

基于深度学习的高噪声图像去噪算法

盖 杉¹ 鲍中运¹

摘 要 为了更有效地实现高噪声环境下的图像去噪, 本文提出一种基于深度学习的高噪声图像去噪算法. 该算法首先采用递增扩充卷积并且融合批量标准化和 Leaky ReLU 函数对输入含噪图像进行特征提取与学习; 然后通过结合递减扩充卷积和 ReLU 函数对提取的特征进行图像重构; 最后通过整合残差学习和批量标准化的端到端网络实现图像与噪声的有效分离. 实验结果表明, 本文提出的算法不仅能够有效地去除高噪声环境下的图像噪声, 获得更高的峰值信噪比 (Peak signal-to-noise ratio, PSNR) 与结构相似度 (Structural similarity index, SSIM), 而且还能够有效地改善图像的视觉效果, 具有较好的实用性.

关键词 深度学习, 图像去噪, 卷积神经网络, 残差学习, 批量标准化

引用格式 盖杉, 鲍中运. 基于深度学习的高噪声图像去噪算法. 自动化学报, 2020, 46(12): 2672–2680

DOI 10.16383/j.aas.c180271

High Noise Image Denoising Algorithm Based on Deep Learning

GAI Shan¹ BAO Zhong-Yun¹

Abstract In order to perform image denoising in high-noise environment more effectively, a high-noise image denoising algorithm based on deep learning is proposed in this paper. Firstly, the proposed algorithm utilized increased expanded convolutional and combined the batch normalization and Leaky ReLU function to extract and learn for the features of noisy image. Secondly, the extracted feature via the decreased expansion convolution and ReLU function for image reconstruction. Finally, the effective separation of image and noise is realized by end-to-end network of integrating the residual learning and batch standardization. The experimental results illustrated that the proposed algorithm removed the image noise more effectively and obtain higher peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) under high noise environment. In addition, the proposed algorithm also significantly improved the visual effects of images, and had a good practicality.

Key words Deep learning, image denoising, convolutional neural networks (CNN), residual learning, batch normalization

Citation Gai Shan, Bao Zhong-Yun. High noise image denoising algorithm based on deep learning. *Acta Automatica Sinica*, 2020, 46(12): 2672–2680

在实际应用中, 数字图像在传输过程中往往会受到成像设备与外部噪声环境干扰等因素的影响, 导致采集的图像质量明显下降. 鉴于计算机视觉等许多科学领域对图像质量提出更高的要求, 因此图像去噪仍然是图像处理领域的热点研究课题之一.

在过去的几十年里, 随着对图像噪声的深入研究, 许多学者不断地提出新的图像去噪算法. 三维块匹配 (Block-matching and 3D filtering, BM3D) 算法^[1] 能够充分挖掘自然图像中存在的自相似特性, 通过对相似块进行域变换进而达到图像的去噪效果. 非局部算法^[2] 从图像整体的角度出发, 同时利用了

局部平滑与全局自相似等特性, 取得了很好的去噪效果.

随着深度学习逐渐成为机器学习领域的研究热点, 深度卷积神经网络^[3–5] 在图像特征提取与识别^[6–8] 等领域的成功应用为解决图像去噪问题提供了新的思路, 尤其是高噪声环境下的图像去噪问题. 与传统的图像去噪方法相比, 深度卷积神经网络具有更强大的学习能力, 通过使用大量含噪图像样本数据进行训练, 能够有效地提高网络模型对不同标准噪声的适应能力, 并使其具有更强的泛化能力.

在文献 [9] 中, Jain 等提出一种全新的卷积神经网络结构, 并将其应用于图像去噪, 实验结果表明该网络模型能够取得与马尔科夫随机场 (Markov random field, MRF) 模型相当甚至更好的去噪效果. Burger 等^[10] 提出的算法将多层感知机 (Multi-layer perceptron, MLP) 成功地应用于图像去噪. 文献 [11] 提出一个可训练的非线性反应扩散模型 TNRD (Trainable nonlinear reaction diffusion), 该模型通过展开固定数量的梯度下降前馈深度网络,

收稿日期 2018-05-02 录用日期 2019-01-09
Manuscript received May 2, 2018; accepted January 9, 2019
国家自然科学基金 (61563037), 江西省杰出青年人才资助计划 (20171BCB23057) 资助
Supported by National Natural Science Foundation of China (61563037) and Outstanding Youth Funding Scheme of Jiangxi Province (20171BCB23057)
本文责任编辑 杨健
Recommended by Associate Editor YANG Jian
1. 南昌航空大学 南昌 330063
1. Nanchang Hangkong University, Nanchang 330063

提高了图像的去噪性能. 文献 [12] 中, Xie 等将堆叠稀疏去噪自编码器方法应用于解决高斯噪声的移除并且实现了与 K-SVD (Singular value decomposition)^[13] 相当的去噪效果. 文献 [14] 中, Zhang 等提出一种基于深度学习的去噪算法 DnCNN (Feed-forward denoising convolutional neural networks, DnCNN). 该算法采用训练单一的去噪模型实现图像去噪的任务, 同时对未知噪声水平的图像也有比较好的去噪效果. 实验结果表明, 该算法的去噪性能和效率均优于 BM3D.

上述这些经典的图像去噪算法虽然在训练目标设计、训练特征选择以及训练集规模上各不相同, 并且在低噪声环境下都能取得很好的去噪效果. 但是, 这些算法在高噪声环境下的去噪效果却不太理想. 为了进一步改善高噪声环境下的图像去噪质量, 本文提出一种对称式扩充卷积残差网络图像去噪算法. 该算法首先通过对称式结构的卷积网络对输入噪声图像进行特征提取与学习, 然后对提取的图像特征进行重构, 最后通过整合残差学习和批量标准化实现图像与噪声的有效分离, 并输出与输入图像尺寸相同的残差图像. 为了解决卷积操作导致的网络内部协变量转移问题, 本文使用批量标准化进行校正, 有效地提高了网络训练的效率. 另外, 本文算法对非卷积后的图像进行零填充操作, 保证在图像大小不变的情况下, 降低了图像的边界伪影. 实验结果表明, 本文提出的算法在去噪性能和效率上都表现得非常好.

