

基于Gabor基算法展开卷积神经网络图像恢复模型

孙璐, 魏伟波, 杨光宇, 宋金涛, 李倩, 潘振宽*

青岛大学计算机科学技术学院, 山东 青岛

收稿日期: 2023年4月25日; 录用日期: 2023年5月24日; 发布日期: 2023年5月31日

摘要

基于变分的图像恢复模型可设计高效的即插即用算法, 但涉及较多的超参数调节, 且去噪器模块通常选择传统深度学习方法, 网络层数和训练参数较多。针对此问题, 在变分模型算法展开——半二次分裂方法基础上采用径向基函数近似非线性扩散模型的激活函数、采用Gabor基近似卷积核, 进而设计了高效紧致的图像恢复深度学习卷积神经网络。Gabor基比离散余弦变换基包含更丰富的方向、尺度信息, 可设计高效、紧致的网络结构。与非线性反应扩散模型相比, 模型针对图像去噪在噪声级为15、25、50对应的峰值信噪比高出0.14 dB、0.7 dB、0.7 dB, 在其它图像问题中也展现出良好的恢复效果。提出的框架不仅能够推广到其它图像恢复任务, 还可拓展到其它图像分析与处理的深度学习卷积神经网络设计。

关键词

图像恢复, 半二次分裂, Gabor基, 卷积神经网络, 算法展开

Image Restoration Based on Gabor—Based Algorithm Unrolling Convolutional Neural Network

Lu Sun, Weibo Wei, Guangyu Yang, Jintao Song, Qian Li, Zhenkuan Pan*

College of Computer Science & Technology, Qingdao University, Qingdao Shandong

Received: Apr. 25th, 2023; accepted: May 24th, 2023; published: May 31st, 2023

Abstract

The image restoration model based on variational can design an efficient plug and play algorithm,

*通讯作者。

文章引用: 孙璐, 魏伟波, 杨光宇, 宋金涛, 李倩, 潘振宽. 基于 Gabor 基算法展开卷积神经网络图像恢复模型[J]. 计算机科学与应用, 2023, 13(5): 1119-1134. DOI: 10.12677/csa.2023.135110

but it involves more hyperparameter adjustment, and the denoising module usually chooses the traditional deep learning method, which has more network layers and training parameters. To solve this problem, an efficient and compact deep learning convolutional neural network for image restoration is designed by using radial basis function to approximate the activation function of nonlinear diffusion model and Gabor basis to approximate convolution kernel on the basis of variational model algorithm expansion—semi-quadratic splitting method. Gabor contains more direction and scale information than discrete cosine transform function, which can design efficient and compact network structure. Compared with the nonlinear reaction-diffusion model, the peak signal-to-noise ratio of the model for image denoising is 0.14 dB, 0.7 dB and 0.7 dB higher when the noise level is 15, 25 and 50, which also shows a good recovery effect in other image problems. The proposed framework can be extended not only to other image restoration tasks, but also to deep learning convolutional neural network design for image analysis and processing.

Keywords

Image Restoration, HQS, Gabor-based, Convolutional Neural Network, Algorithm Unrolling

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

图像恢复是计算机视觉、图像处理的基本问题之一，是图像处理与分析的基础，其目的是依据图像退化模型由观测图像恢复出理想图像。图像退化模型通常为图像出现模糊、失真的条件下附加额外噪声。作为非适定性的逆问题，其解不唯一，而变分法是解决这种病态问题的有效的经典方法[1]。该方法作为图像处理领域中的一类基础性方法，具有完善的数学理论体系，可通过最小化能量泛函对问题求解。对于能量泛函的求解，存在多种方法。Perona and Malik 采用梯度下降法对非线性扩散方程求解[2]。为简化计算、提高计算效率，Chalasanani 等设计了临近梯度下降法(PG: Proximal Gradient Method) [3]，Chan 等设计了交替方向乘子法(ADMM: Alternating Direction Method of Multiplier) [4]，Ono 设计了原始对偶变量方法(PD: Primal-Dual Method) [5]。这些优化方法具有较好的可解释性，且在执行过程中无需大量数据便可获得有效结果。但该类方法往往需要引入可调整的惩罚参数，模型的计算需反复迭代完成。

基于变分模型算法展开策略设计深度学习神经网络是解决上述困难的有效途径[6]。Gavaskar 等将迭代收缩阈值算法(ISTA: Iterative Shrinkage Thresholding Algorithm)展开为具有可学习参数的深度神经网络(PnP-ISTA: Plug-and-Play Iterative Shrinkage Thresholding Algorithm)以解决线性逆问题[7]。Zhang 等训练一组快速有效的卷积神经网络去噪器作为模型先验，利用半二次分裂法(HQS: Half Quadratic Splitting)将其集成到基于模型的优化方法中，以解决不同的逆问题[8]。该算法因解耦保真度项和正则化项带来的简单性和快速收敛性被大量应用，解耦之后可得到规则项相关的高斯去噪器。

