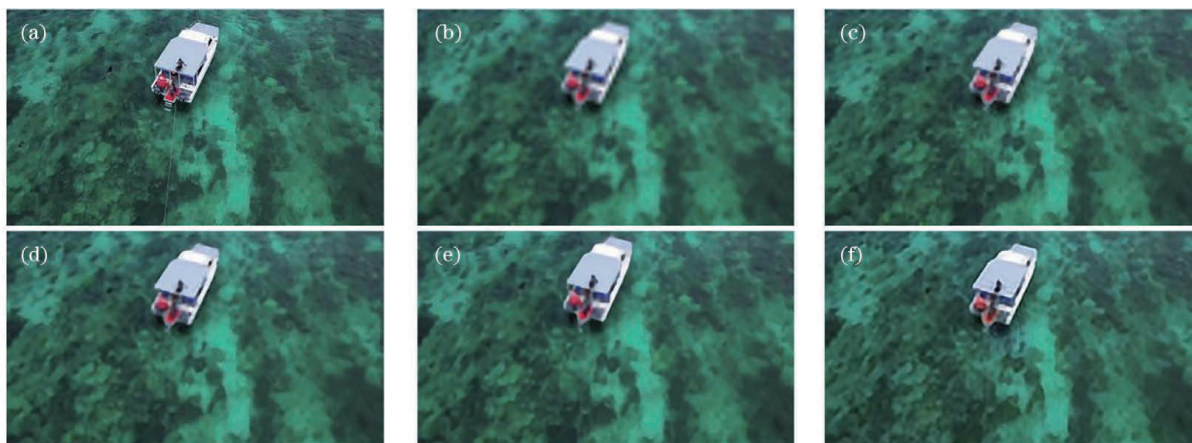


图 9 水上船只图像的对比实验结果图。(a)原图;(b)Bicubic 算法;(c)SRCNN 算法;(d)VDSR 算法;(e)VSR-DUF 算法;(f)所提算法

Fig. 9 Comparison of experimental results of ship images. (a) Original image; (b) Bicubic algorithm; (c) SRCNN algorithm; (d) VDSR algorithm; (e) VDR-DUF algorithm; (f) proposed algorithm

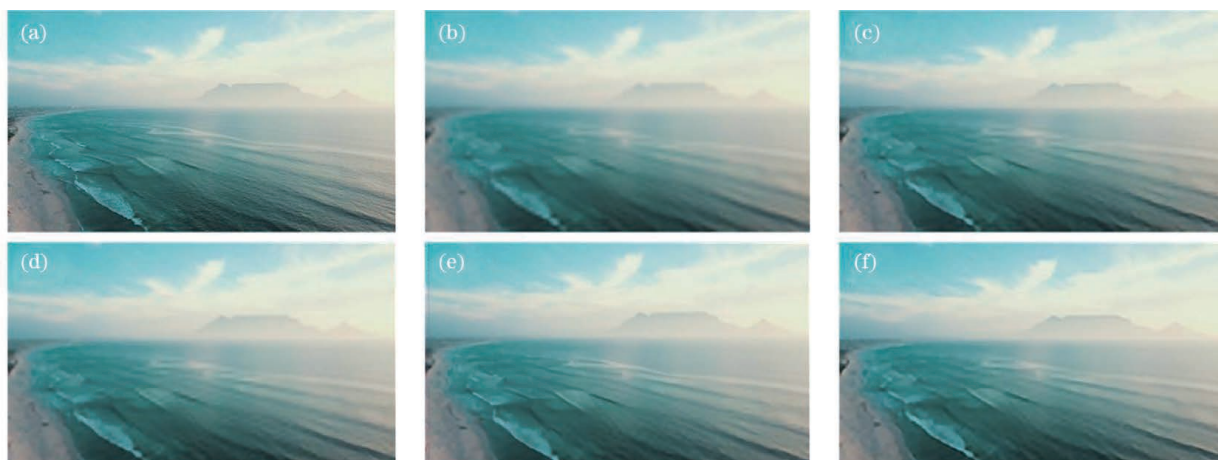


图 10 海滩图像对比实验结果图。(a)原图;(b)Bicubic 算法;(c)SRCNN 算法;(d)VDSR 算法;(e)VSR-DUF 算法;(f)所提算法
Fig. 10 Comparison of experimental results of beach images. (a) Original image; (b) Bicubic algorithm; (c) SRCNN algorithm; (d) VDSR algorithm; (e) VDR-DUF algorithm; (f) proposed algorithm