1 深度卷积神经网络相关技术原理

1.1 残差学习

卷积神经网络的残差学习最初的灵感来自于在不断增加网络的深度时, 会出现一个训练精度退化的问题, 即神经网络的训练精度先上升然后达到饱和, 再持续增加网络深度时则会导致训练精度下降. 因此, 残差学习的提出起初是用来解决神经网络性能的退化问题. 在这里, 它是通过假设残差映射是很容易被学习并且有一个浅层的网络已达到了饱和的准确率, 后面再加上几个 $y = x$ 的全等映射层, 起码误差不会增加, 即更深的网络不应该带来训练集误差的上升问题, 这里提到的使用全等映射直接将前一层输出传递到后一层的思想就是残差网络通过一些堆叠层明确地学习残差映射思想的来源. 采用残差学习策略使得深度卷积神经网络很容易被训练并且能更好地提升图像分类和目标检测的精确度. 本文与之不同之处在于采用两个复合残差单元来预测残差图像.

1.2 批量标准化

目前小批量随机梯度下降算法 (Stochastic gradient descent, SGD) 已广泛用于神经网络模型的训练. 该算法虽然简单有效, 但是在网络训练过程中容易出现内部协变量转移, 进而改变网络内部非线性输入分配, 导致训练效率降低. 因此, 批量标准化 (Batch normalization, BN)^[15] 的提出可以用来解决网络内部协变量转移, 克服深度神经网络难以训练的缺陷, 有效预防梯度弥散. 本文中, 我们在神经网络每一层非线性化之前引入批量标准化操作, 然后结合归一化、缩放与移位变化等操作来解决网络内部协变量转变的问题. 每一个激活函数在批量标准化过程中只需要添加两个参数, 它们通过反向传播进行更新. 这种机制能够加快训练过程中的收敛速度, 提高模型的精确度, 降低模型初始化具有的敏感性.

1.3 高噪声环境下 Leaky ReLU 函数的作用

Leaky ReLU 函数^[16-17] 的提出是基于 ReLU 函数的改进, 它不仅具有 ReLU 函数的优点, 而且能够修正数据分布, 在网络的训练过程中使得阈值为负的神经元得到了训练和更新, 这样阈值为负的神经元信息就得到了一定的保留, 同时也提升了网络的拟合能力. 鉴于高噪声环境下, 图像内部可利用的有效信息相对较少, 因此在网络训练过程中, 我们需要更加充分地挖掘和利用图像的有效信息, 提取和学习图像特征. Leaky ReLU 函数在网络训练过程中能够更好地保留图像的有效信息, 获得最佳的去噪效果. 另外, 本文采用 Msra 高斯分布对神经网络的权重值进行初始化, 并且该权重初始化服从 $(0, 2/n)$ 的高斯分布. 实验证明这种初始化方法能够更好地适应 Leaky ReLU 函数, 加快梯度在网络结构中更新, 从而使得本文算法在高噪声环境下也能够取得较好的去噪效果.

2 基于深度学习的高噪声图像去噪算法

本文研究的高噪声图像去噪算法是针对加性白高斯噪声, 其模型定义为

$$Y = X + N \quad (1)$$

其中, X 是原始无噪图像, N 是分布为 $N(0, \sigma^2)$ 的噪声图像, σ 为噪声标准差, Y 是含噪图像. 本文提出的基于对称式扩充卷积残差网络模型结构示意图如图 1 所示.

2.1 网络深度

根据 VGGNet^[18] 中的原则, 首先移除所有的池化层, 然后改进卷积核的大小设置. 目前深度卷积神经网络中最常用的卷积核大小为 3×3 , 相应深度为

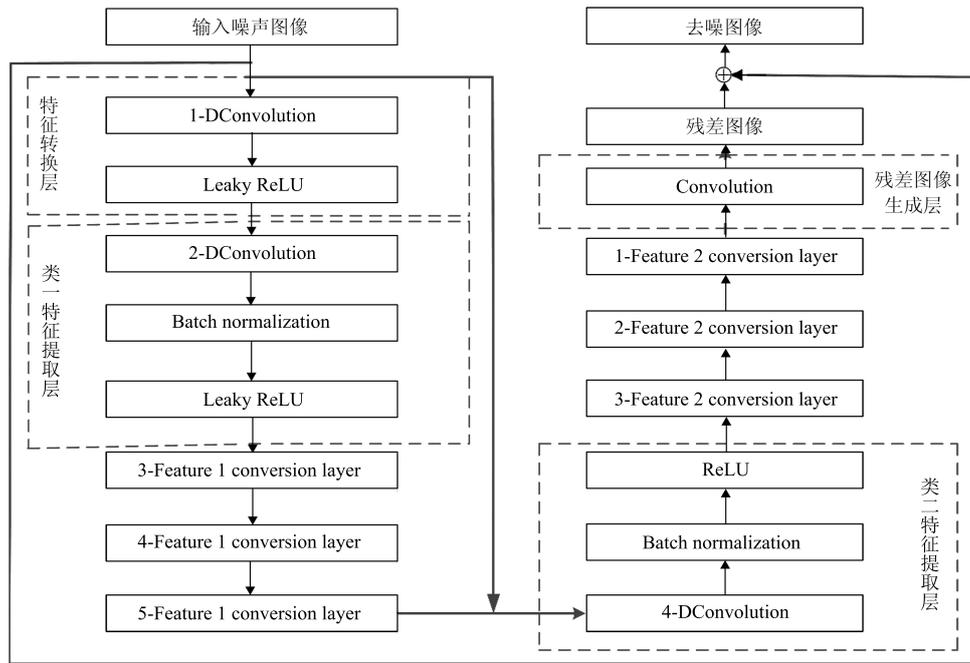


图 1 本文模型结构示意图

Fig. 1 The structure of the proposed model

f 的卷积神经网络的感受野大小为 $(2f + 1) \times (2f + 1)$. 众所周知, 图像的上下文信息有助于对图像中损坏像素进行恢复. 我们在卷积神经网络中, 通过前向卷积操作扩大网络的感受野, 进而得到上下文信息. 尤其在高噪声环境下, 为了恢复图像中损坏的像素, 需要更大的感受野来获得上下文信息. 这增加了网络的计算负担, 需要网络付出计算复杂度和计算效率的代价.

