

# 图像多尺度密集网络去模糊模型

宋昊泽, 吴小俊\*

江南大学物联网工程学院, 江苏 无锡 214122

**摘要** 使用基于深度学习的端到端去模糊方法, 将模糊图像编码后再解码成高清图像。针对编码过程中网络模型存在提取特征信息不足, 导致重建的去模糊图像质量下降的问题, 提出两种网络结构改进方法: 在自编码网络中添加密集网络结构以提高网络提取特征信息的能力; 引入多尺度感受野结构, 该结构由 4 个尺度的平均池化层和上采样层组成, 从而提取更多输入图像的上下文特征信息。在 GOPRO 数据集和 Kohler 数据集, 两种网络改进方法均取得了较好的图像重建效果。

**关键词** 图像处理; 图像去模糊; 多尺度结构; 平均池化层; 上采样层

中图分类号 TP301.6

文献标识码 A

doi: 10.3788/LOP56.211001

## Deblurring Model of Image Multi-Scale Dense Network

Song Haoze, Wu Xiaojun\*

School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China

**Abstract** This study uses an end-to-end method for image deblurring based on deep learning to encode the blurred image and to subsequently decode it into a high-definition image. However, the lack of extracted feature information during encoding decreases the quality of the reconstructed deblurred image. To solve this problem, we propose two methods for improving the network structure. First, a dense network structure is added to the autoencoder network for extracting considerable feature information. Second, a multiscale perceptual field structure is introduced to extract considerable contextual feature information, comprising 4 scales of average pooling layers and up-sampling layers. The two improved methods achieve good image deblurring effects using the GOPRO and Kohler datasets.

**Key words** image processing; image deblurring; multi-scale structure; average pooling layer; up-sampling layer

**OCIS codes** 100.3010; 100.2960; 150.1135

## 1 引言

近些年在人工智能和模式识别领域, 图像的去模糊方法得到了越来越多的关注<sup>[1-7]</sup>。经典的方法<sup>[4-5, 8-9]</sup>主要基于已知模糊核的模糊计算公式, 这些方法使用模糊核进行反卷积操作, 进而计算出清晰图像。但在这一过程中很难寻找一个模糊核来满足所有像素点。在此基础上通过判断模糊核种类来搭建去模糊模型<sup>[10]</sup>, 该方法的效果有所提升, 但是算法耗时太长。

随着深度学习的广泛应用, 卷积神经网络<sup>[11]</sup>已被用于图像去模糊<sup>[12-14]</sup>, Chakrabarti<sup>[15]</sup>使用卷积神

经网络预测模糊核, 训练卷积神经网络时输入已知模糊核生成的模糊图像, 期望输出的模糊核尽量与已知的相对应的模糊核相近, 测试时通过卷积神经网络获得模糊图像模糊核, 再使用反卷积方法计算出去模糊图像。文献[16]中使用全连接网络(FCN)预测运动轨迹, 并以此获得去模糊图像。以上模型只将深度学习技术方法作为体系的一部分, Nah等<sup>[17]</sup>和Noroozi等<sup>[18]</sup>使用卷积神经网络端到端地直接生成去模糊图像, 这种方法在测试时, 将模糊图像输入到网络模型中, 输出即为去模糊图像。基于深度学习的端到端模型可用来解决未知运动类型的模糊问题, 这样的模糊往往是由相机的抖动或

收稿日期: 2019-03-15; 修回日期: 2019-04-11; 录用日期: 2019-04-30

基金项目: 国家自然科学基金(61672265, 61373055)

\* E-mail: wu\_xiaojun@jiangnan.edu.cn

拍摄对象的移动造成的。为进一步提升效果, Kupyn 等<sup>[19]</sup>提出了 DeblurGAN 去模糊模型, 这一模型使用 Wasserstein 距离对抗生成网络(WGAN)的改进版(WGAN-GP)<sup>[20-21]</sup>, 使重建的去模糊图像具有更多纹理信息。

DeblurGAN 模型为提取到更多输入图像的特征信息, 其生成器中含有残差网络结构(ResNets)<sup>[22]</sup>。而在提取特征能力方面, 文献[23]中提出一种新型卷积网络——密集网络, 它是残差网络的改进结构, 相比残差网络, 密集网络对参数的使用效率更高且抗过拟合能力更强, 在密集网络中每一层都可以更容易地对之前所有层的特征进行重复利用。文献[24-25]将密集网络以密集网络卷积块(Dense block)的形式应用在网络模型中, 完成图像语义分割和图像去雨任务, 并取得了更好的效果。

