

# 基于多尺度稠密残差网络的JPEG压缩伪迹去除方法

陈书贞 张祎俊 练秋生\*

(燕山大学信息科学与工程学院 秦皇岛 066004)

(河北省信息传输与信号处理重点实验室 秦皇岛 066004)

**摘要:** JPEG在高压缩比的情况下,解压缩后的图像会产生块效应、边缘振荡效应和模糊,严重影响了图像的视觉效果。为了去除JPEG压缩伪迹,该文提出了多尺度稠密残差网络。首先把扩张卷积引入到残差网络的稠密块中,利用不同的扩张因子,使其形成多尺度稠密块;然后采用4个多尺度稠密块将网络设计成包含2条支路的结构,其中后一条支路用于补充前一条支路没有提取到的特征;最后采用残差学习的方法来提高网络的性能。为了提高网络的通用性,采用具有不同压缩质量因子的联合训练方式对网络进行训练,针对不同压缩质量因子训练出一个通用模型。经实验表明,该文方法不仅具有较高的JPEG压缩伪迹去除性能,且具有较强的泛化能力。

**关键词:** JPEG压缩; 压缩伪迹; 多尺度稠密块; 扩张卷积

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2019)10-2479-08

DOI: 10.11999/JEIT180963

## JPEG Compression Artifacts Reduction Algorithm Based on Multi-scale Dense Residual Network

CHEN Shuzhen ZHANG Yijun LIAN Qiusheng

(Institute of Information Science and Technology, Yanshan University, Qinhuangdao 066004, China)

(Hebei Key Laboratory of Information Transmission and Signal Processing, Qinhuangdao 066004, China)

**Abstract:** In the case of high compression rates, the JPEG decompressed image can produce blocking artifacts, ringing effects and blurring, which affect seriously the visual effect of the image. In order to remove JPEG compression artifacts, a multi-scale dense residual network is proposed. Firstly, the proposed network introduces the dilate convolution into a dense block and uses different dilation factors to form multi-scale dense blocks. Then, the proposed network uses four multi-scale dense blocks to design the network into a structure with two branches, and the latter branch is used to supplement the features that are not extracted by the previous branch. Finally, the proposed network uses residual learning to improve network performance. In order to improve the versatility of the network, the network is trained by a joint training method with different compression quality factors, and a general model is trained for different compression quality factors. Experiments demonstrate that the proposed algorithm not only has high JPEG compression artifacts reduction performance, but also has strong generalization ability.

**Key words:** JPEG compression; Compression artifacts; Multi-scale dense blocks; Dilate convolution

### 1 引言

随着信息化技术的不断发展,图像作为一种重要的信息载体,高质量图像的数据量逐渐增加。受

传输带宽和存储能力的限制,图像压缩技术在现代工作生活中越来越重要。目前图像压缩方法一般分为2种:无损压缩和有损压缩。JPEG是一种典型的有损压缩方法,其具有较高的压缩比,已经广泛应用于数码相机和互联网等领域。在高压缩比的情况下, JPEG解压缩图像会产生压缩伪迹。JPEG的压缩比越大,解压缩图像质量越差,越影响人们的视觉体验。因此,去除JPEG压缩伪迹在图像压缩领域具有十分重要的作用。

现有的JPEG压缩伪迹去除算法主要分为3类:基于图像增强的算法,基于图像恢复的算法

收稿日期: 2018-10-15; 改回日期: 2019-03-05; 网络出版: 2019-04-02

\*通信作者: 练秋生 lianqs@ysu.edu.cn

基金项目: 国家自然科学基金(61471313), 河北省自然科学基金(2019203318)

Foundation Items: The National Natural Science Foundation of China (61471313), The Natural Science Foundation of Hebei Province (2019203318)

和基于学习的算法。基于图像增强的算法通常利用像素域滤波和变换域滤波来去除压缩伪迹。如Foi等人<sup>[1]</sup>提出了基于形状自适应离散余弦变换(Shape-Adaptive DCT, SA-DCT)的图像滤波算法。该算法有效地去除了图像中的块效应和边缘振荡效应。Yoo等人<sup>[2]</sup>提出了基于块间相关性的去块效应算法,该算法采用平滑和分组滤波的方式分别去除平坦区域和边缘区域的伪迹。基于图像增强的算法一般计算复杂度较低,但在高压缩比的情况下,去块效果不佳。基于图像恢复的算法通常利用自然图像,压缩算法和压缩参数的先验知识,试图从压缩图像中恢复出高质量的图像。如Zhao等人<sup>[3]</sup>提出了一种在最大后验概率的框架上同时利用结构稀疏表示先验和量化约束先验的图像去块效应算法。该算法利用结构稀疏表示先验来同时加强内在的局部稀疏性和自然图像的非局部自相似性,再结合量化约束先验得到更可靠的估计。基于图像恢复的算法一般需要复杂的迭代运算,计算复杂度较高。