为了解决上述问题, 本文使用一种扩充卷积在神经网络感受野的尺寸大小与网络深度之间进行权衡. 扩充卷积不但具有扩充感受野的能力, 同时也保留了传统 3×3 卷积的优点而得到广泛使用. 在本文构建的神经网络中, 从第 1 层至倒数第 2 层均采用扩充卷积, 其扩充因子分别设置为 1, 2, 3, 4, 5, 4, 3, 2, 1, 最后一层使用 3×3 卷积, 网络总深度设置为 10 层. 其中, 扩充卷积中的扩充因子是由尺寸大小为 $(2r + 1) \times (2r + 1)$ 的稀疏过滤器构成.

因此, 扩充卷积的每一层等效的感受野大小为 3, 5, 7, 9, 11, 9, 7, 5, 3. 在相同的网络深度前提下, 扩充卷积网络的感受野大小为 51×51 , 可接受域大小为 21×21 . 相比较而言, 传统的 3×3 卷积要想得到相同大小的可接受域, 则需要将网络深度扩展至 25 层. 为了验证本文设计的网络模型在高噪声环境下去噪的有效性, 我们通过实验对比了本文算法与上述两种算法的去噪效果和效率. 其训练收敛情况如图 2 所示.

从实验结果可以看出, 感受野大小为 51×51 的 25 层网络需要训练迭代 60 轮才能达到收敛状态, 而感受野大小为 21×21 的 10 层网络和本文设计的网络模型分别只需要训练迭代 40 轮和 45 轮就达到了收敛状态, 在效率上节约了训练的收敛时间. 另外, 在对比实验中, 我们把噪声标准差固定在 $\sigma = 50$ 的条件下进行实验, 使用相同的训练集, 训练三个不同的模型, 并且在 BSD68 数据集上进行不同模型的去噪性能评估. 实验结果表明, 本文提出的网络模型在测试集上最终的平均峰值信噪比为 26.42 dB, 比感受野大小为 21×21 的 10 层网络模型的平均峰值信噪比 25.87 dB 多 0.55 dB. 同时, 本文提出的网络模型平均峰值信噪比也非常接近感受野大小为 51×51 的 25 层网络的平均峰值信噪比 26.45 dB.

2.2 网络模型结构

本文模型的输入是含有高斯噪声的图像, 表示为 $Y = X + N$. 本文提出的去噪模型通过学习映射函数 $f(Y) = X$ 来预测潜在的无噪图像, 相应的损失函数定义为

$$L = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{w \times h} \sum_{j=1}^w \sum_{k=1}^h \|f_i(j, k) - X_i(j, k)\|^2 \right) \quad (2)$$

其中, f 表示去噪后的图像, X 表示原始无噪图像,

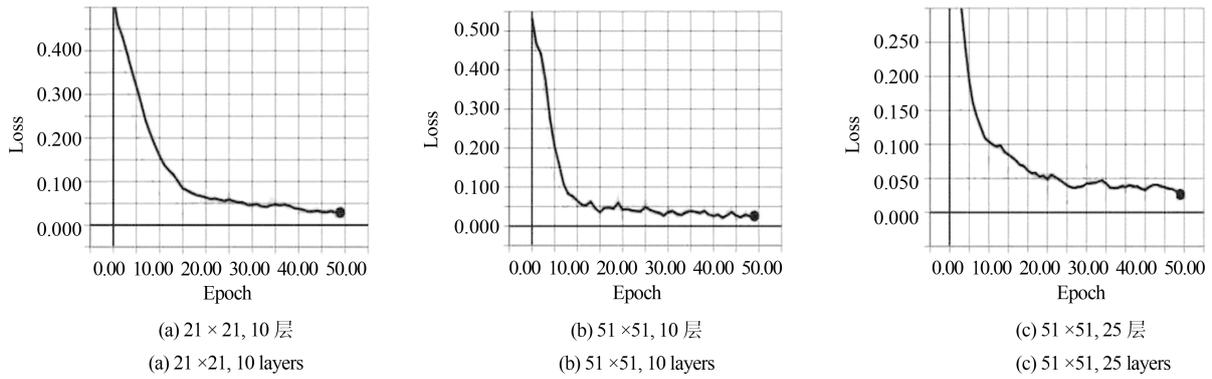


图2 不同网络结构模型训练收敛图

Fig. 2 The training convergence graph of different network

n 为每个训练批次的样本数量, w, h 表示每个样本的宽度和高度.

接下来, 我们采用残差学习来训练残差映射 $R(Y) \approx N$, 然后通过 $X = Y - R(Y)$ 计算期望残差图像与预测残差图像之间的均方误差. 定义为

$$\ell(w, b, X_i, Y_i) = \frac{1}{2N} \sum_{i=1}^N \|R_{w,b}(Y_i) - (Y_i - x_i)\|_F^2 \quad (3)$$

其中, w, b 表示神经网络中可学习的参数, $\{Y_i, X_i\}_{i=1}^N$ 表示 N 个噪声与干净图像样本, $\ell(w, b; X_i, Y_i)$ 表示网络预测的 N 个样本噪声图像 $R_{w,b}(Y_i)$ 与期望噪声图像 $(Y - X)$ 之间的均方误差即代价函数.

为了计算函数 $\ell(w, b; X, Y)$ 的最小值, 本文提出的模型采用随机梯度下降算法迭代优化和更新参数矩阵 w 和参数向量 b , 定义为

$$w_{ij}^{(l+1)} = w_{ij}^{(l)} - \alpha \frac{\partial}{\partial w_{ij}^{(l)}} \ell(w, b) \quad (4)$$

$$b_i^{(l+1)} = b_i^{(l)} - \alpha \frac{\partial}{\partial b_i^{(l)}} \ell(w, b) \quad (5)$$

其中, α 表示学习率, l 表示参数所在的层序数, $\ell(w, b; X, Y)$ 表示目标损失函数. 采用神经网络的前向传播和反向传导算法, 利用上述迭代规则, 我们能够训练得到预期的网络模型.