最近在网络结构上的创新工作也很多, 文献[26]提出一种可以增加网络感受野的模型——多尺度感受野模型。这种多尺度感受野结构使用不同窗口大小的平均池化层, 同时并行处理图像, 由此通过更多的感受野提取图像更多的上下文特征信息, 将多尺度感受野结构嵌入到网络中进行图像语义分割, 取得了更好的分割效果。随后 Zhang 等<sup>[27]</sup>提出多尺度去雾模型, 在网络中引入多尺度结构进行图像去雾, 同样取得了良好的效果。而图像的上下文信息对于图像去模糊也同样重要。

为进一步提升去模糊效果, 本文对网络结构做出更大的改进。在 DeblurGAN 模型的基础上加入密集网络结构和多尺度感受野结构, 简称 MSDN。目的是通过使用密集网络提取到更多的图像特征, 并通过多尺度结构提取到更多图像局部上下文信息的特征, 由此生成更清晰的去模糊图像。

## 2 基础工作及密集网络

### 2.1 DeblurGAN 方法及其损失函数

DeblurGAN 是目前去模糊效果最好的模型之一, 它是一种基于对抗生成网络的去模糊方法。网络架构由生成器和判别器 2 部分组成。判别器的作用是使生成器生成的去模糊图像更贴近真实的高清图像。生成器的作用是将模糊图像转换为清晰图像, 为提取更多输入图像的特征信息, 生成器中加入残差网络结构。DeblurGAN 训练过程中生成器损失函数为

$$L_{\text{total}} = L_{\text{adv}} + \gamma L_{\text{content}} \quad (1)$$

(1)式中使用  $L_{\text{content}}$  保留图像的语义信息, 在此基础上使用  $L_{\text{adv}}$  保留图像的细节信息,  $L_{\text{adv}}$  起辅助作用, 在  $L_{\text{total}}$  中占比小。参考 DeblurGAN 模型<sup>[19]</sup>中的参数设置,  $\gamma$  具体取值为 100。将生成的去模糊图像和高清图像输入到预训练的 VGG-19 网络中, 得到的输出特征图为

$$L_{\text{content}} = \frac{1}{W_{i,j} H_{i,j}} \sum_{a=1}^{W_{i,j}} \sum_{b=1}^{H_{i,j}} \oint_{i,j} (\mathbf{x})_{a,b} - \oint_{i,j} (\bar{\mathbf{x}})_{a,b}, \quad (2)$$

式中,  $W_{i,j}$  和  $H_{i,j}$  表示 VGG 输出特征图的维度,  $\bar{\mathbf{x}}$  表示清晰图像,  $\mathbf{x}$  表示生成的去模糊图像,  $\oint_{i,j}$  表示在 VGG-19 网络里第  $i$  个卷积块第  $j$  个卷积层的输出,  $a, b$  分别为输出的特征图的长和宽。此处输出 VGG-19 网络中第 5 个卷积块中第 4 个卷积层。

使用的  $L_{\text{adv}}$  来自于 WGAN-GP<sup>[21]</sup>。  $L_{\text{adv}}$  的计算公式为

$$L_{\text{adv}} = -D[G(\mathbf{x})], \quad (3)$$

式中,  $D$  代表对抗生成网络中的判别器,  $G$  代表生成器。WGAN-GP 网络更加稳定, 但是由于惩罚采样的原因, 在网络中无法使用批次归一化标准化法(BN)。

### 2.2 密集网络

在密集网络卷积块中, 每一层的输入都与所有之前层的输出有关。  $\mathbf{x}_i$  代表块结构中第  $i$  个卷积层的输出, 在经典的卷积神经网络中,  $\mathbf{x}_i$  由一个非线性操作  $T_i$  计算得到。  $T_i$  操作通常代表一个卷积计算后接一个 ReLU 激活函数,  $\mathbf{x}_i$  表示为

$$\mathbf{x}_i = T_i(\mathbf{x}_{i-1}). \quad (4)$$

密集网络<sup>[23]</sup>将当前层之前所有层输出的特征在特征图的通道上进行累积后作为当前层输入。则在密集网络中当前层  $\mathbf{x}_i$  的输出为

$$\mathbf{x}_i = T_i([\mathbf{x}_{i-1}, \mathbf{x}_{i-2}, \dots, \mathbf{x}_0]), \quad (5)$$