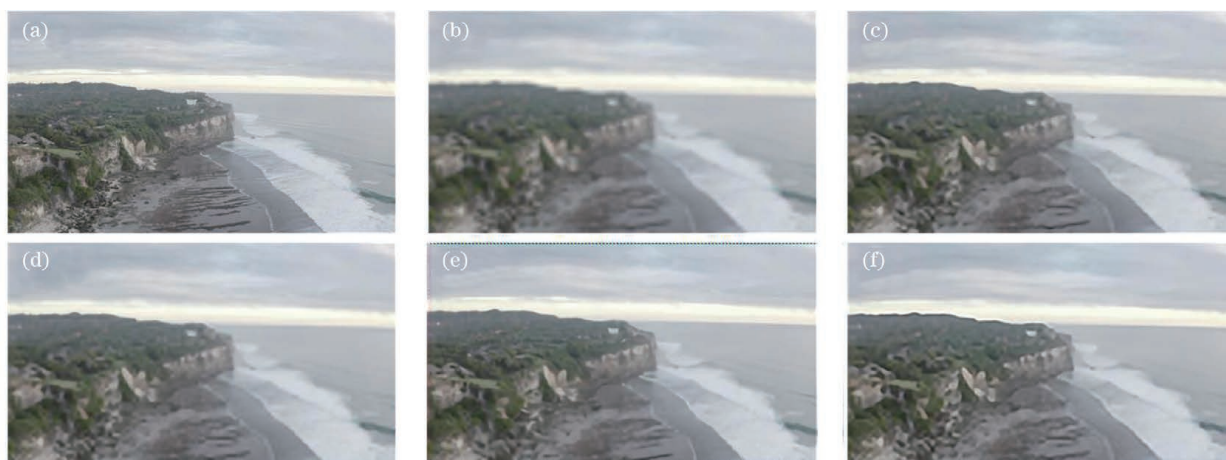


图 11 断崖海岸图像对比实验结果图。(a)原图;(b)Bicubic 算法;(c)SRCNN 算法;(d)VDSR 算法;(e)VSR-DUF 算法;(f)所提算法

Fig. 11 Comparison of experimental results of cliff coast images. (a) Original image; (b) Bicubic algorithm; (c) SRCNN algorithm; (d) VDSR algorithm; (e) VDR-DUF algorithm; (f) proposed algorithm