本文提出的网络结构模型共设置为 10 层, 具体分析如下:

1) 1-DConv + Leaky ReLU: 网络模型的第一层采用扩充率大小为 1, 核尺寸大小为 $1D \times c$ 的 32 个卷积核生成 32 个子特征, 并使用 Leaky ReLU ($\max(0.01x, x)$) 对子特征实现非线性变换形成 1 个非线性特征映射. 其中, 步长设为 1×1 , 这里 c 表示图像的通道数.

2) 2/3/4/5-DConv + BN + Leaky ReLU: 网络模型的第 2~5 层, 分别使用大小为 $2D \times 32$ 的 64 个卷积核, $3D \times 64$ 的 128 个卷积核, $4D \times 128$ 的 256 个卷积核和 $5D \times 256$ 的 c 个卷积核. 其中, 步长均设置为 1×1 , 在扩充卷积层和 Leaky ReLU 函数之间进行批量标准化操作^[15].

3) 4/3/2/1-DConv + BN + ReLU: 网络模型的 4~1 层分别设置大小为 $4D \times c$ 的 256 个卷积核, $3D \times 256$ 的 128 个卷积核, $2D \times 128$ 的 64 个卷积核和 $1D \times 64$ 的 32 个卷积核. 其中, 步长均设置为 1×1 .

4) Conv 层: 网络的最后一个卷积层通过 3×3 的 c 个卷积核, 及 1×1 步长, 实现特征图像重构并输出.

考虑到图像在高噪声环境下内部可利用的有效信息较少的特点以及整个网络结构的计算复杂度, 本文融合扩充卷积和双函数对称残差网络对数据进行建模, 并引入残差学习和批量归一化加速网络训练速度并提升网络的去噪性能, 构造了 10 层对称式扩充卷积残差网络实现高噪声环境下的图像去噪任务. 本文提出的算法在高噪声环境下以及盲高噪声环境下都取得了较好的去噪效果, 同时也提升了去噪效率. 本文的算法的主要优点和贡献总结如下:

1) 通过实验发现扩充卷积和 Leaky ReLU 函数更有利于卷积神经网络实现高噪声环境下的图像去噪任务. 扩充卷积操作和 Leaky ReLU 函数的运用能够在不增加网络计算负担的同时增大网络的感受野, 使网络可以获得并充分利用图像中更多的有效的上下文信息来实现去噪图像中损坏像素的恢复, 完成图像去噪任务.

2) 通过构造对称式扩充残差卷积神经网络, 设计对称式的渐变特征提取、双激活函数和双残差块, 并结合批量归一化算法, 实现了一种端到端的可训练的去噪卷积神经网络模型用于高噪声环境下的图像去噪任务. 通过与现有的经典去噪方法相比, 本文

算法通过迭代优化噪声误差值达到优化网络模型的目的。

3) 本文提出的网络模型能够通过特定高斯噪声水平下和随机高斯噪声水平下训练单一的网络模型实现高噪声图像去噪任务, 而且去噪效果要优于那些在特定高斯噪声水平下的经典去噪方法。

降低边界伪影, 在许多低层视觉应用中, 通常要求网络的输出图像尺寸和输入图像尺寸保持一致。由于卷积的特点, 如果在卷积操作过程中, 不对卷积前的图像做任何预处理, 则将会导致最后的图像边界出现伪影。因此, 本文提出的算法在每次非 1×1 卷积操作之后直接对图像块进行零填充, 并且我们希望设计的卷积神经网络能够捕获模型图像边界, 确保中间隐含层的每一个特征映射与输入图像有相同的尺寸。我们发现简单的零填充策略没有导致任何边界伪影, 这种好的性能或许一方面归功于本文模型的强大性能, 另一方面则归功于在对图像进行裁剪时所做的贡献。当把图像裁剪成小的图像块时, 它能够使卷积神经网络获得更多的边界信息, 即对图像的边界信息进行了增强。本文也测试了使用更大尺寸的图像块时的网络性能, 实验结果表明, 更大尺寸的图像块只能给网络的去噪性能带来很小的改善, 但是训练图像块的尺寸小于网络可接受域大小时, 网络性能则会下降。

2.3 融合残差学习与批量标准化的图像去噪

对称式扩充卷积残差网络结构展示了该网络训练残差映射 $R(Y)$ 来预测 N 的过程。通过文献 [19] 可知, 当原始映射近似为一个同等映射时, 残差映射将更容易被优化。尤其是, 含噪图像 Y 比起残差图像 $R(Y)$ 更接近潜在的清晰图像, 即原始映射比 $R(Y)$ 更加接近于同等映射, 因此, 残差学习法更适合图像去噪。

虽然采用随机梯度下降算法能够加速训练速度和提高网络性能, 但是网络的结构设计也是一个重要的因素。批量标准化和残差学习是最常用的两个结构设计技术, 它们广泛应用于卷积神经网络的结构设计中。尤其是批量标准化, 定义如下:

$$\bar{Z}^{(i)} = \gamma Z_{\text{norm}}^{(i)} + \beta \quad (6)$$

其中, γ, β 表示可调参数, \bar{Z} 是调整参数 γ, β 后的 Z_{norm} 的分布, Z_{norm} 定义如下:

$$Z_{\text{norm}}^{(i)} = \frac{Z^{(i)} - \mu}{\sqrt{\sigma^2 + \xi}} \quad (7)$$

其中, ξ 为一个取值很小的参数, 用于保证分母不为 0, Z_{norm} 表示正则化, Z 表示未激活前的神经元, μ, σ^2 分别表示样本均值和方差, μ, σ^2 有如下定义:

$$\mu = \frac{1}{m} \sum_i Z^i \quad (8)$$

$$\sigma^2 = \frac{1}{m} \sum_i (Z^i - \mu)^2 \quad (9)$$

它可以用于对在训练期间的每个卷积层所提取的图像特征数据实现归一化处理, 使各层网络学习相同的特征数据分布, 提高网络的训练效率。鉴于在高斯去噪的过程中, 残差学习和批量标准化都与高斯分布有关, 因此整合批量标准化和残差学习非常有利于高斯图像去噪。更准确地说, 它不仅加速和稳定网络的训练, 而且还能获得更好的去噪性能。