近几年来,基于深度学习的算法已经广泛应用在图像处理领域,如图像恢复<sup>[4,5]</sup>、目标检测<sup>[6,7]</sup>、语义分割<sup>[8]</sup>等,并取得优异的效果。在去除JPEG压缩伪迹中,Dong等人<sup>[9]</sup>最早提出一种去除伪迹的卷积神经网络(Artifacts Reduction Convolutional Neural Networks, ARCNN)。该网络由4个卷积层组成,分别用于特征提取、特征增强、非线性映射和重构。ARCNN的优点是网络结构简单,但复原图像仍存在较多的伪迹。Svoboda等人<sup>[10]</sup>提出了包含8个卷积层和跳跃连接的残差网络。在高压缩比的情况下,该网络的收敛速度快,但复原图像仍存在部分伪迹。Zhang等人<sup>[11]</sup>设计了一个更深的残差网络,并且在网络里加入批量归一化层。该网络利用残差训练,加快了网络的收敛速度,同时解决了因层数过多而出现的梯度消失问题。Kim等人<sup>[12]</sup>利用GoogLeNet的Inception V1<sup>[13]</sup>的思想,用 $7 \times 7$ 滤波器代替池化层,并在不同尺寸滤波器提取多尺度特征之前,使用 $1 \times 1$ 的卷积来减少参数的数量。该网络的恢复图像缺少部分细节信息。Liu等人<sup>[14]</sup>利用小波分解原理,提出一种类似U-Net<sup>[15]</sup>的网络结构。该网络把U-Net框架中的下采样部分用小波变换替代,上采样部分用小波逆变换替代,从而增大了网络的感受野。该网络把图像的高低频信息分开处理,有利于恢复较清晰的图像。上述方法(除Zhang等人<sup>[11]</sup>提出的残差网络)均采用对每个压缩质量因子进行单独训练的方式,使用该方式训练的模型不能有效地恢复与训练样本不同压缩质量因子的压缩图像。

最近,生成对抗网络受到了越来越多的关注。Guo等人<sup>[16]</sup>提出了一对多网络,用于去除JPEG压缩伪迹。该网络首先将辅助变量和压缩图像作为生成器中两个不同分支的输入,然后使用残差单元将两个分支级联在一起,最后得到去伪迹图像。一对多网络使用感知损失、对抗损失和JPEG损失进行优化,通过改变辅助变量,使得恢复的图像包含不同细节信息。Galteri等人<sup>[17]</sup>提出了基于深度生成对抗网络去除JPEG压缩伪迹的算法。该算法利用生成器和判别器相互对抗的原则,通过优化感知损失和对抗损失,使得恢复的图像不仅去除了大部分伪迹还保留了部分细节信息。虽然生成对抗网络能获得不错的性能,但该类网络仍存在不稳定性、训练时间长和训练过程中可能引入其它伪迹等缺陷。

受稠密连接卷积网络<sup>[18]</sup>的启发,利用不同尺度滤波器能提取不同尺度特征的原理,本文在稠密块里分别使用具有不同扩张因子的扩张卷积<sup>[19]</sup>,使其形成多尺度稠密块。然后利用多尺度稠密块,构造出包含2条支路的多尺度稠密残差网络,其中后一条支路用于补充前一条支路没有提取到的特征。为了使网络更具有通用性,本文采用具有不同压缩质量因子的图像对网络进行联合训练,训练好的网络模型可适用于不同压缩质量因子的JPEG压缩图像,即同一模型可恢复由各种压缩质量因子所产生的压缩图像。

## 2 图像JPEG压缩

对图像进行JPEG压缩,首先需要把图像转换成YCbCr颜色空间,然后将图像分成 $8 \times 8$ 的不重叠小块,最后将每一个小块依次进行离散余弦变换(DCT)、量化和编码。每个 $8 \times 8$ 图像块压缩解压过程数学表达式为(本文不考虑熵编解码过程)

$$\mathbf{Y}^{\text{dct}}(i, j) = \text{round}(\mathbf{X}^{\text{dct}}(i, j) / \mathbf{Q}(i, j)) \cdot \mathbf{Q}(i, j) \quad (1)$$

其中, $\mathbf{X}$ 和 $\mathbf{Y}$ 分别为真实图像和解压缩图像, $\mathbf{X}^{\text{dct}}$ 和 $\mathbf{Y}^{\text{dct}}$ 分别为 $\mathbf{X}$ 和 $\mathbf{Y}$ 的DCT系数, $\mathbf{Q}$ 为量化表, $i$ 和 $j$ 为DCT域 $8 \times 8$ 块的索引, $\text{round}(\cdot)$ 表示4舍5入取整。

图像进行块离散余弦变换后,在 $8 \times 8$ 的DCT系数矩阵中,低频信息集中在块的左上角,高频信息集中在块的右下角;然后进行量化,将块离散余弦变换后的系数除以量化表中对应的量化系数之后取整,该量化过程使相邻块边缘产生不连续性从而形成块效应<sup>[10]</sup>。在标准量化表中,低频分量对应较小的量化步长,高频分量对应较大的量化步长,因此在量化过程中损失了部分高频信息。JPEG压缩图像因损失部分高频信息引起的感应振荡,从而形

成了边缘振荡效应和模糊。

### 3 多尺度稠密残差网络

#### 3.1 网络结构

为了去除JPEG压缩伪迹，本文提出了多尺度稠密残差网络，如图1所示。该网络首先将JPEG压缩图像 $x$ 经过2个 $3 \times 3$ 的卷积层，对 $x$ 进行特征提取；然后利用4个多尺度稠密块将网络设计成包含两条支路的结构，其中后一条支路B用于补充前一

条支路A没有提取到的特征；最后采用残差学习的方法得到去除伪迹的图像 $y$

$$y = x + G_A(x; \theta_A) + G_B(x; \theta_B) \quad (2)$$

其中， $G_A(\cdot)$ 表示从JPEG压缩图像 $x$ 到支路A的输出， $G_B(\cdot)$ 表示从JPEG压缩图像 $x$ 到支路B的输出， $\theta_A$ 和 $\theta_B$ 分别表示 $G_A(\cdot)$ 和 $G_B(\cdot)$ 过程中卷积层的超参数。网络中所有卷积层(除每条支路的最后一个 $3 \times 3$ 卷积层)的激活函数都用ReLU<sup>[20]</sup>。