式中,  $[\cdot]$  表示在特征图通道上的连接操作。密集网络卷积块中第  $i$  层的输出有  $j$  个特征图。在  $i$  层后  $[\mathbf{x}_{i-1}, \mathbf{x}_{i-2}, \dots, \mathbf{x}_0]$  连接成  $i \times j$  个特征图, 此连接方式加强了对之前层输出的特征信息的重复使用。

## 3 网络结构及训练

多尺度密集网络模型使用卷积神经网络端对端地将模糊图像转换生成去模糊图像, 在 DeblurGAN 工作的基础上, 对网络结构进行改进。训练时, 损失函数为 DeblurGAN 方法提出的组合损失函数。训

练时使用的数据为 GOPRO 数据集提供的图片对, 即一张模糊的图片和对应的一张高清图片。模型测试时只使用生成器, 输入一张模糊图像, 输出一张去模糊图像。这种使用卷积神经网络将图像进行编码-解码得到新图像的方法目前在图像处理领域应用广泛, 不仅应用于去模糊领域, 还应用于图像超分辨率<sup>[28]</sup>、图像风格转换<sup>[29-30]</sup>、3D 图像重建<sup>[31]</sup>等领域。

### 3.1 密集网络结构

图 1 为残差网络卷积块 (Res block) 和 dense block 示意图。Dense block 对网络有更强的表示能力, 每一层对之前所有层输出的特征信息的重复利用能力更强。最初的密集网络用于解决图像分类的问题, 文献[16]中说明密集网络的网络结构可以提取更多输入图像的特征信息, 使很深的神经网络可以更容易地在训练的时候学习图像分类损失函数。这样的能力也适用于本文图像到图像的转换方法。去模糊方法的思想为通过编码-解码操作将模糊图像转换成去模糊图像, 则在此网络中应用密集网络可以提取到更多模糊图像的特征信息, 从而通过更多的特征信息, 解码器生成质量更高的去模糊图像。

图 1 中, 在 Res block 中, 符号 C 代表加操作, 在 dense block 中, 符号 C 代表连接操作。由于使用密集网络的网络层数比较多, 在这种非常深的网络里, 为使网络更加稳定, 损失函数需更快地收敛。同时使用 WGAN-GP 损失函数和批次归一化方法使用<sup>[32]</sup>, 在最初的密集网络卷积块<sup>[16]</sup>中每个密集网络层均使用 BN 方法。但 Arpit 等<sup>[33]</sup>指出 BN 会

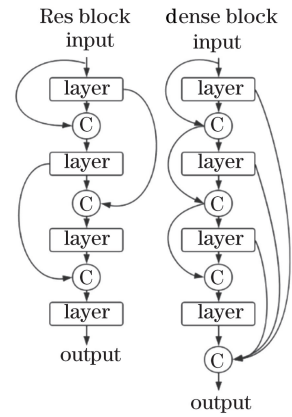


图 1 Res block 和 Dense block 结构比较  
Fig. 1 Comparison of structures of Res block and Dense block

因为最小批次的选取规则引入误差, 并提出权重归一化 (WN), 通过分离网络权重的方向和大小对网络的权重参数进行标准化, 并在文献[33]中论证 WN 不依赖于网络输入批次的大小, 相比于 BN 对网络的优化速度更快, 这正适用于 WGAN-GP。随后 Xiang 等<sup>[34]</sup>提出一种衡量 GANs 生成图像质量的方法并证明在图像生成领域使用 WN 相比于 BN 在生成图像中具有更多细节信息。所以这里对密集网络卷积块中每层使用文献[31]提出的标准化方法, 对每一卷积层进行权重归一化并后接 TPRReLU 激活函数。密集网络卷积块中每一个密集网络卷积层的结构如图 2 中 dense layer 所示, 结构为一个核 (kernel) 为 3、步长 (stride) 为 1 的卷积层 (convolutional layer), 并后接权重归一化操作 (weight normalization)、TPReLU 激活函数、dropout 层<sup>[35]</sup>。

dense layer	down layer	up layer
convolutional layer (kernel size: 3, stride: 1)	convolutional layer (kernel size: 1, stride: 1)	transposed convolutional layer (kernel size: 3, stride: 2)
weight normalization	weight normalization	weight normalization
TPReLU	TPReLU	TPReLU
dropout: 0.2	dropout: 0.2	
	max pooling layer (size: 2)	