综上所述, 本文提出的基于深度学习的高噪声图像去噪算法具体步骤如下:

算法 1. 基于深度学习的高噪声图像去噪算法

输入层. 网络模型的输入层是加噪图像样本 Y 。

输出层. 网络模型的输出层是预测噪声图像 $R(Y)$ 。

步骤 1. 加噪图像首先通过构建模型的前半部分递增式扩充卷积操作, 进行图片特征的学习, 同时在每层卷积之后, 利用式 (6) 实现批量标准化, 避免了卷积之后出现网络内部协变量转移问题, 最后通过式 Leaky Relu 函数进行图像特征非线性映射, 实现图像特征提取。

步骤 2. 步骤 1 得到的特征图像通过模型的后半部分递减式扩充卷积操作、批量标准化和 Relu 函数的非线性映射进行图像特征重构, 最后输出残差图像。

步骤 3. 根据步骤 2 中得到的残差图像对式 (3) 的损失函数值进行计算。

步骤 4. 利用式 (4) 和式 (5) 对式 (3) 中相应参数 w, b 进行反向迭代更新和优化, 最终得到理想的去噪模型。

步骤 5. 用步骤 4 中的理想模型对测试集中的加噪图片进行去噪测试, 并对去噪结果进行对比分析和评估。

3 实验结果与分析

3.1 实验设置

1) 训练和测试数据. 对于已知不同标准差的噪声水平的高斯图像去噪, 本文使用文献 [11] 中尺寸大小为 180×180 像素的 400 幅无噪图像作为原始训练集。为了训练所提出模型在特定的高斯噪声环境下的高斯去噪, 我们设置了 5 种噪声水平, 即噪声标准差 σ 分别为 15, 25, 40, 50 和 60。根据网络的深度以及感受野的大小设置训练数据集, 本文对原始的 400 幅无噪图像进行数据增强, 最终得到 128×1600 幅大小为 53×53 像素的图像样本训练本

文提出的网络模型. 此外, 我们也在随机噪声环境下训练本文提出的网络模型, 同样使用上述大小为 180×180 像素的 400 幅无噪图像样本作为原始数据集, 然后通过数据增强获得 128×2400 幅大小为 53×53 像素的训练样本集, 并且对其加入标准差 $\sigma \in [0, 60]$ 的随机噪声, 实现网络模型的训练. 我们在训练网络模型前对增强后的图像加入不同标准差的加性高斯噪声来实现本文噪声训练样本的生成.

与此同时我们为所有竞赛方法的性能评估设置同样的测试集, 它们来自于一个广泛使用的数据集: Berkeley segmentation dataset^[20], 它包含了 68 幅自然图像. 这些测试图像不包括在训练样本集里面.

2) 参数设置和网络训练. 为了权衡网络的感受野大小和网络深度, 本文对在特定噪声标准差和随机噪声标准差下训练得到的网络模型深度均设置为 10 层. 利用式 (5) 计算网络学习的残差映射 $R(Y)$ 与真实的残差 $(Y - X)$ 之间的损失代价. 本文在优化网络参数 w, b 的阶段, 采用文献 [21] 的方法初始化权重以及使用 0.0001 的权重衰减系数, 衰减率为 0.9 的 SGD 进行优化. 为了提高效率, 对网络的输入数据采用小批量输入法, mini-batch 设置为 128. 如果训练误差大小在连续 6 轮训练以内, 固定在某一个值上不变, 我们就停止迭代训练. 通过实验, 我们最终对模型进行 10 轮的迭代训练, 在训练过程中学习率初始化为 0.1, 当训练误差停止下降时, 我们固定到 0.0001 进行模型的训练. 基于数据增强的图像翻转、修剪是在图像批量学习中完成的.

3.2 实验分析

本文实验中深度卷积神经网络模型的整个训练过程中需要较多的数据集, 涉及到矩阵运算和图像处理单元, 因此我们在 Inter(R) Core(TM) i7-8700 CPU 3.30 GHz 和 GPU 上完成卷积神经网络的图像去噪模型训练和测试任务, 使用的 GPU 为 Nvidia Titan X GPU, 同时使用 TensorFlow (1.9.0, Python3.6.0) 深度学习框架进行实验. 为了验证本文算法的有效性, 我们将本文提出的去噪算法与

BM3D^[1]、WNNM^[22]、MLP^[10]、TNRD^[11]、EPLL^[23]、CSF^[24] 以及 DnCNN^[14] 的去噪效果进行比较. 本文主要采用峰值信噪比 (Peak signal to noise ratio, PSNR) 与结构相似度 (Structural similarity, SSIM) 两个指标衡量模型的去噪效果.

采用不同的去噪算法对 BSD68 数据集样本进行去噪的结果如表 1 所示. 从表 1 可以看出, 当 $\sigma = 15$ 时, WNNM^[22], MLP^[10], TNRD^[11] 的峰值信噪比 (PSNR) 比 BM3D^[1] 提高约 0.3 dB, 比 DnCNN^[14] 提高约 0.9 dB. 本文算法的峰值信噪比在 DnCNN^[14] 的基础上又提高了约 0.2 dB, 比 WNNM^[22], TNRD^[11] 提高约 0.5 dB, 比 BM3D^[1] 提高约 0.9 dB. 当对 $\sigma = 25, \sigma = 40, \sigma = 50, \sigma = 60$ 的高斯噪声图像进行去噪时, 由表 1 的对比实验数据可知, 本文算法的去噪性能同样表现得非常可观, 尤其在 $\sigma = 60$ 的高噪声环境下去噪效果要明显优于其他经典的去噪算法. 另外, 我们训练的随机噪声模型在特定标准差的噪声环境下去噪效果仍然表现得可观, 在上述 5 种特定的噪声水平下, 本文算法取得了较高的峰值信噪比, 尤其在 $\sigma = 60$ 时, 其峰值信噪比 (PSNR) 较 BM3D^[1] 算法的峰值信噪比提高约 2 dB, 与特定噪声模型的去噪效果相当.