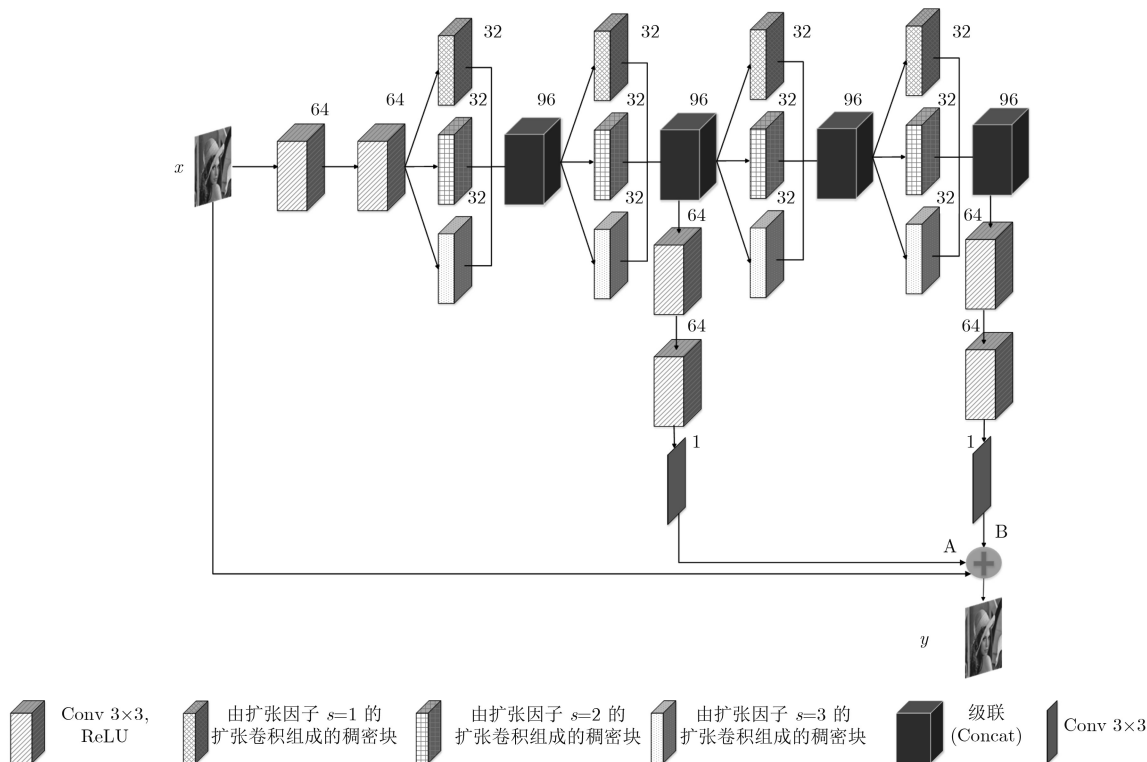


图1 多尺度稠密残差网络

#### 3.2 多尺度稠密块

稠密连接卷积网络<sup>[18]</sup>最早用于分类和目标识别任务。受稠密连接卷积网络的启发，本文采用稠密块作为多尺度结构的分支。本文对5个卷积层进行稠密连接，即对于每一层，它前面所有层的特征映射的级联用作此层的输入，并且此层的特征映射用作后续所有层的输入，如图2所示。图中 $n$ 表示卷积核的个数， $k$ 表示卷积核的大小， $s$ 表示扩张因子。为了防止因特征级联而产生过多的参数，稠密块的第1层设置64个滤波器，其余4层设置32个滤波器。稠密块的每一层都与后面所有层级联，因此可以实现特征重用<sup>[18]</sup>，有利于提高网络的学习效率。

利用不同尺度滤波器能提取不同尺度特征的原理，本文采用3个稠密块来构造多尺度结构。3个稠密块分别用尺寸为 $3 \times 3$ 、 $5 \times 5$ 和 $7 \times 7$ 的滤波器提取特征，小的滤波器可以提取小尺度特征，大的滤

波器可以提取大尺度特征，从而提高网络的表示能力，最后使每个稠密块提取的特征进行级联作为下一层的输入。为了减少网络的计算量，本文选择用扩张卷积来代替普通卷积。如扩张因子 $s = 2$ 时的扩张卷积的卷积核尺寸为 $(2s + 1) \times (2s + 1)$ ，也即卷积核大小为 $5 \times 5$ ，一个卷积核的参数仍为 $3 \times 3$ ，因此扩张卷积可在不增加计算量的同时增大了感受野。使用扩张因子 $s = 1, 2, 3$ 的扩张卷积来代替普