图 2 层结构图  
Fig. 2 Structure of layers

### 3.2 多尺度结构

在神经网络里, 网络的感受野与卷积层的卷积核大小或池化层窗口的尺寸直接相关。从理论上讲, 文献[36]中残差网络的感受野尺寸应该比输

入图像的尺寸要大, 但是从经验角度上分析, Deng 等<sup>[37]</sup>提出在卷积神经网络上使用的感受野应远小于理论值, 但这也会造成很多网络不能充分利用图像上下文语义信息的结果。尤其在神经网络





考文献[24],本文添加速连接(skip connection)的处理方法减少了损失,如图3中①②③④所示,速连接表示缩小层的输出直接与对应的放大层的输出在颜色通道上连接。

### 3.4 训练策略

用Pytorch框架<sup>[43]</sup>编写代码完成所有对比实验。训练参数设置参考文献[19],实验代码修改自文献[19]。生成器和判别器的初始学习率设置为 $10^{-4}$ ,200个周期后学习率随周期的增长线性减小到 $10^{-8}$ 。训练周期设置为300。训练生成器和判别器的优化器应用Adam<sup>[44]</sup>优化器。训练使用的数据为成对的数据,即模糊图像和对应的真实的高清图像。将输入图像尺寸调整(resize)到256 pixel×256 pixel,图像格式为RGB,模型输出图像尺寸为256 pixel×256 pixel。生成器和判别器(图3)分开训练。生成器训练5次,判别器训练1次。训练判别器时,固定生成器参数,将模糊图像输入到生成器中,生成器输出去模糊图像,将真实的高清图像和去模糊图像再分别输入到判别器中训练判别器,期望判别器判断去模糊图像为假,判断真实的高清图像为真,由此期望获得一个辨别“真假”能力强的判别器。训练生成器时,固定判别器参数,将模糊图像输入到生成器中,生成器输出为去模糊图像,再将去模糊图像输入到判别器中

表1 GOPRO数据集的实验对比结果

Table 1 Comparison of experimental results with GOPRO dataset

Method	Ref. [9]	Ref. [14]	Ref. [19]-wild	Ref. [19]	DN	MSDN
PSNR	23.54	24.60	26.82	28.31	28.42	28.64
SSIM	0.8222	0.8420	0.8940	0.8880	0.8920	0.8950

在GOPRO数据集上,深度网络模型具有充分的训练数据。相比于文献[9]、文献[14]中两种经典去模糊方法,文献[19]中方法、DN和MSDN均使用端到端的深度去模糊网络模型,PSNR和SSIM提升显著。对比实验中,文献[19]中方法在训练时使用的数据更多,但本文实验结果依然比文献[19]中方法有所提高。相比于使用同等数据集训练的文献[19]-wild方法,MSDN在PSNR和SSIM上提高显著,其中密集网络结构使网络特性提升明显,多尺度结构进一步提升网络模型性能,改进后生成的图像纹理信息更加丰富,更与真实的高清图像相近。

### 4.2 Kohler数据集

Kohler数据集<sup>[46]</sup>中含有4张高清图像,每张高清图像对应12张不同模糊类型的模糊图像。在此

期望其被判别器判定为真,由此期望生成器可以生成“真假”难辨的图像。在测试时,只需生成器工作,将模糊图像输入到生成器中,输出即为去模糊图像。

## 4 实验比较与分析

### 4.1 GOPRO数据集

GOPRO数据集<sup>[18]</sup>中共有3114对模糊图像和对应的高清图像。图像分辨率为720 pixel。其训练集由2013对图像组成,测试集由1111对图像构成,测试部分固定为图像最中间256 pixel×256 pixel区域。此数据集模糊图像和高清图像均来自真实世界,模糊主要由相机的移动和物体的移动造成。表1为本文方法和其他对比模型的实验效果对比,对比方法为文献[9]的方法和文献[14]中的深度学习方法。使用的对比指标为峰值信噪比(PSNR)和结构相似性(SSIM)<sup>[45]</sup>,其可对图像重建的整体效果进行评价。实验对比图如图4所示。对比实验中文献[19]中方法的训练数据为GOPRO数据集附加其论文制作数据集。文献[19]-wild为只将GOPRO数据集作为训练数据集的方法。在实验中将本文方法分为2种,一种为只在网络结构中改进添加密集网络结构的DN,另一种为MSDN。