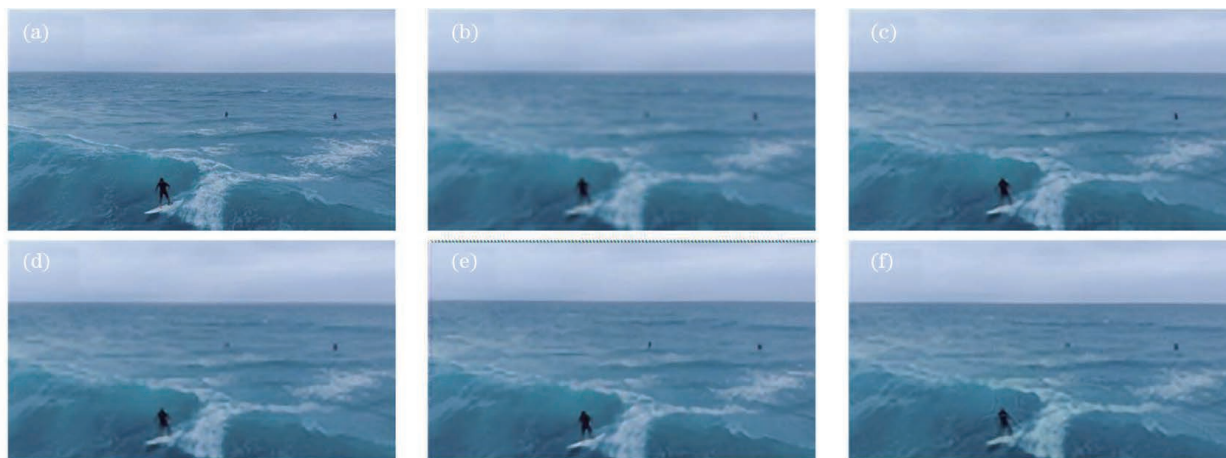


图 12 海面冲浪图像对比实验结果图。(a)原图;(b)Bicubic 算法;(c)SRCNN 算法;(d)VDSR 算法;(e)VSR-DUF 算法;(f)所提算法

Fig. 12 Comparison of experimental results of sea surfing images. (a) Original image; (b) Bicubic algorithm; (c) SRCNN algorithm; (d) VDSR algorithm; (e) VDR-DUF algorithm; (f) proposed algorithm

3.3 主观评价

无论是水面船只的俯视图、海滩图像、断崖图像还是海面冲浪图像,相比于传统算法 Bicubic 算法而言,深度学习对图像的超分辨率重建结果更好。SRCNN 算法及 VDSR 算法的表现良好,但是边缘细节不够丰富,比如海浪边缘、悬崖边缘等地方。VSR-DUF 算法的多帧图像超分辨率重建效果良好,但图像边缘会出现伪影,生成过多虚假像素点。本文所提方法具有较好的多帧图像超分辨重建能力,有效利用了帧与帧之间的时空相关性,并对高频

分量采用残差网络进行学习,因此高频边缘的恢复效果良好,包括船只轮廓、断崖轮廓以及冲浪人物图像,对多种海上图像都具有良好的重建能力。

3.4 客观评价

对于图像超分辨率重建效果的客观评价,大多采用峰值信噪比(PSNR,单位为 dB)及结构相似度(SSIM)进行评价,但客观评价结果并不能代表主观感受,对于一些恢复出更多像素点的超分辨图像,客观评价结果有时不好。表 1 所示为对比实验的客观评价结果。

表 1 各算法的客观评价指标

Table 1 Objective evaluation indexes of each algorithm

Image	Bicubic algorithm		SRCNN algorithm		VDSR algorithm	
	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM
Fig. 8	28.63	0.94	28.88	0.95	29.68	0.74
Fig. 9	30.57	0.92	32.00	0.93	34.05	0.93
Fig. 10	29.33	0.83	30.18	0.85	31.19	0.85
Fig. 11	29.64	0.95	30.73	0.97	31.62	0.98
Image	VSR-DUF algorithm		Proposed algorithm			
	PSNR /dB	SSIM	PSNR /dB	SSIM		
Fig. 8	25.95	0.89	30.04	0.96		
Fig. 9	28.45	0.90	33.77	0.95		
Fig. 10	24.32	0.81	30.79	0.85		
Fig. 11	25.01	0.95	32.78	0.98		

从客观评价的角度来看,因受到伪影及不确定像素点的影响,VSR-DUF 算法的客观评价结果差强人意。SRCNN 及 VDSR 算法的表现良好,相比于传统 Bicubic 算法,峰值信噪比与结构相似度均有提升。可以看出,所提算法利用卷积神经网络得到光流估计与运动补偿图像,帧与帧之间的信息得到了更加充分的利用,因此相较于传统插值算法及单帧图像超分辨算法,重建效果更好;相较于 VSR-DUF 多帧图像超分辨算法,所提算法有效地减少了人工伪影,提高了峰值信噪比及结构相似度。相较于传统插值算法,所提算法重建的低分辨率图像的 PSNR 提升了 5%,SSIM 提升了 3%。由此可见,所提算法具有良好的多帧图像超分辨重建能力。

4 结 论

提出了一种多帧图像超分辨重建算法,并将其应用于海面图像,利用多帧低分辨率图像恢复出一帧高分辨率图像。为了增强图像细节与边缘轮廓,

采用残差网络与光流估计同时处理的方法,对比实验结果显示,所提算法具有良好的图像重建能力。但是所提算法在多水域及纹理细节的重建方面还有待改善,今后的研究重点将考虑扩充数据集,在加强网络深度的同时减少参数及训练时间,以达到最优效果。

