



基于边缘指导的双通道卷积神经网络 单图像超分辨率算法

摘要

当前基于卷积神经网络(CNN)的超分辨率(SR)重建算法,虽然取得了很大的成功,但是重建图像高频纹理的效果仍然不能令人满意,其高分辨率(HR)图像局部边缘存在明显的震荡.本文提出一种结合形态学成分分析(MCA)分解的边缘指导双通道CNNSR算法:待处理的低分辨率(LR)图像通过MCA分解为纹理部分和平滑结构部分;纹理部分和原LR图像共同组成双通道,输入到改进的网络结构中重建HR纹理部分;结合HR纹理输出与LR平滑结构部分重建HR图像.训练过程采用最小化纹理损失与原图像损失之和最优化网络模型参数.后处理包括:执行网络输出与LR输入图像的直方图匹配使色调保持一致,提升感官效果;应用迭代的反向映射使HR重建与LR输入保持退化算子一致性提高PSNR值.实验结果显示:该方法能够很好地恢复HR图像的纹理细节,对纹理细节丰富的图像恢复效果更好.

关键词

超分辨率;卷积神经网络;形态学成分分析;双通道输入

中图分类号 TP391.41

文献标志码 A

收稿日期 2017-08-28

资助项目 国家自然科学基金(61372184);北京市自然科学基金(4162056)

作者简介

李春平,女,硕士生,研究方向为图像处理、深度学习.chunpingli113@163.com

周登文(通信作者),男,教授,研究方向为图像处理、深度学习.d.w.zhou@163.com

0 引言

单图像超分辨率(SR, Super-Resolution)技术是指通过软件方式把一幅低分辨率(LR, Low-Resolution)图像恢复成高分辨率(HR, High-Resolution)图像的过程.该技术具有广泛的应用,如视频监控、医学成像、遥感卫星成像等.现有的超分辨率算法可分为3类:基于插值^[1]、基于重建^[2]和基于学习的方法^[3-10].其中基于学习的图像SR算法,学习LR和HR图像之间的函数映射关系,生成最终的HR图像,具有更好的复原效果,是近年来研究的热点.

Chang等^[4]引入局部线性嵌入法(LLE, Locally Linear Embedding)计算重建图像加权平均的权值.假设HR图像块和对应的LR图像块在几何上具有相似性,通过LLE算法计算出一组最优的线性组合系数,使得样本库中的 K 个最近邻样本LR块经过这组系数加权平均得到的图像与输入LR图像块之间的误差最小;再将这组系数直接应用于 K 个样本中的HR图像块,从而得到HR图像.Yang等^[6-7]对LR和HR样本图像块构成的样本库进行稀疏表示,并通过联合训练的方式找到LR和HR图像块对应的过完备字典.这种联合训练的方式迫使对应的LR块和HR块通过各自的字典所获得的稀疏系数相同,并由此建立起LR与HR之间的桥梁.基于稀疏字典的图像SR算法^[7]建立稀疏先验约束,在稀疏编码过程自动选择参与SR重建的字典原子数量,而非LLE中人为的设定.这类方法较LLE方法重建质量更高,但是稀疏编码和重建过程需要多次迭代,算法复杂度较大.Jing等^[8]改进了Yang等^[6-7]的算法,分别重建图像高频纹理部分和平滑部分,纹理部分采用稀疏表示的方法重构,平滑部分简单地用插值的方法恢复,然后组合HR图像,增强了图像纹理细节.

以上基于学习的SR算法主要分为特征提取、学习和重建3个阶段.各阶段分别独立设计优化,且学习模型的特征提取与表达能力有限.近年来,深度学习引起了广泛的关注,它的出现弥补了传统基于学习方法的不足.其中,Dong等^[9]率先将卷积神经网络(CNN, Convolutional Neural Network)引入到图像SR复原问题中,设计了基于深度CNN的图像SR复原方法(SRCNN).该方法主要思想是:以深度学习与传统的稀疏编码之间的关系为依据,将网络分为图像块提取、非线性

¹ 华北电力大学 控制与计算机工程学院,北京,102206

性映射和重建 3 个阶段,再将这 3 个阶段统一到一个深度 CNN 的框架中,学习 LR 图像到 HR 图像之间的端到端映射关系.网络结构上,将稀疏编码过程看作卷积操作,利用 3 层卷积层完成图像块提取、非线性映射和重建操作.优化方式上,传统的基于学习的 SR 方法对 3 个阶段分别独立优化处理,忽略了整体框架的优化.SRCNN 采用联合优化方式,图像重建性能有了较大的提高.SRCNN 模型证明了直接学习 LR 和 HR 之间端到端的映射简单可行,效果也很好,但重构出的高频特征仍然不能令人满意.