数据集上进行测试,表2为本文方法和其他几种方法的对比。实验对比图如图5所示。

在Kohler数据集上,相比于传统去模糊方法(文献[9]、[14]中方法),文献[19]中方法、DN、MSDN使用端到端的去模糊网络模型得到的去模糊图像SSIM较高,表明模型对图像中模糊的消除处理效果依然显著。PSNR提升并不明显,原因在于基于深度网络的端到端的去模糊方法依赖于训练数据,当测试数据与训练数据相关程度不高时,模型没有此类数据的先验知识,重建的去模糊图像在纹理上较差。在部分数据集中端到端去模糊网络模型相比于传统方法提升显著,但在部分数据集中提升并不明显,训练数据是关键问题,在未来具有海量、准确的训练数的情况下,端到端去模糊网络模型泛化能力应有更大提升。

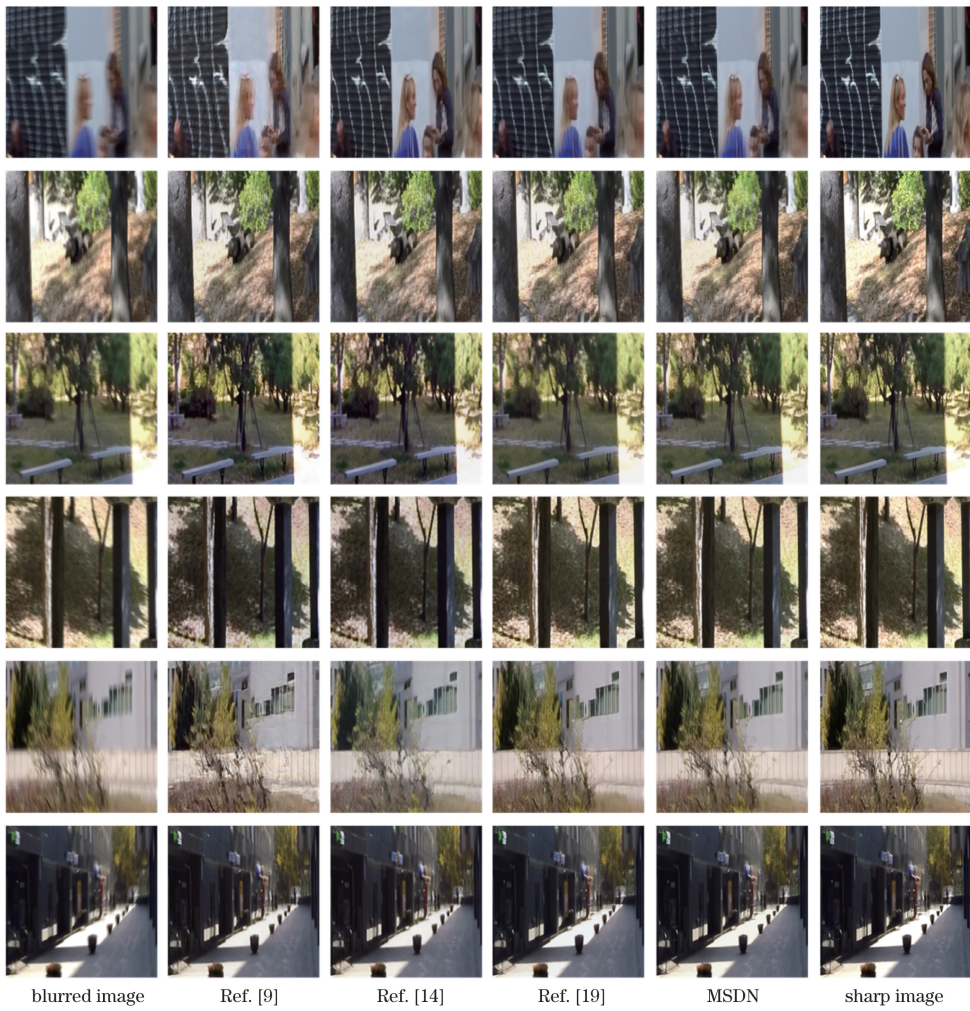


图 4 实验结果对比图

Fig. 4 Comparison of experiment results

表 2 Kohler 数据集的实验对比结果

Table 2 Comparison of experimental results with Kohler dataset