参 考 文 献

- [1] Zhou J H, Zhou C, Zhu J J, et al. A method of super-resolution reconstruction for remote sensing image based on non-subsampled contourlet transform [J]. Acta Optica Sinica, 2015, 35(1): 0110001.
周靖鸿,周瑾,朱建军,等.基于非下采样轮廓波变换遥感影像超分辨重建方法[J].光学学报,2015,35(1): 0110001.
- [2] Liu K W, Ma Y, Xiong H X, et al. Medical-image super-resolution reconstruction method based on residual channel attention network [J]. Laser & Optoelectronics Progress, 2020, 57(2): 021014.

- 刘可文, 马圆, 熊红霞, 等. 基于残差通道注意力网络的医学图像超分辨率重建方法[J]. 激光与光电子学进展, 2020, 57(2): 021014.
- [3] Ur H, Gross D. Improved resolution from subpixel shifted pictures [J]. *CVGIP: Graphical Models and Image Processing*, 1992, 54(2): 181-186.
- [4] Dong C, Loy C C, He K M, et al. Learning a deep convolutional network for image super-resolution [M]//Fleet D, Pajdla T, Schiele B, et al. *Computer vision-ECCV 2014. Lecture notes in computer science*. Cham: Springer, 2014, 8692: 184-199.
- [5] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 105-114.
- [6] Yuan P Y, Zhang Y P. Image super-resolution reconstruction method using dual discriminator based on generative adversarial networks [J]. *Laser & Optoelectronics Progress*, 2019, 56(23): 231010.
袁飘逸, 张亚萍. 双判别器生成对抗网络图像的超分辨率重建方法[J]. 激光与光电子学进展, 2019, 56(23): 231010.
- [7] Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1637-1645.
- [8] Shi Z T, Wang Z R, Wang R, et al. Single image super-resolution based on convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2018, 55(12): 121001.
史紫腾, 王知人, 王瑞, 等. 基于卷积神经网络的单幅图像超分辨[J]. 激光与光电子学进展, 2018, 55(12): 121001.
- [9] Li S M, Lei G Q, Fan R. Depth map super-resolution reconstruction based on convolutional neural networks [J]. *Acta Optica Sinica*, 2017, 37(12): 1210002.
李素梅, 雷国庆, 范如. 基于卷积神经网络的深度图超分辨率重建[J]. 光学学报, 2017, 37(12): 1210002.
- [10] Sun C, Lü J W, Li J W, et al. Method of rapid image super-resolution based on deconvolution [J]. *Acta Optica Sinica*, 2017, 37(12): 1210004.
孙超, 吕俊伟, 李健伟, 等. 基于去卷积的快速图像超分辨率方法 [J]. 光学学报, 2017, 37(12): 1210004.
- [11] Jo Y, Oh S W, Kang J, et al. Deep video super-resolution network using dynamic upsampling filters without explicit motion compensation [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: 3224-3232.
- [12] Tao X, Gao H Y, Liao R J, et al. Detail-revealing deep video super-resolution [C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 4482-4490.
- [13] Caballero J, Ledig C, Aitken A, et al. Real-time video super-resolution with spatio-temporal networks and motion compensation [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2848-2857.
- [14] Xue T F, Chen B A, Wu J J, et al. Video enhancement with task-oriented flow [J]. *International Journal of Computer Vision*, 2019, 127(8): 1106-1125.
- [15] Ranjan A, Black M J. Optical flow estimation using a spatial pyramid network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2720-2729.
- [16] Jaderberg M, Simonyan K, Zisserman A. Spatial transformer networks [C]//Advances in Neural Information Processing Systems (NIPS), December 11-12, 2015, Canada, Montreal. Canada: NIPS, 2015: 2017-2025.
- [17] He K M, Zhang X Y, Ren S Q, et al. Identity mappings in deep residual networks [M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9908: 630-645.