本文还使用结构相似度这个指标来衡量网络模型的去噪效果, 把在特定噪声标准差下训练的去噪模型的去噪效果与经典的去噪模型的去噪效果进行对比, 实验结果如表 2 所示. 从表 2 中的实验数据可以看出, 在不同的噪声环境下, 本文算法的去噪效果均优于其他经典去噪算法. 尤其是在 $\sigma \geq 40$ 的高噪声环境下, 本文算法的去噪效果要明显优于其他经典去噪方法. 在 $\sigma = 60$ 时, 它们的去噪效果在结构相似度上要高于 DnCNN^[8] 约 0.005. 为了进一步验证本文算法去噪的可视化效果, 图 3 与图 4 分别给出了高噪声环境下 ($\sigma = 50$), 不同去噪算法对含噪图像 Lighthouse 与 Butterfly 的去噪效果.

对于图像去噪方法优劣的另一个重要衡量标准就是去噪模型的去噪效率. 本文在噪声标准差 $\sigma = 25$

表 1 不同去噪算法在 BSD68 数据集下的峰值信噪比 (PSNR) (dB)

Table 1 The PSNR value using different denoising algorithms at the BSD68 data set (dB)

σ	BM3D	WNNM	MLP	TNRD	DnCNN	EPLL	CSF	特定噪声模型	随机噪声模型
15	31.07	31.37	-	31.42	31.73	31.21	31.24	31.94	31.85
25	28.57	28.83	28.96	28.92	29.23	28.68	28.74	29.46	29.38
40	26.22	26.33	-	26.49	26.88	26.26	26.30	27.11	27.06
50	25.62	25.87	26.03	25.97	26.23	25.67	-	26.48	26.47
60	23.18	-	23.55	23.43	23.73	23.24	23.27	24.01	24.06

表 2 不同去噪算法在 BSD68 数据集下的结构相似度
Table 2 The SSIM value using different denoising algorithms at the BSD68 data set

σ	BM3D	WNNM	MLP	TNRD	DnCNN	本文方法 1	本文方法 2
15	0.8772	0.8774	0.8792	0.8826	0.8826	0.8831	0.8827
25	0.8017	0.8019	0.8120	0.8157	0.8190	0.8193	0.8190
40	0.7223	0.7237	0.7294	0.7310	0.7322	0.7334	0.7331
50	0.6869	0.6871	0.6956	0.7029	0.7076	0.7102	0.7100
60	0.6521	0.6544	0.6643	0.6712	0.6745	0.6796	0.6799

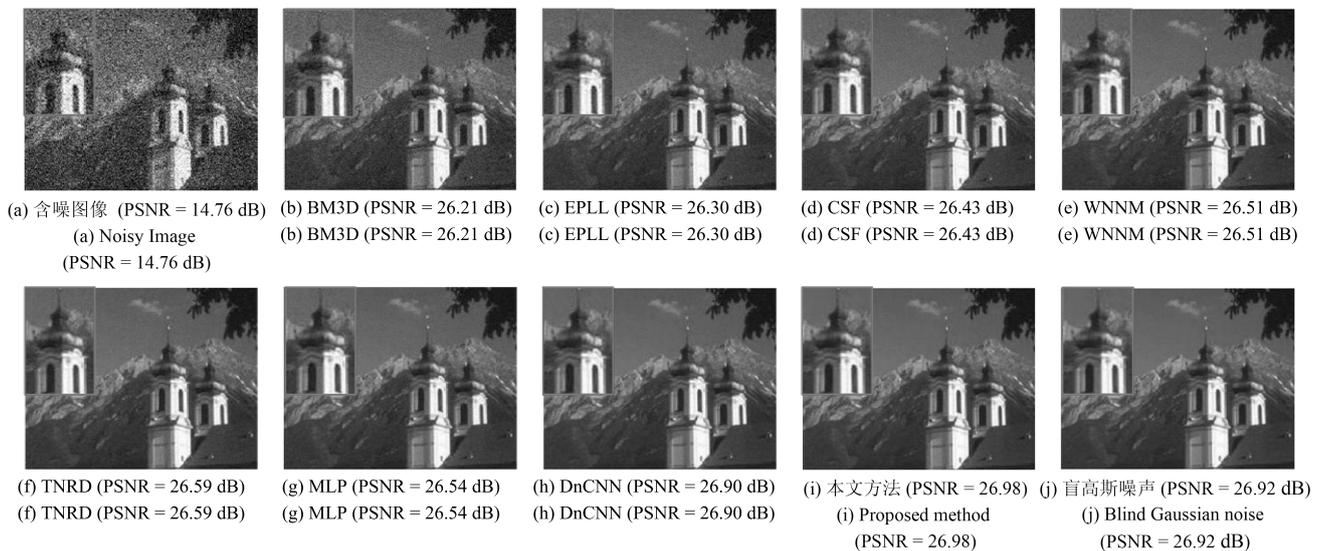


图 3 不同的去噪方法对 lighthouse 的去噪效果对比图

Fig. 3 Comparison of denoising effects of lighthouse with different denoising methods

且图像块的尺寸分别为 256×256 像素, 512×512 像素, 1024×1024 像素的条件下, 对比了本文提出的算法与其他去噪算法的运行时间, 如表 3 所示. 从表 3 中可以看出, 在相同条件下, 与算法 DnCNN^[8] 相比较, 本文算法在去噪性能与去噪效率上均得到一定程度的提升.

目前的去噪算法虽然取得了较好的去噪效果, 但是从去噪性能和去噪效率来看其仍然存在一些缺陷, 具体可归纳总结为以下三点:

1) 以 BM3D^[1] 为代表的传统经典去噪算法, 在高噪声环境下, 由于图像内部能够利用的有效信息较少, 去噪效果相对较差, 因此, 在高噪声环境下去噪效果表现得不是很理想.