Method	Ref. [9]	Ref. [14]	Ref. [19]-wild	Ref. [19]	DN	MSDN
PSNR	24.62	25.22	26.10	25.35	25.72	26.12
SSIM	0.793	0.773	0.812	0.802	0.814	0.814

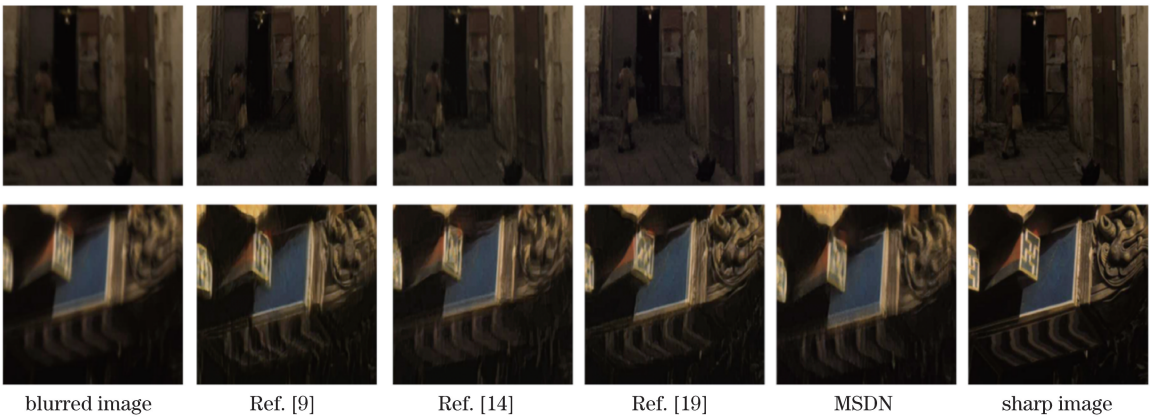


图 5 实验结果对比图

Fig. 5 Comparison of experiment results

## 5 结 论

多尺度密集网络去模糊模型通过对 DeblurGAN 模型网络框架进行改进,使生成的去模糊图像与输入的模糊图像共享更多的特征信息。改进主要从两方面进行:第一方面在网络中使用密集网络,相比于 DeblurGAN 中应用的残差网络,密集网络提取特征的能力更强;第二方面在网络框架中加入多尺度感受野结构,通过四个不同尺度的感受野,提取更多输入模糊图像的上下文信息。经过实验发现,MSDN 实验效果明显优于 DeblurGAN 模型。未来工作中希望进一步地改进模型的网络结构,本文的改进方法主要针对编码过程,之后的工作将对解码过程进行进一步改进,从而重建质量更高的去模糊图像。