受文献[8-9]等方法的启发,本文提出一种结合形态学成分分析(MCA, Morphological Component Analysis)^[11]分解的边缘指导双通道 CNN 算法.首先,利用 MCA 算法,将待处理的 LR 图像分解为边缘纹理部分和平滑结构部分;边缘纹理部分和原 LR 图像共同组成双通道,输入到 CNN 中得到 HR 纹理图像输出;再将 HR 纹理输出图像与原 LR 图像合并,得到重建的 HR 图像;最后通过最小化纹理损失和原图像损失之和和训练网络结构参数.实验结果表明,本文提出的边缘指导双通道模型超分辨率算法在结果图像中能够保留更多高频细节信息,同时也可以很好地减弱振铃现象.实验的评价指标 PSNR 值也证明了本文模型算法的有效性.

1 本文算法

本章首先介绍如何提取边缘纹理信息构建双通道输入网络结构;然后,进一步解释如何将边缘先验信息嵌入到深度 CNN 中,以便更好地预测图像 SR 的高频细节.

1.1 MCA 图像分解算法

MCA^[11]分解是基于稀疏表示的图像分解方法,通过稀疏表示和全变分(TV, Total Variation)方法将图像分解为不同形态的空间特征.输入图像 X 假设总共包含 R 个像素, X 是 2 个不同部分的线性组合^[12]:纹理部分 X_t 和结构部分 X_n ,则有 $X = X_t + X_n$.

为了分离包含图像纹理部分的 X_t 和包含图像结构部分的 X_n ,把每一部分用一个给定的联合字典稀疏表示,即 $X_t = T_t \alpha_t$, $X_n = T_n \alpha_n$,其中 $T_t, T_n \in M^{R \times L}$. MCA 分解的目的是求解以下最优化问题:

$$\begin{aligned} \{\alpha_k, \alpha_n\} &= \arg \min_{\{\alpha_t, \alpha_n\}} \|\alpha_k\|_0 + \|\alpha_n\|_0, \\ \text{s.t. } X_t &= T_t \alpha_t + T_n \alpha_n. \end{aligned} \quad (1)$$

由于式(1)是一个 NP 难问题不容易求解,可以用 l_1 代替 l_0 ,得到一个凸优化问题:

$$\{\alpha_k, \alpha_n\} = \arg \min_{\{\alpha_t, \alpha_n\}} \|\alpha_k\|_1 + \|\alpha_n\|_1,$$

$$\text{s.t. } \|X - T_t \alpha_t - T_n \alpha_n\| \leq \xi, \quad (2)$$

式(2)中,参数 ξ 表示图像噪声约束.由于 TV 模型能够有效地复原分段平滑图像的显著边缘,因此在图像平滑成分分解中增加一个 TV 约束项,如式(3)所示.惩罚项 $\gamma TV\{T_n \alpha_n\}$ 使得分解的平滑图有一个稀疏的梯度,更接近分段平滑图像.

$$\begin{aligned} \{\alpha_k, \alpha_n\} &= \arg \min_{\{\alpha_t, \alpha_n\}} \|\alpha_k\|_1 + \|\alpha_n\|_1 + \lambda \|X - \\ &T_t \alpha_t - T_n \alpha_n\|_2^2 + \gamma TV\{T_n \alpha_n\}. \end{aligned} \quad (3)$$

对图像进行 MCA 分解时,需要采用已知的变换矩阵 T_k 和 T_n .经过实验本文选用基于局部的 DCT 变换分解图像的纹理成分,基于曲波变换分解图像的结构成分.但由于 MCA 分解是有损的,直接使用 $T_n \alpha_n$ 进行 SR 重建得到的 HR 图片效果不佳,本文使用 LR 图像与其纹理部分的差值 $Y - T_t \alpha_t$ 代替分解后的 LR 平滑结构图片 N_t ,用于 SR 重建.