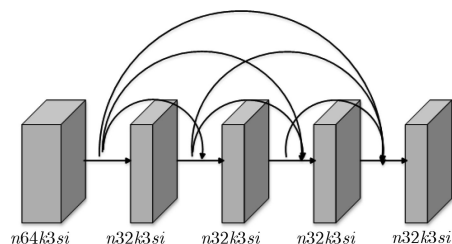


图2 由扩张因子 $s = i$ 的扩张卷积组成的稠密块

通卷积,实现了多尺度的思想,增大了网络的感受野,同时又加快了网络的训练速度。

本文选择  $s = 1, 2, 3$ , 利用3种不同的扩张因子,形成了3种不同尺度的稠密块。 $f_i(\mathbf{x}_n; \theta'_{i,n})$ ,  $i = 1, 2, 3$  分别表示3种不同尺度的稠密块,  $\mathbf{x}_n$  表示稠密块的输入,  $\theta'_{i,n}$  表示第  $i$  个尺度稠密块的超参数。输入的特征经过3个不同尺度稠密块后级联形成了多尺度稠密块

$$F(\mathbf{x}_n; \theta_n) = C\left(f_1(\mathbf{x}_n; \theta'_{1,n}), f_2(\mathbf{x}_n; \theta'_{2,n}), f_3(\mathbf{x}_n; \theta'_{3,n})\right), n = 1, 2, 3, 4 \quad (3)$$

其中,  $\mathbf{x}_n$  和  $F(\mathbf{x}_n; \theta_n)$  分别表示第  $n$  个多尺度稠密块的输入和输出,  $\theta_n = \{\theta'_{1,n}, \theta'_{2,n}, \theta'_{3,n}\}$  表示第  $n$  个多尺度稠密块中卷积层的超参数,  $C(\cdot)$  表示级联操作。

### 3.3 网络中的分支

为了防止因网络层数较少使恢复的图像仍存在部分伪迹,如算法ARCNN<sup>[9]</sup>, L8 Residual<sup>[10]</sup>等,本文选择使用4个多尺度稠密块来构造网络;为了防止因网络层数较多可能会使恢复的图像丢失部分细节信息,如Kim等人<sup>[12]</sup>提出的网络,本文将4个多尺度稠密块设计成包含两条支路的结构,其中后

一条支路用于补充前一条支路没有提取到的特征。网络中的支路A和支路B的输出分别表示为

$$G_A(\mathbf{x}; \theta_A) = Y(F(F(P(\mathbf{x}; \xi_1); \theta_1); \theta_2); \xi_2) \quad (4)$$

$$G_B(\mathbf{x}; \theta_B) = Y(F(F(F(F(P(\mathbf{x}; \xi_1); \theta_1); \theta_2); \theta_3); \theta_4); \xi_3) \quad (5)$$

式(4)和式(5)中,  $G_A(\cdot)$  为从JPEG压缩图像  $\mathbf{x}$  到支路A的输出,  $G_B(\cdot)$  为从JPEG压缩图像  $\mathbf{x}$  到支路B的输出,  $P(\cdot)$  为网络中前2个  $3 \times 3$  卷积层的输出,  $F(\cdot)$  为每个多尺度稠密块的输出,  $Y(\cdot)$  为每一条支路最后3个  $3 \times 3$  卷积层的输出,  $\theta_A = \{\xi_1, \theta_1, \theta_2, \xi_2\}$  为  $G_A(\cdot)$  过程中卷积层的超参数,  $\theta_B = \{\xi_1, \theta_1, \theta_2, \theta_3, \theta_4, \xi_3\}$  为  $G_B(\cdot)$  过程中卷积层的超参数。

多尺度稠密残差网络从输入图像  $\mathbf{x}$  到分支A是对输入的压缩图像进行伪迹提取,此时提取的是图像的伪迹信息和部分因压缩而丢失的细节信息;从  $F(\mathbf{x}_2; \theta_2)$  到分支B可补充从  $F(\mathbf{x}_2; \theta_2)$  到分支A过程中没有提取到的图像伪迹和细节信息;最后,基于残差学习的思想,把分支A和B与输入图像  $\mathbf{x}$  相加,最终得到一个清晰的图像  $\mathbf{y}$ 。网络中每条分支的输出图像,如图3所示,从图中可知,分支B补充了分支A没有提取到的部分伪迹和细节信息。

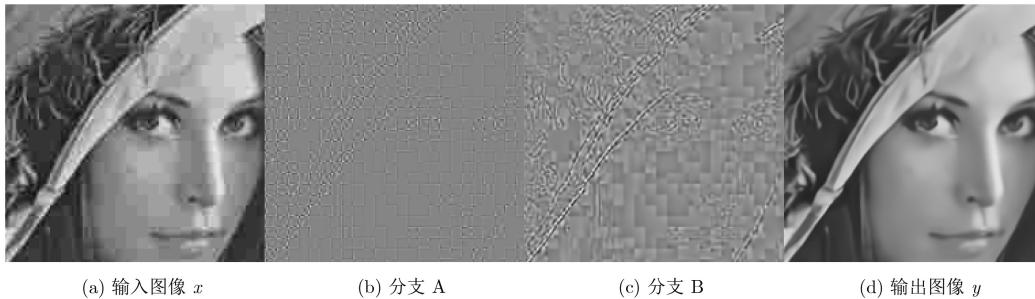


图3 多尺度稠密残差网络中每个分支的输出图像

### 3.4 损失函数

在训练过程中,本文网络使用  $l_2$  范数损失函数

$$L(\theta_A, \theta_B) = \frac{1}{N} \sum_{i=1}^N \|\hat{\mathbf{x}}_i - \mathbf{x}_i - G_A(\mathbf{x}_i; \theta_A) - G_B(\mathbf{x}_i; \theta_B)\|_2^2 \quad (6)$$

其中,  $N$  表示图像样本数,  $\mathbf{x}_i$  为JPEG压缩图像,  $\hat{\mathbf{x}}_i$  为  $\mathbf{x}_i$  对应的压缩前的图像,  $G_A(\cdot)$  表示从JPEG压缩图像  $\mathbf{x}_i$  到支路A的输出,  $G_B(\cdot)$  表示从JPEG压缩图像  $\mathbf{x}_i$  到支路B的输出,  $\theta_A$  和  $\theta_B$  分别表示  $G_A(\cdot)$  和  $G_B(\cdot)$  过程中卷积层的超参数。