## 参 考 文 献

- [1] Chen L S, Jiang Y G. Restoration of blurred image by convolution of arbitrary point spread function[J]. Chinese Journal of Lasers, 1989, 16(9): 541-545.  
陈林森, 姜亚光. 与任意点扩散函数卷积的模糊图像的恢复[J]. 中国激光, 1989, 16(9): 541-545.
- [2] Yang G G, Leith E N. A new image deblurring method[J]. Acta Optica Sinica, 1983, 3(1): 1-9.  
羊国光, Leith E N. 图象消模糊的一种新方法[J]. 光学学报, 1983, 3(1): 1-9.
- [3] Fu Z L, Feng H J, Xu Z H, *et al.* Restoration of image blurred by any form motion in two-dimension [J]. Acta Optica Sinica, 2010, 30(2): 369-372.  
付中梁, 冯华君, 徐之海, 等. 二维任意运动形式模糊图像的恢复[J]. 光学学报, 2010, 30(2): 369-372.
- [4] Whyte O, Sivic J, Zisserman A, *et al.* Non-uniform deblurring for shaken images [J]. International Journal of Computer Vision, 2012, 98(2): 168-186.
- [5] Gupta A, Joshi N, Lawrence Zitnick C, *et al.* Single image deblurring using motion density functions[M]// Daniilidis K, Maragos P, Paragios N. Computer vision-ECCV 2010. Lecture notes in computer science. Berlin, Heidelberg: Springer, 2010, 6311: 171-184.
- [6] Yang A P, Wang J B, Yang B W, *et al.* Joint deep denoising prior for image blind deblurring[J]. Acta Optica Sinica, 2018, 38(10): 1010003.  
杨爱萍, 王金斌, 杨炳旺, 等. 联合深度去噪先验图像盲去模糊[J]. 光学学报, 2018, 38(10): 1010003.
- [7] Hirsch M, Schuler C J, Harmeling S, *et al.* Fast removal of non-uniform camera shake [C] // 2011 International Conference on Computer Vision, November 6-13, 2011, Barcelona, Spain. New York: IEEE, 2011: 463-470.
- [8] Alam M S. Deblurring using fringe-adjusted joint transform correlation[J]. Optical Engineering, 1998, 37(2): 556-564.
- [9] Hyun Kim T, Mu Lee K. Segmentation-free dynamic scene deblurring [C] // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23-28, 2014, Columbus, Ohio. New York: IEEE, 2014: 2766-2773.
- [10] Fergus R, Singh B, Hertzmann A, *et al.* Removing camera shake from a single photograph [J]. ACM Transactions on Graphics, 2006, 25(3): 787-794.
- [11] Razavian A S, Azizpour H, Sullivan J, *et al.* CNN features off-the-shelf: an astounding baseline for recognition[C] // The IEEE conference on computer vision and pattern recognition workshops (CVPR), June 23-28, 2014, Columbus, Ohio. New York: IEEE, 2014: 806-813.
- [12] Xu L, Ren J S J, Liu C, *et al.* Deep convolutional neural network for image deconvolution [C] // Proceedings of the 27th International Conference on Neural Information Processing Systems, December 8-13, 2014, MA, USA. Canada: NIPS, 2014: 1790-1798.
- [13] Schuler C J, Hirsch M, Harmeling S, *et al.* Learning to deblur[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(7): 1439-1451.
- [14] Sun J, Cao W F, Xu Z B, *et al.* Learning a convolutional neural network for non-uniform motion blur removal [C] // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, Massachusetts, USA. New York: IEEE, 2015: 769-777.
- [15] Chakrabarti A. A neural approach to blind motion deblurring[M] // Leibe B, Matas J, Sebe N, *et al.* Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9907: 221-235.
- [16] Gong D, Yang J, Liu L Q, *et al.* From motion blur to motion flow: a deep learning solution for removing heterogeneous motion blur [C] // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, Hawaii. New York: IEEE, 2017: 2319-2328.