算法基本流程描述如下:

步骤 1:初始化.

阈值因子 $L_{\max} = 255$,参数 $\lambda = 1, \gamma = 0.8$,迭代次数 $N = 30$,令 $X_n = X, X_k = 0, \gamma = \lambda L_{\max}$,执行 N 次迭代.

步骤 2:固定 X_k ,更新 X_n .

计算冗余误差: $R = X - X_n - X_k$;计算 $X_n + R$ 的小波变换: $\alpha_n = T_n^+(X_n + R)$;使用阈值 δ 进行阈值处理,得到 $\hat{\alpha}_n$,通过 $\hat{\alpha}_n$ 重构 $X_n, X_n = T_n \hat{\alpha}_n$.

步骤 3:固定 X_n ,更新 X_k .

计算冗余误差: $R = X - X_n - X_k$;计算 $X_n + R$ 的局部 DCT 变换: $\alpha_k = T_k^+(X_n + R)$;使用阈值 δ 进行阈值处理,得到 $\hat{\alpha}_k$,通过 $\hat{\alpha}_k$ 重构 $X_k, X_k = T_k \hat{\alpha}_k$.

步骤 4:求解 TV 约束项,更新阈值 δ .

通过以下公式对 TV 约束项调整:

$$\begin{aligned} X_n &= X_n - \mu \frac{\partial TV\{X_n\}}{\partial X_n} = \\ &X_n - \mu \nabla \cdot \left(\frac{\nabla X_n}{|\nabla X_n|} \right), \end{aligned} \quad (4)$$

$$\delta = \delta - \lambda/N. \quad (5)$$

步骤 5:判断停止条件是否满足:如果 $\delta > \lambda$,则返回步骤 2,否则,结束.

通过以上算法模型得到图像的结构部分和纹理部分如图 1 所示.

1.2 边缘指导的双通道输入 CNN 结构

本文提出一种端到端可训练的边缘指导双通道输入卷积神经网络算法.该算法拓宽了输入图像的



a. Butterfly原始图像 b. Butterfly结构部分 c. Butterfly纹理部分

图1 图像 Butterfly 的 MCA 分解

Fig. 1 Decomposition of image Butterfly based on MCA,

with (a) for original image, (b) for structure and (c) for texture

维度,图2是提出的边缘指导双通道输入网络结构,输入是提取的LR图像纹理部分和原LR图像组成的图像块.算法包括以下几个部分:

1) LR边缘纹理提取.在1.1节详细提到利用基于稀疏表示和全变分的MCA方法提取LR图像的边缘纹理信息,这也可以推广到建模其他自然图像先验信息上.这里表述如何在深度网络中嵌入边缘信息,具体地,将LR图像中提取的边缘纹理特征与原LR图像共同作为双通道输入,旨在预测HR输出图像的高频纹理特征,然后将其用于复原HR图像.

网络的输入特征 F_{input} 是原LR图像 Y 与纹理特征 X'_r 的级联:

$$F_{input} = [Y, T_l]. \quad (6)$$

2) 边缘指导双通道输入深度网络.该网络结构旨在学习高、低分辨率纹理图像之间的映射函数.网络由 n 层组成:第1层块的提取和表示层,从输入图像中紧密的裁剪小块,并表示成一个特征向量,如图2采用64个大小为 3×3 的滤波器提取特征,相当于滤波器作用在2个输入通道的 3×3 空间区域(特征映射)上,该特征提取层是一个线性操作,激活函数ReLU对特征提取层输出的所有特征映射作非线性处理,得到特征图(或图像块)第1层的操作如下所示:

$$F_{T_1}(Y) = \max(0, W_1 \times F_{input} + b_1), \quad (7)$$

最后一层重构层用单一的滤波器大小为 $3 \times 3 \times 64$ 将HR边缘纹理小块拼接成一幅完整的图像:

$$F_T(Y) = \max(0, W_n \times F_{T_{n-1}}(Y) + b_{n-1}). \quad (8)$$

中间层由若干个特征映射层组成,不同层可以逐步恢复不同频域的HR图像特征,即把不同频域的LR图像小块的低维特征向量表示成高维特征向量,最终组合成一幅HR纹理输出,其中每一层相当于大小为 $3 \times 3 \times 64$ 滤波器作用在64个通道的 3×3 的空间区域上.表示如下:

$$F_{T_{mid}}(Y) = \max(0, W_{mid} \times F_{T_{mid-1}}(Y) + b_{mid}). \quad (9)$$

3) HR图像预测.因为LR图像包含重要的LR细节,基于边缘指导的双通道网络主要恢复高频成分,HR图像重建过程依然需要LR图像细节的辅助,因此最终HR图像输出是用 $Y - T_l \alpha_l$ 替代后的LR平滑结构图片 N_l 与HR纹理图像输出的和表示,即 $\hat{X} = N_l + \hat{T}_l$.