## 4 训练和测试

本文实验采用和文献<sup>[9]</sup>相同的训练样本,共

400张图像。从BSDS500数据集的验证图像中随机选取4张作为网络的验证样本。为了更有效地训练网络,400张训练样本需进行数据扩充处理,即对样本进行缩放和旋转。其中,每张图像缩放0.9, 0.8, 0.7和0.6倍,旋转  $90^\circ$ ,  $180^\circ$  和  $270^\circ$ 。400张图像经过数据扩充后,训练样本数增加至3200张。为了进一步增大样本数量,本文对图像进行分块处理。JPEG压缩后的3200张训练样本分成  $51 \times 51$  的图像块,取块步长为40,训练批量数设为64。同样,4张验证样本也分成64个  $51 \times 51$  大小的图像块。测试样本选用LIVE1数据集和Classic5数据集。鉴于人眼对亮度变化较为敏感,因此在实验中的所有样本只考虑亮度通道即Y通道。本文选用压缩质量因子QF为10, 15, 20, 25, 30, 35, 40和45,用MAT-

LAB中的JPEG编码器分别对训练样本、验证样本和测试样本进行压缩，解压后的图像根据不同的压缩质量因子会产生不同强度的块效应。

网络中的超参数包括每层的权重和偏置，所有卷积层的权重选用Xavier方法进行初始化，偏置初始化为0。为了使网络的输入和输出图像大小相同，网络中每个卷积层都进行补零操作。网络使用ADAM算法进行优化，初始学习率设为 $1e^{-4}$ 。当网络训练趋于稳定时，学习率降为 $1e^{-5}$ ，继续训练趋于稳定时，学习率再降为 $1e^{-6}$ ，再继续训练直至收敛。本文使用Tensorflow深度学习框架进行训练，所有实验在Nvidia Tesla K80 GPU平台下完成。

## 5 实验结果

### 5.1 与其它算法比较

为了说明网络结构的性能，本文与4种近年来提出的JPEG压缩伪迹去除算法进行比较，分别为ARCNN<sup>[9]</sup>，L4 Residual<sup>[10]</sup>，L8 Residual<sup>[10]</sup>和DnCNN-3<sup>[11]</sup>。本文算法在LIVE1和Classic5数据集上进行测试，利用峰值信噪比(PSNR)和结构相似度(SSIM)作为评价网络性能的指标。本文中实验对比算法的数据都来自相关文献的网络模型和图像的网址。

目前，去除JPEG压缩伪迹的方法一般采用对于每个压缩质量因子所产生的压缩图像进行单独训练的方式。为了说明单独训练方式的优缺点，本文以ARCNN算法为例。表1中ARCNN(QF = 10, 20, 30或40)表示训练样本分别使用压缩质量因子为10, 20, 30和40时训练好的模型，这4种模型分别对

LIVE1数据集在压缩质量因子为10, 20, 30和40上进行测试。由表1可知，模型ARCNN(QF = 10)恢复QF为40的压缩图像时，平均PSNR比JPEG压缩图像低，说明该模型不能恢复QF为40的压缩图像。其它情况，4种模型分别对QF为10, 20, 30和40的压缩图像进行恢复时，均表现出一定的去除压缩伪迹的能力。由表1的PSNR数据对比中可知，每个模型在恢复与训练样本相同压缩质量因子的压缩图像时，恢复的效果相对最佳，但不能有效地恢复与训练样本不同压缩质量因子的压缩图像。

为了使网络具有通用性和更高的泛化能力，本文采用对于网络进行“盲去块”的联合训练方式，训练一个统一的模型来恢复因各种压缩质量因子所产生的压缩图像。网络采用“盲去块”的联合训练方式时，3200张训练样本分别用QF为10, 20, 30和40进行压缩，压缩后的训练样本数达到12800张。这4种不同程度的压缩图像对网络进行联合训练，训练好的统一模型分别对LIVE1和Classic5数据集在各种压缩质量因子上进行测试。

由表2和表3可知，算法DnCNN-3在LIVE1和

表1 ARCNN的4个模型在LIVE1数据集上的PSNR(dB)对比

模型	QF			
	10	20	30	40
JPEG	27.77	30.07	31.41	32.35
ARCNN(QF = 10)	<b>28.96</b>	30.79	31.51	31.90
ARCNN(QF = 20)	28.78	<b>31.30</b>	32.53	33.30
ARCNN(QF = 30)	28.60	31.25	<b>32.69</b>	33.61
ARCNN(QF = 40)	28.48	31.14	32.62	<b>33.63</b>

表2 本文方法在LIVE1数据集上的PSNR(dB)/SSIM对比

方法	QF			
	10	20	30	40
JPEG	27.77/0.7905	30.07/0.8683	31.41/0.9000	32.35/0.9173
ARCNN	28.96/0.8217	31.30/0.8871	32.69/0.9161	33.63/0.9303
L4 Residual	29.08/0.8241	31.42/0.8900	32.80/0.9174	33.78/0.9322
L8 Residual	-	31.51/0.8911	-	-
DnCNN-3	29.20/0.8262	31.59/0.8936	32.98/ <b>0.9204</b>	33.96/0.9346
本文方法	<b>29.49/0.8329</b>	<b>31.81/0.8952</b>	<b>33.08/0.9196</b>	<b>34.14/0.9367</b>