2) 以 DnCNN^[14] 为代表的这种利用人工神经网络和残差学习方法通过学习自然图像的统计特性, 借助外部信息来为去噪过程服务的去噪算法. 它需要大量的数据集和更深层的网络结构, 并且通常在训练阶段涉及到复杂的迭代优化问题, 从而使得去

噪过程消耗大量的时间. 尤其在处理本文高噪声环境下图像去噪问题时, 网络在没有有效的牺牲耗时和计算复杂度的情况下很难实现高性能, 而且它在高噪声环境下的去噪质量还有很大的提升空间.

3) 目前用深度卷积神经网络无论是进行图像分类、识别还是图像去噪, 它们通常所使用的激活函数都是 ReLU 函数, 但是 ReLU 函数在神经网络训练的时候表现得很脆弱, 容易出现神经元“坏死”并且不可再被激活, 这就会导致在网络训练过程中部分神经元得不到非线性运算和更新, 从而使得图像部分特征信息丢失, 影响了网络的去噪性能.

本文在弥补以上算法缺点的同时充分兼顾去噪效率和去噪性能的基础上, 首次提出同时使用线性整流函数的变体 Leaky ReLU 函数和 ReLU 函数作为神经网络隐藏层的激活函数, 并整合残差学习和批量标准化来构建一种对称式扩充卷积残差网络, 在实现以少量的扩充卷积层就能达到与传统的更深卷积层相当的去噪效果的同时进一步增强了网络的

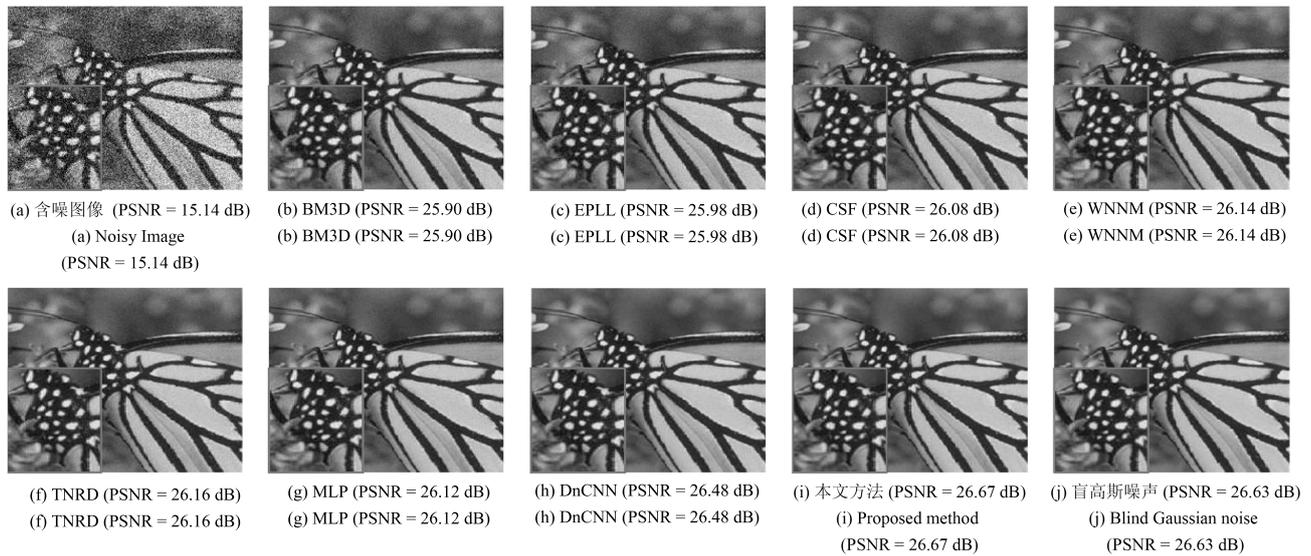


图 4 不同的去噪方法对 butterfly 的去噪效果对比图

Fig. 4 Comparison of denoising effects of butterfly with different denoising methods

表 3 不同尺寸大小的测试图像去噪运行时间比较 ($\sigma = 25$) (s)Table 3 The running time of test images denoising with different size ($\sigma = 25$) (s)

图像块大小 (像素)	配置	BM3D	WNNM	TNRD	MLP	EPLL	CSF	DnCNN	特定噪声模型	随机噪声模型
256 × 256	CPU/GPU	0.65	203.1	0.45/0.010	1.42	25.4	2.11/-	0.74/0.014	0.68/0.016	0.97/0.020
512 × 512	CPU/GPU	2.85	773.2	1.33/0.032	5.51	45.5	5.67/0.92	3.41/0.051	2.98/0.072	3.68/0.083
1024 × 1024	CPU/GPU	11.89	2536.4	4.61/0.116	19.4	422.1	40.8/1.72	12.1/0.200	10.7/0.160	13.7/0.173

去噪性能和效率. 通过上述的大量实验数据表明, 本文的实验模型能够在高噪声环境下产生更好的高斯去噪结果比经典的方法 BM3D^[1]、TNRD^[11]、WNNM^[22]、MLP^[10]、DnCNN^[14]、EPLL^[23] 和 CSF^[24] 的去噪效果还要好. 另外, 在去噪效率上也得到了一定程度的提升, 弥补了这些经典去噪算法的不足.

4 结论

本文提出了一种基于深度学习的高噪声图像去噪算法, 该算法通过构造对称式扩充卷积残差网络, 并结合 Leaky ReLU 和 ReLU 双函数, 能够更好地适应高噪声环境下的图像去噪, 同时得到的去噪图像不存在边缘细节的受损、边界伪影和清晰度变差问题. 此外, 本文提出的算法在继承和提高现有经典算法优点的同时也弥补了它们存在的不足之处. 实验结果表明, 本文算法得到了更高的 PSNR、SSIM 与良好的视觉效果, 在提高了高噪声环境下图像去噪能力的同时, 也一定程度地提升了高噪声环境下的去噪效率. 这也为后续的目标检测、识别和跟踪等应用提供了很好的图像样本数据, 具有很好的实用

性.