4) 训练.重构损失采用HR图像和HR边缘纹理损失联合训练边缘指导的双通道输入网络结构,把高频纹理特征的恢复应用到网络训练的误差上,使恢复图像更接近真实的HR图像.

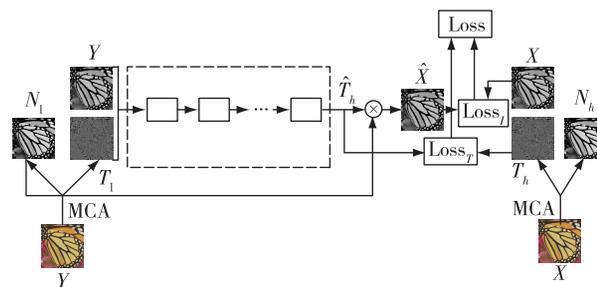


图2 基于边缘指导的双通道输入卷积神经网络结构

Fig. 2 Structure of edge guided two-dual input convolutional neural network

1.3 网络训练

本文依旧采用最小化欧式距离来优化模型参数得到最终的模型.给定一组数据集 $\{X_i, Y_i\}_{i=1}^n$, X_i 表示一组真实的HR图像, Y_i 表示的是一组LR图像.本文的学习模型, \hat{X} 是目标图像HR图像的估计,因此,只要使得均值误差最小即可.值得一提的是双通道输入结构中从LR图像中提取纹理信息 $\{T_l^i\}$ 、平滑结构信息 $\{N_l^i\}$,从HR图像中提取纹理信息 $\{T_h^i\}$ 、平滑信息 $\{N_h^i\}$,则新的LR训练样本为 $Y'_i = \{T_l^i, Y_i\}$,相对应的真实HR图像为 $X'_i = \{X_i\}$,损失函数为:

$$L(\Theta') = \frac{1}{n'} \sum_{i=1}^{n'} (\|F(Y'_i; \Theta') - X'_i\|^2 + \lambda \|F_T(Y'_i; \Theta') - T_h^i\|^2), \quad (10)$$

其中 $\Theta' = \{W'_1, W'_2, W'_3, B'_1, B'_2, B'_3\}$, $F(\cdot)$ 和 $F_T(\cdot)$ 分别表示由输入的LR图像和LR纹理图像得到HR图像和HR纹理图像的映射函数, λ 是平衡图像损失与纹理损失的权重参数,经验值设为1.这里把每个训练样本分为边缘纹理部分和平滑结构部分,因此 n' 等于 n 为训练样本的数目.

2 实验结果及分析

为了验证本文提出的基于双通道CNN SR算法

的有效性,LR 图像先利用双三次插值(Bicubic)得到初始 HR 大小的图像,然后用 A+算法^[13]、ScSR 算法^[7]、SCN 算法^[14]、SRCNN 算法^[9]以及本文的算法 SR 重建得到的结果进行对比.实验测试使用的仿真软件为:Matlab R2016a,实验的环境为 GeForce GTX TITAN GPU,32 GB 内存,Ubuntu 操作系统.图像效果用主观效果和客观效果评价,测试图像选用 Set5、Set14 和 BSD200.

1)模型细节:实验过程中选取 SRCNN 实验中使用的 91 幅图像,采取高斯核模糊($\delta=1.2$),并双三次插值下采样得到训练集.随机裁剪大小为 33×33 的小块,采用零均值,标准方差 0.001(偏置项为 0)的高斯分布初始化每一层的权重,权重衰减项 0,动力 0.9,采用 Adam 的策略,无 dropout.作为后处理步骤,执行网络的图像输出与 LR 图像输入的直方图匹配提升感官效果;应用迭代的反向映射使 HR 图像重建与 LR 图像输入保持退化算子一致性,提高图像复原质量.