- [17] Nah S, Hyun Kim T, Mu Lee K. Deep multi-scale convolutional neural network for dynamic scene deblurring[C] // The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), July 21-26, 2017, Honolulu, Hawaii. New York: IEEE, 2017: 3883-3891.
- [18] Noroozi M, Chandramouli P, Favaro P. Motion deblurring in the wild [M] // Roth V, Vetter T. German conference on pattern recognition. GCPR 2017. Lecture notes in computer science. Cham: Springer, 2017, 10496: 65-77.
- [19] Kupyn O, Budzan V, Mykhailych M, *et al.* DeblurGAN: blind motion deblurring using conditional adversarial networks [C] // The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 8183-8192.
- [20] Arjovsky M, Chintala S, Bottou L. Wasserstein GAN[J/OL]. (2017-12-06) [2018-12-04]. <https://arxiv.org/abs/1701.07875>.
- [21] Gulrajani I, Ahmed F, Arjovsky M, *et al.* Improved training of wasserstein GANs [C] // Advances in Neural Information Processing Systems 30 (NIPS 2017), December 4-9, 2017, Long Beach, Philippines. Canada: NIPS, 2017: 5767-5777.
- [22] Szegedy C, Ioffe S, Vanhoucke V, *et al.* Inception-v4, inception-RESNET and the impact of residual connections on learning [C] // Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17), February 4-9, 2017, San Francisco, California, USA. USA: AAAI, 2017: 4278-4284.
- [23] Huang G, Liu Z, van der Maaten L, *et al.* Densely connected convolutional networks [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2261-2269.
- [24] Jégou S, Drozdal M, Vazquez D, *et al.* The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation[C]//The IEEE Conference on Computer Vision and Pattern Recognition Workshops(CVPR), July 21-26, 2017, Honolulu, Hawaii. New York: IEEE, 2017: 11-19.
- [25] Zhang H, Patel V M. Density-aware single image de-raining using a multi-stream dense network[C]//The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 695-704.
- [26] Zhao H S, Shi J P, Qi X J, *et al.* Pyramid scene parsing network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6230-6239.
- [27] Zhang H, Patel V M. Densely connected pyramid dehazing network [C] // The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 3194-3203.
- [28] Chen Z M, Tong Y G. Face super-resolution through wasserstein GANs [J/OL]. (2018-05-06) [2018-12-04]. <https://arxiv.org/abs/1705.02438>.
- [29] Choi Y, Choi M, Kim M, *et al.* StarGAN: unified generative adversarial networks for multi-domain image-to-image translation [C] // The IEEE Conference on Computer Vision and Pattern Recognition(CVPR), June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 8789-8797.
- [30] Huang Z W, Kratzwald B, Paudel D P, *et al.* Face translation between images and videos using identity-aware cycleGAN [J/OL]. (2017-12-04) [2018-12-04]. <https://arxiv.org/abs/1712.00971>.
- [31] Smith E, Meger D. Improved adversarial systems for 3D object generation and reconstruction [J/OL]. (2017-10-30) [2018-12-04]. <https://arxiv.org/abs/1707.09557>.
- [32] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[J/OL]. (2015-03-02) [2018-12-04]. <https://arxiv.org/abs/1502.03167>.
- [33] Arpit D, Zhou Y B, Kota B U, *et al.* Normalization propagation: a parametric technique for removing internal covariate shift in deep networks [J/OL]. (2016-07-12) [2018-12-04]. <https://arxiv.org/abs/1603.01431>.
- [34] Xiang S T, Li H. On the effects of batch and weight normalization in generative adversarial networks [J/OL]. (2017-12-04) [2018-12-04]. <https://arxiv.org/abs/1704.03971>.
- [35] Gupta A, Duggal R. P-TELU: parametric Tan Hyperbolic Linear Unit activation for deep neural networks[C] // 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 974-978.
- [36] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern



- Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [37] Deng J, Dong W, Socher R, *et al.* ImageNet: a large-scale hierarchical image database [C] // 2009 IEEE Conference on Computer Vision and Pattern Recognition, June 20-25, 2009, Miami, FL, USA. New York: IEEE, 2009: 248-255.
- [38] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// Proceedings of the 25th International Conference on Neural Information Processing Systems, December 3-6, 2012, Lake Tahoe, Nevada. Canada: NIPS, 2012: 1097-1105.
- [39] Ou T S, Huang Y H, Chen H H. SSIM-based perceptual rate control for video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(5): 682-691.
- [40] He K M, Zhang X Y, Ren S Q, *et al.* Delving deep into rectifiers: surpassing human-level performance on ImageNet classification [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1026-1034.
- [41] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [J/OL]. (2016-01-07) [2018-12-04]. <https://arxiv.org/abs/1511.06434>.
- [42] Tao X, Gao H Y, Shen X Y, *et al.* Scale-recurrent network for deep image deblurring [C] // The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22, 2018, Salt Lake City, Utah. New York: IEEE, 2018: 8174-8182.
- [43] PyTorch[EB/OL]. [2018-12-04]. <http://pytorch.org>.
- [44] Kingma D P, Ba J. Adam: a method for stochastic optimization [J/OL]. (2017-01-30) [2018-12-04]. <https://arxiv.org/abs/1412.6980>.
- [45] Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM [C] // 2010 20th International Conference on Pattern Recognition, August 23-26, 2010, Istanbul, Turkey. New York: IEEE, 2010: 2366-2369.
- [46] Köhler R, Hirsch M, Mohler B, *et al.* Recording and playback of camera shake: benchmarking blind deconvolution with a real-world database [M] // Fitzgibbon A, Lazebnik S, Perona P, *et al.* Computer vision-ECCV 2012. Lecture notes in computer science. Berlin, Heidelberg: Springer, 2012, 7578: 27-40.