References

- Kostadin D, Alessandro F, Vladimir K, Karen E. Image denoising by sparse 3-D transform domain collaborative filtering. *IEEE Transactions on Image Processing*, 2007, **16**(8): 2080–2095
- Sun Wei-Feng, Dai Yong-Shou. Non-local mean image denoising using multi-level residual filtering. *Journal of Electronics and Information Technology*, 2016, **38**(8): 1999–2006
(孙伟峰, 戴永寿. 采用多级残差滤波的非局部均值图像去噪方法. *电子与信息学报*, 2016, **38**(8): 1999–2006)
- Zhou Fei-Yan, Jin Lin-Peng, Dong Jun. A review of convolutional neural networks. *Chinese Journal of Computers*, 2017, **40**(6): 1229–1251
(周飞燕, 金林鹏, 董军. 卷积神经网络研究综述. *计算机学报*, 2017, **40**(6): 1229–1251)
- Luo Jian-Hao, Wu Jian-Xin. A survey of fine-grained image classification based on deep convolution features. *Acta Automatica Sinica*, 2017, **43**(8): 1306–1318
(罗建豪, 吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述. *自动化学报*, 2017, **43**(8): 1306–1318)
- Tang Xian-Lun, Du Yi-Ming, Liu Yu-Wei, Li Jia-Xin, Ma Yi-Wei. Image recognition method based on conditional

- depth convolution to generate anti-network. *Acta Automatica Sinica*, 2018, **44**(5): 855–864
(唐贤伦, 杜一铭, 刘雨微, 李佳歆, 马艺玮. 基于条件深度卷积生成对抗网络的图像识别方法. 自动化学报, 2018, **44**(5): 855–864)
- 6 Jin Lian-Wen, Zhong Zhuo-Yao, Yang Zhao, Yang Wei-Xin, Xie Ze-Cheng, Sun Jun. Application of deep learning in handwritten Chinese character recognition. *Acta Automatica Sinica*, 2016, **42**(8): 1125–1141
(金连文, 钟卓耀, 杨钊, 杨维信, 谢泽澄, 孙俊. 深度学习在手写汉字识别中的应用综述. 自动化学报, 2016, **42**(8): 1125–1141)
- 7 Li Wen-Ying, Cao Bin, Cao Chun-Shui, Huang Yong-Zhen. A bronze inscription recognition method based on deep learning. *Acta Automatica Sinica*, 2018, **44**(11): 2023–2030
(李文英, 曹斌, 曹春水, 黄永祯. 一种基于深度学习的青铜器铭文识别方法. 自动化学报, 2018, **44**(11): 2023–2030)
- 8 Yao Nai-Ming, Guo Qing-Pei, Qiao Feng-Chun, Chen Hui, Wang Hong-An. Robust facial expression recognition based on generative confrontation network. *Acta Automatica Sinica*, 2018, **44**(5): 865–877
(姚乃明, 郭清沛, 乔逢春, 陈辉, 王宏安. 基于生成式对抗网络的鲁棒人脸表情识别. 自动化学报, 2018, **44**(5): 865–877)
- 9 Jain V, Seung H S. Natural image denoising with convolutional networks. In: Proceedings of the 22nd Annual Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc. 2008. 769–776
- 10 Burger H C, Schuler C J, Harmeling S. Image denoising: can plain neural networks compete with BM3D? In: Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, RI, USA: IEEE, 2012. 2392–2399
- 11 Chen Y J, Pock T. Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **39**(6): 1256–1272
- 12 Xie J Y, Xu L L, Chen E H. Image denoising and inpainting with deep neural networks. In: Proceedings of the 2012 International Conference on Neural Information Processing Systems. Curran-Associates Inc. 2012: 341–349
- 13 Zhang Q, Li B X. Discriminative K-SVD for dictionary learning in face recognition. In: Proceedings of the 23rd IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, USA: IEEE, 2010. 2691–2698
- 14 Zhang K, Zuo W M, Chen Y J, Meng D Y, Zhang L. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 2017, **26**(7): 3142–3155
- 15 Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: Proceedings of the 2015 IEEE Conference on Machine Learning. Miami, FL, USA: IEEE, 2015. 448–456
- 16 Maas A, Hannun A, Ng A. Rectifier nonlinearities improve neural network acoustic models. In: Proceedings of the 30th IEEE Conference on Machine Learning (ICML). Columbus, USA: IEEE, 2013.
- 17 He K M, Zhang X Y, Ren S Q, Sun J. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the 2015 IEEE Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 1026–1034
- 18 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: Proceedings of the 2014 International Conference for Learning Representations. San Diego, CA, USA: 2014. 1–14
- 19 He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 770–778
- 20 Marill T. Emulating the human interpretation of line-drawings as three-dimensional objects. *International Journal of Computer Vision*. 1991, **6**(2): 147–161
- 21 Christos T, Vasileios M, Ioannis P. Linear maximum margin classifier for learning from uncertain data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, **40**(12): 2948–2962
- 22 Gu S H, Zhang L, Zuo W M, Feng X C. Weighted nuclear norm minimization with application to image denoising. In: Proceedings of the 2014 Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014. 2862–2869
- 23 Daniel Z, Yair W. From learning models of natural image patches to whole image restoration. In: Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Barcelona, Spain: IEEE, 2011. 479–486
- 24 Schmidt U, Roth S. Shrinkage fields for effective image restoration. In: Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, USA: IEEE, 2014. 2774–2781



盖杉 南昌航空大学信息工程学院副教授. 主要研究方向为图像处理, 模式识别, 计算机视觉, 人工智能. 本文通信作者. E-mail: gaishan@nchu.edu.cn
(**GAI Shan** Associate professor at the School of Information Engineering, Nanchang Hangkong University. His research interest covers image processing, pattern recognition, computer vision, and artificial intelligence. Corresponding author of this paper.)



鲍中运 南昌航空大学信息工程学院硕士研究生. 主要研究方向为图像处理, 模式识别, 计算机视觉, 深度学习. E-mail: baozhongyun1234@163.com
(**BAO Zhong-Yun** Master student at the School of Information Engineering, Nanchang Hangkong University. His research interest covers image processing, pattern recognition, computer vision, and deep learning.)