2)主观效果:限于篇幅,仅给出纹理丰富、对本文算法非常敏感的测试图像 Butterfly 的视觉结果.如图 3,是一个 3 倍放大的 Butterfly,LR 图像是以因子为 3 下采样原 HR 图像得到的.可以看到:由于双三次插值方法没有考虑沿边缘纹理插值,蝴蝶翅膀较大的内部支架纹理不丰富、偏模糊;ScSR, A+, SRCNN 方法中蝴蝶的翅膀较大的内部支架和本文方法一样较清晰,但由于没有考虑高频纹理的特征提取,翅膀的细小边缘部分恢复比较模糊;本文的方法提取了更多的高频纹理特征用于 SR 重建,并逐一恢复不同频率的高频细节,蝴蝶翅膀的细小边缘细节恢复更加清晰.因此从主观上来看,本文算法效果整体优于对比算法.

3)客观效果:采用峰值信噪比(PSNR)进行对比.PSNR 对重建图像和原始图像的误差进行定量计算,值越大复原的效果越好.计算式(11)所示:

$$R_{\text{PSN}} = 10 \times \log_{10} \left(\frac{255^2}{E_{\text{MS}}} \right), \quad (11)$$

$$E_{\text{MS}} = \frac{1}{HW} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} \| I(i,j) - K(i,j) \|^2, \quad (12)$$

I 和 K 分别表示原 HR 图像和重构的 HR 图像,尺寸为 $H \times W$.从表 1、表 2、表 3 中得出:不同放大因子,本文方法的 PSNR 均值都高于对比方法,当放大倍数为 3 时在 Set5 上平均 PSNR 值比 SRCNN 和 SCN 方法高出约 0.4 dB.对于所有的测试图像,本文的方法

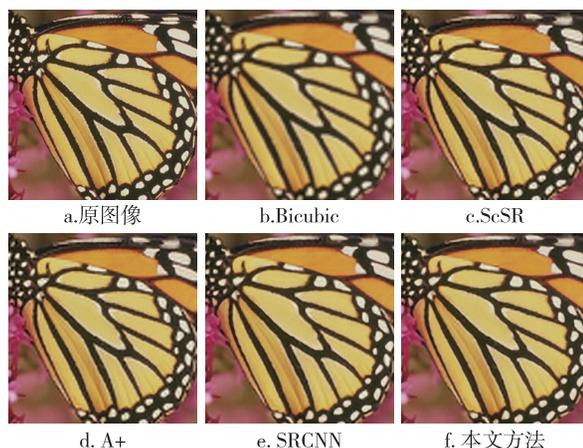


图 3 3 倍放大的 Butterfly 图像

Fig. 3 Parts of the 3×upscale image Butterfly reconstructed by different algorithms, with (a) for original image and (f) for the proposed method

都胜出了其他方法.对于显著边和丰富纹理的 Butterfly 本文超出其他方法近 1 dB.

表 1 不同算法在 Set5 上的均值 PSNR 比较

Table 1 Comparison of PSNR among different algorithms for 3×upscale on Set5

方法	上采样倍数			dB
	2	3	4	
Bicubic	33.66	30.39	28.42	
A+	36.56	32.60	30.30	
ScSR	36.04	31.78	29.48	
SRCNN	36.66	32.75	30.49	
SCN	36.76	33.04	30.82	
本文方法	37.04	33.19	30.93	

表 2 不同算法在 Set14 上的均值 PSNR 比较

Table 2 Comparison of PSNR among different algorithms for 3×upscale on Set14

方法	上采样倍数			dB
	2	3	4	
Bicubic	30.23	27.54	26.00	
A+	32.14	29.07	27.28	
ScSR	31.70	28.31	26.50	
SRCNN	32.44	29.30	27.50	
SCN	32.48	29.37	27.62	
本文方法	32.63	29.52	27.69	

3 结束语

本文提出了结合 MCA 分解的边缘指导双通道

表3 不同算法在 BSD200 上的均值 PSNR 比较

方法	上采样倍数		
	2	3	4
Bicubic	29.70	27.26	25.97
A+	30.78	28.18	26.77
ScSR	30.04	27.15	26.02
SRCNN	31.34	28.27	26.72
SCN	31.40	28.50	27.02
本文方法	31.66	28.61	27.06