表3 本文方法在Classic5数据集上的PSNR(dB)/SSIM对比

方法	QF			
	10	20	30	40
JPEG	27.82/0.7800	30.12/0.8541	31.48/0.8844	32.43/0.9011
ARCNN	29.04/0.8108	31.16/0.8691	32.52/0.8963	33.34/0.9098
DnCNN-3	29.40/0.8201	31.63/0.8775	32.90/0.9011	33.77/0.9141
本文方法	<b>29.68/0.8275</b>	<b>31.87/0.8798</b>	<b>33.03/0.9013</b>	<b>33.95/0.9166</b>

Classic5数据集上的PSNR和SSIM的数据明显高于算法ARCNN, L4 Residual和L8 Residual。当QF为10, 20, 30和40时, 本文方法在LIVE1数据集上的平均PSNR分别比算法DnCNN-3高0.29 dB, 0.22 dB, 0.10 dB和0.18 dB; 本文方法在Classic5数据集上的平均PSNR分别比算法DnCNN-3高0.28 dB, 0.24 dB, 0.13 dB和0.18 dB。除了QF为30时, 本文方法在LIVE1数据集上的平均SSIM比算法DnCNN-3低0.0008。其它情况, 本文方法的平均SSIM比算法DnCNN-3高。总体来说, 本文所提出的网络在去除JPEG压缩伪迹的性能上要高于算法ARCNN, L4 Residual, L8 Residual和DnCNN-3。

除了PSNR和SSIM评价指标外, 本文还选取了LIVE1数据集中的sailing3图像, 用于各个算法的视觉比较。由图4可知, 在QF为10时, 本文方法与算法ARCNN, L4 Residual和DnCNN-3在图像上的视觉比较, 可以看出本文方法恢复的图像更清晰。

本文用训练样本中使用的压缩质量因子之外的

压缩质量因子对网络进行测试。如表4所示, 当QF为15, 25, 35和45时, 本文方法在LIVE1数据集上的平均PSNR分别比算法DnCNN-3高0.22 dB, 0.15 dB, 0.15 dB和0.17 dB。本文方法在LIVE1数据集上的平均SSIM与算法DnCNN-3相比, 也有一定程度的提高。因此, “盲去块”的联合训练方式可以使网络具备更高的泛化能力。

## 5.2 不同尺度对网络性能的影响

表5中单一尺度( $3 \times 3$ )表示网络中多尺度稠密块的每个卷积层都采用扩张因子为1的扩张卷积, 单一尺度( $5 \times 5$ )表示网络中多尺度稠密块的每个卷积层都采用扩张因子为2的扩张卷积, 考虑到训练时间的问题, 本文不考虑网络中多尺度稠密块的每个卷积层都采用扩张因子为3的扩张卷积。两个单一尺度网络与本文的多尺度网络进行比较, 如表5所示, 当QF为10, 15, 20和25时, 本文方法在LIVE1数据集上的平均PSNR和SSIM比2种单一尺度的网络高。因此, 为了提高网络的性能, 本文方法采用多尺度结构。