卷积神经网络算法,LR 边缘纹理信息从原图像中分离出来作为输入特征的一部分,引导图像的 SR 重建.采用最小化纹理损失与原图像损失之和最优化网络模型参数以确保训练的平稳性,并执行直方图匹配的后处理与 LR 图像保持相同的色调和反差,提高图像的感官效果.实验结果表明,本文算法较好地恢复了一般图像的边缘和细节信息,尤其对细节比较丰富的图像复原效果更加明显.下一步工作是将传统的超分辨率方法与基于深度学习的 SR 方法结合做进一步研究.

参考文献

References

- [1] Zhang L, Wu X L. An edge-guided image interpolation algorithm via directional filtering and data fusion [J]. IEEE Transactions on Image Processing, 2006, 15 (8): 2226-2238
- [2] Rasti P, Demirel H, Anbarjafari G. Image resolution enhancement by using interpolation followed by iterative back projection [C] // IEEE Signal Processing and Communications Applications Conference, 2013, DOI: 10.1109/SIU.2013.6531593
- [3] Freeman W T, Jones T R, Pasztor E C. Example-based super-resolution [J]. IEEE Computer Graphics and Applications, 2002, 22 (2): 56-65
- [4] Chang H, Yeung D Y, Xiong Y M. Super-resolution through neighbor embedding [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, DOI: 10.1109/CVPR.2004.1315043
- [5] Chan T M, Zhang J P, Pu J, et al. Neighbor embedding based super-resolution algorithm through edge detection and feature selection [J]. Pattern Recognition Letters, 2009, 30 (5): 494-502
- [6] Yang J C, Wright J, Huang T S, et al. Image super-resolution as sparse representation of raw image patches [C] // IEEE Conference on Computer Vision and Pattern Recognition, 2008: 1-8
- [7] Yang J C, Wright J, Huang T S, et al. Image super-resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19 (11): 2861-2873
- [8] Jing G D, Shi Y H, Lu B. Single-image super-resolution based on decomposition and sparse representation [C] // International Conference on Multimedia Communications, 2011: 127-130
- [9] Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38 (2): 295-307
- [10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C] // International Conference on Neural Information Processing Systems, 2012: 1097-1105
- [11] Bobin J, Starck J L, Fadili J M, et al. Morphological component analysis: An adaptive thresholding strategy [J]. IEEE Transactions on Image Processing, 2007, 16 (11): 2675-2681
- [12] Starck J L, Elad M, Donoho D L. Image decomposition via the combination of sparse representations and a variational approach [J]. IEEE Transactions on Image Processing, 2005, 14 (10): 1570-1582
- [13] Timofte R, De Smet V, Van Gool L. A+: Adjusted anchored neighborhood regression for fast super-resolution [C] // Asian Conference on Computer Vision, 2014: 111-126
- [14] Wang Z W, Liu D, Yang J C, et al. Deeply improved sparse coding for image super-resolution [J]. arXiv e-print, 2015, arXiv: 1507.08905

Edge guided dual-channel convolutional neural network for single image super resolution algorithm

LI Chunping¹ ZHOU Dengwen¹ JIA Huimiao¹

¹ School of Control and Computer Engineering, North China Electric Power University, Beijing 102206

Abstract At present, although the super-resolution (SR) reconstruction algorithm based on the Convolutional Neural Network (CNN) has achieved great success, it cannot well reconstruct the high-frequency texture of the image. As a result, there exists obvious shake in local edge of the high-resolution (HR) image. We present an edge guided dual-channel CNN SR reconstruction algorithm integrated with Morphological Component Analysis (MCA).

The low-resolution (LR) image to be processed is decomposed into texture part and structure part by MCA, then the texture part and the original LR image form a dual channel together, which is then input into the modified network structure to reconstruct the HR texture part. The reconstruction loss of both the HR image and HR texture are chosen simultaneously for training. As for post-processing step, we perform histogram matching between our network output and the LR input to strengthen the visual effect and apply an iterative back projection refinement to improve the PSNR. As shown in experiment results, this method with dual-channel input can restore texture details of the image, especially restore the image with rich texture.

Key words super resolution; convolutional neural network; morphological component analysis (MCA); dual-channel input