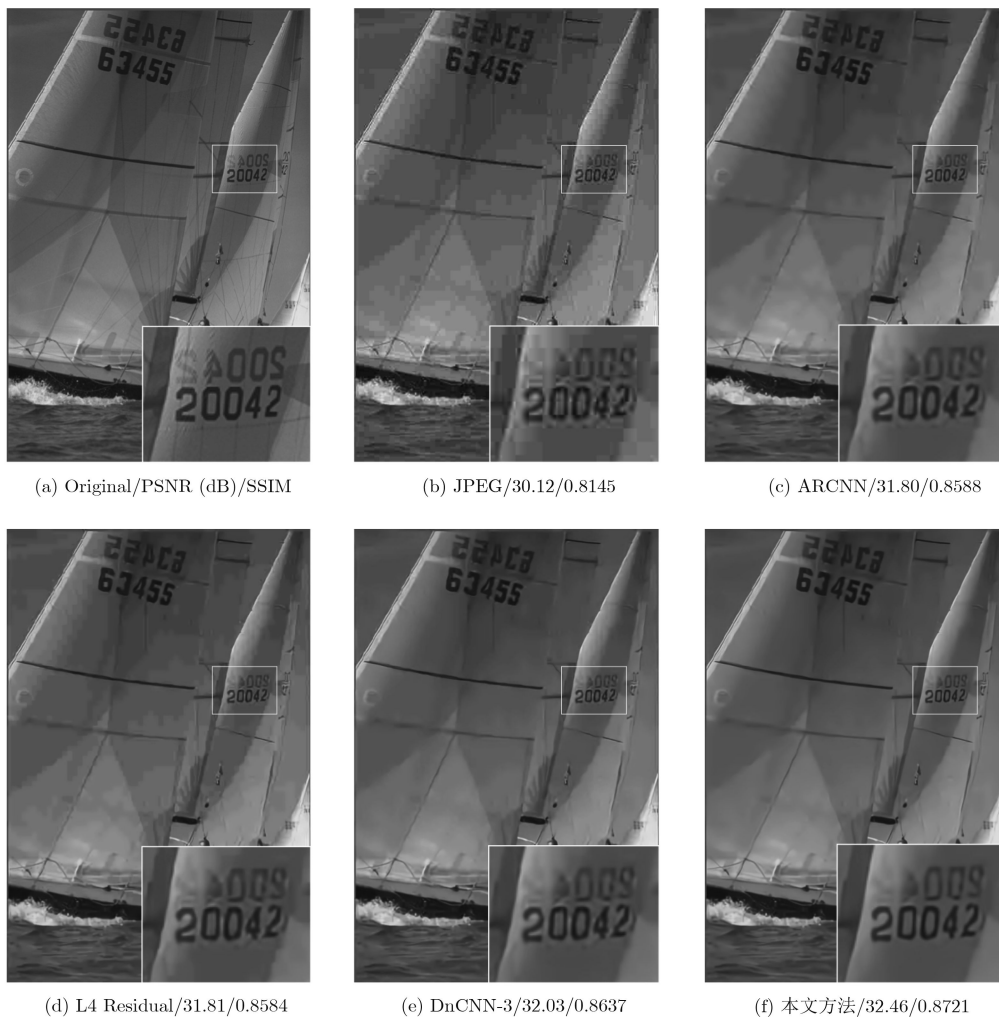


图4 QF为10时, 图像sailing3在各个算法中的视觉比较

表4 本文方法在LIVE1数据集上的PSNR(dB)/SSIM对比

方法	QF			
	15	25	35	45
JPEG	29.13/0.8402	30.81/0.8869	31.93/0.9101	32.78/0.9241
DnCNN-3	30.61/0.8697	32.35/0.9094	33.53/0.9287	34.39/0.9400
本文方法	<b>30.83/0.8733</b>	<b>32.50/0.9095</b>	<b>33.68/0.9303</b>	<b>34.56/0.9416</b>

表5 不同尺度的选择在LIVE1数据集上的PSNR(dB)/SSIM对比

不同尺度	QF			
	10	15	20	25
单一尺度(3×3)	29.42/0.8309	30.78/0.8719	31.75/0.8942	32.46/0.9086
单一尺度(5×5)	29.44/0.8316	30.79/0.8719	31.76/0.8945	32.46/0.9090
本文方法	<b>29.49/0.8329</b>	<b>30.83/0.8733</b>	<b>31.81/0.8952</b>	<b>32.50/0.9095</b>

### 5.3 不同网络层数的对比实验

本文网络中的每个稠密块都由3层(Dense3), 4层(Dense4), 5层(Dense5)或6层(Dense6)组成, 分别代表了4种不同的网络层数。如表6所示, 当QF为10, 15, 20和25时, 使用Dense3的网络到使用Dense5的网络在LIVE1数据集上的平均PSNR和SSIM的数据呈递增趋势, 使用Dense6的网络与使用Dense5的网络性能相当。因此, 本文网络的每个稠密块都采用5个卷积层组成。

表6 不同网络层数在LIVE1数据集上的PSNR(dB)/SSIM对比

不同层数	QF			
	10	15	20	25
Dense3	29.45/0.8318	30.79/0.8714	31.77/0.8947	32.47/0.9090
Dense4	29.47/0.8325	30.81/0.8728	31.79/0.8950	32.49/0.9092
Dense5	<b>29.49/0.8329</b>	<b>30.83/0.8733</b>	<b>31.81/0.8952</b>	<b>32.50/0.9095</b>
Dense6	29.47/0.8324	30.81/0.8735	31.79/0.8951	32.49/0.9099

### 5.4 使用稠密块对网络性能的影响

本文对比了分别使用普通块和稠密块对网络性能的影响, 其中普通块和稠密块均由5个卷积层组成, 但稠密块的每一层都与后面所有层相级联。如表7所示, 当QF为10, 15, 20和25时, 使用稠密块的网络在LIVE1数据集上的平均PSNR分别比使用普通块的网络高0.10 dB, 0.08 dB, 0.10 dB和0.09 dB; 使用稠密块的网络在LIVE1数据集上的平均SSIM

表7 使用普通块和稠密块在LIVE1数据集上的PSNR(dB)/SSIM对比

方法	QF			
	10	15	20	25
普通块	29.39/0.8303	30.75/0.8712	31.71/0.8938	32.41/0.9081
稠密块	<b>29.49/0.8329</b>	<b>30.83/0.8733</b>	<b>31.81/0.8952</b>	<b>32.50/0.9095</b>

分别比使用普通块的网络高0.0026, 0.0021, 0.0014和0.0014。因此, 为了提高网络的性能, 本文选用稠密块作为网络中多尺度结构的分支。

## 6 结束语

本文提出了去除JPEG压缩伪迹的多尺度稠密残差网络。网络中的稠密块实现了特征重用, 有利于提高网络的学习效率。为了提取不同尺度的特征, 本文采用不同扩张因子的扩张卷积, 形成了多尺度稠密块。本文利用多尺度稠密块将网络设计成包含2条支路的结构, 其中后一条支路用于补充前一条支路没有提取到的特征。最后, 使用残差学习的方式来提高网络的性能。在训练时, 本文使用“盲去块”的联合训练方式增强网络的泛化能力。在LIVE1和Classic5数据集上的实验结果表明, 本文所提出的方法不仅具有较高的性能, 还具有较强的泛化能力。

## 参考文献

- [1] FOI A, KATKOVNIK V, and EGI AZARIAN K. Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images[J]. *IEEE Transactions on Image Processing*, 2007, 16(5): 1395–1411. doi: 10.1109/TIP.2007.891788.
- [2] YOO S B, CHOI K, and RA J B. Post-processing for blocking artifact reduction based on inter-block correlation[J]. *IEEE Transactions on Multimedia*, 2014, 16(6): 1536–1548. doi: 10.1109/TMM.2014.2327563.
- [3] ZHAO Chen, ZHANG Jian, MA Siwei, et al. Reducing image compression artifacts by structural sparse representation and quantization constraint prior[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2017, 27(10): 2057–2071. doi: 10.1109/TCSVT.2016.2580399.

- [4] 吕晓琪, 吴凉, 谷宇, 等. 基于深度卷积神经网络的低剂量CT肺部去噪[J]. 电子与信息学报, 2018, 40(6): 1353–1359. doi: [10.11999/JEIT170769](https://doi.org/10.11999/JEIT170769).
- LÜ Xiaoqi, WU Liang, GU Yu, *et al.* Low dose CT lung denoising model based on deep convolution neural network[J]. *Journal of Electronics & Information Technology*, 2018, 40(6): 1353–1359. doi: [10.11999/JEIT170769](https://doi.org/10.11999/JEIT170769).
- [5] DONG Chao, LOY C C, HE Kaiming, *et al.* Image super-resolution using deep convolutional networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295–307. doi: [10.1109/TPAMI.2015.2439281](https://doi.org/10.1109/TPAMI.2015.2439281).
- [6] 郭智, 宋萍, 张义, 等. 基于深度卷积神经网络的遥感图像飞机目标检测方法[J]. 电子与信息学报, 2018, 40(11): 2684–2690. doi: [10.11999/JEIT180117](https://doi.org/10.11999/JEIT180117).
- GUO Zhi, SONG Ping, ZHANG Yi, *et al.* Aircraft detection method based on deep convolutional neural network for remote sensing images[J]. *Journal of Electronics & Information Technology*, 2018, 40(11): 2684–2690. doi: [10.11999/JEIT180117](https://doi.org/10.11999/JEIT180117).
- [7] REN Shaoqing, HE Kaiming, GIRSHICK R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137–1149. doi: [10.1109/TPAMI.2016.2577031](https://doi.org/10.1109/TPAMI.2016.2577031).
- [8] BADRINARAYANAN V, KENDALL A, and CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481–2495. doi: [10.1109/TPAMI.2016.2644615](https://doi.org/10.1109/TPAMI.2016.2644615).
- [9] DONG Chao, DENG Yubin, LOY C C, *et al.* Compression artifacts reduction by a deep convolutional network[C]. 2015 IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 576–584. doi: [10.1109/ICCV.2015.73](https://doi.org/10.1109/ICCV.2015.73).
- [10] SVOBODA P, HRADIS M, BARINA D, *et al.* Compression artifacts removal using convolutional neural networks[J/OL]. arXiv preprint arXiv: 1605.00366. <http://arxiv.org/abs/1605.00366>, 2016.
- [11] ZHANG Kai, ZUO Wangmeng, CHEN Yunjin, *et al.* Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising[J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3142–3155. doi: [10.1109/TIP.2017.2662206](https://doi.org/10.1109/TIP.2017.2662206).
- [12] KIM Y, HWANG I, and CHO N I. A new convolutional network-in-network structure and its applications in skin detection, semantic segmentation, and artifact reduction[J/OL]. arXiv preprint arXiv: 1701.06190. <http://arxiv.org/abs/1701.06190>, 2017.
- [13] SZEGEDY C, LIU Wei, JIA Yangqing, *et al.* Going deeper with convolutions[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 1–9. doi: [10.1109/CVPR.2015.7298594](https://doi.org/10.1109/CVPR.2015.7298594).
- [14] LIU Pengju, ZHANG Hongzhi, ZHANG Kai, *et al.* Multi-level Wavelet-CNN for image restoration[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, USA, 2018: 886–895. doi: [10.1109/CVPRW.2018.00121](https://doi.org/10.1109/CVPRW.2018.00121).
- [15] RONNEBERGER O, FISCHER P, and BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]. The 18th International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015: 234–241. doi: [10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [16] GUO Jun and CHAO Hongyang. One-to-many network for visually pleasing compression artifacts reduction[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 4867–4876. doi: [10.1109/CVPR.2017.517](https://doi.org/10.1109/CVPR.2017.517).
- [17] GALTERI L, SEIDENARI L, BERTINI M, *et al.* Deep generative adversarial compression artifact removal[C]. 2017 IEEE International Conference on Computer Vision, Venice, Italy, 2017: 4836–4845. doi: [10.1109/ICCV.2017.517](https://doi.org/10.1109/ICCV.2017.517).
- [18] HUANG Gao, LIU Zhuang, WEINBERGER K Q, *et al.* Densely connected convolutional networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017, 1: 2261–2269. doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243).
- [19] ZHANG Kai, ZUO Wangmeng, GU Shuhang, *et al.* Learning deep CNN denoiser prior for image restoration[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017, 2: 2808–2817. doi: [10.1109/CVPR.2017.300](https://doi.org/10.1109/CVPR.2017.300).
- [20] KRIZHEVSKY A, SUTSKEVER I, and HINTON G E. ImageNet classification with deep convolutional neural networks[C]. The 25th International Conference on Neural Information Processing Systems, Lake Tahoe, USA, 2012: 1097–1105.
- 陈书贞: 女, 1968年生, 副教授, 研究方向为图像处理、压缩感知、深度学习、相位恢复。
- 张祎俊: 女, 1994年生, 硕士生, 研究方向为深度学习, JPEG压缩伪迹去除。
- 练秋生: 男, 1969年生, 教授, 博士生导师, 研究方向为稀疏表示、深度学习、压缩感知及相位恢复。