对于去噪器，可采用成熟模型作为先验应用到网络中，如 Zhang 等基于多个卷积与激活块结合，提出了利用 IRCNN 设计去噪器的方法[8]，又通过残差网络与批量归一化设计了 DnCNN 框架[9]；在 FFDNet 增加噪声映射图的基础上[10]，DRUNet 模型将残差块整合到 U-Net 中，实现有效的去噪先验建模[11]；为确保观测一致性，DPDNN 设计了一个简洁的 U-Net 子网络作为去噪器[12]，并与图像重建模块共同组成图像恢复的深度学习神经网络。此类模型不仅利用了 DNN 强大的去噪能力，即插即用的特性使得参数

只需训练一次便可直接应用到其它图像任务中。尽管如此,这种经典的深度学习模型往往需要大量小尺寸卷积核与传统激活函数 ReLU (Linear Rectification Function)等来实现非线性近似,其层数较深,参数量较大。

上述变分模型的规则项与其梯度降偏微分方程的扩散项对应,通过对扩散项的有效近似可设计紧凑、高效的深度学习网络。Chen 等引入专家场模型(FoE: Fields of Experts)作为规则项[2] [13],并将相应的非线性反应扩散模型中的扩散项用流函数表示,该流函数可设计为用径向基函数近似的神经网络激活函数。根据传统差分算子与卷积核的对应关系,文献[2]还基于离散余弦变换(DCT: Discrete Cosine Transformation)基设计了不同通道中的卷积核[14]。所设计的可训练非线性反应扩散网络(TNRD: Trainable Nonlinear Reaction Diffusion)具有较好的可解释性,并具紧致、高效特点,但 DCT 基仅包含图像变化的频谱特性,不能反映图像变化的尺度、方向等信息,而 Gabor 基函数包含了上述三类信息,并已在图像识别等应用的特征提取方面显现优越的性能[15] [16] [17]。

受此启发,本文将采用 Gabor 基函数设计卷积核,并在变分模型的半二次分裂迭代算法展开基础上设计端到端的图像恢复深度学习神经网络,其中的激活函数采用通用的高斯径向基函数近似。

2. 基于 Gabor 基与算法展开卷积神经网络的图像恢复模型

2.1. 图像恢复模型

图像退化模型为:

$$y = Ax + n \quad (1)$$

其中, y 为观测图像, x 为理想图像, n 为噪声, A 代表不同的退化问题。当 A 是子采样矩阵算子时,图像恢复问题为超分辨率问题;当 A 为模糊矩阵算子时,图像恢复问题为去模糊问题;当 A 为单位矩阵时,图像恢复问题为去噪问题。

公式(1)对应的图像恢复变分模型为:

$$x = \arg \min \left\{ \frac{1}{2} \|y - Ax\|_2^2 + \lambda J(x) \right\} \quad (2)$$

其中, $\frac{1}{2} \|y - Ax\|_2^2$ 为数据项,表示图像恢复前后的相似程度, $J(x)$ 代表规则项,决定图像的恢复质量, λ 为惩罚参数。

2.2. 半二次分裂框架

引入辅助变量 v , 令 $v = x$, (2)转化为:

$$(x, v) = \arg \min_{x, v} \left\{ \frac{1}{2} \|y - Ax\|_2^2 + \frac{\eta}{2} \|x - v\|_2^2 + \lambda J(v) \right\} \quad (3)$$

η 为新引入的惩罚参数,对公式(3)交替优化得:

$$\begin{cases} x^{t+1} = \arg \min_x \left\{ \frac{1}{2} \|y - Ax\|_2^2 + \frac{\eta}{2} \|x - v^t\|_2^2 \right\} \\ v^{t+1} = \arg \min_v \left\{ \frac{\eta}{2} \|x^{t+1} - v\|_2^2 + \lambda J(v) \right\} \end{cases} \quad (4)$$

其中第一个子优化问题的梯度下降法求解公式为:

$$x^{t+1} = x^t - \delta \left[A^T (Ax^t - y) + \eta (x^t - v^t) \right] \quad (5)$$

第二个子优化问题对应非线性扩散过程。当采用 FOE 规则项

$$J(v) = \sum_{i=1}^{N_k} \langle \phi_i(|K_i v|), 1 \rangle \tag{6}$$

时，其解与以下扩散模型对应

$$\frac{\partial v}{\partial t} = -\lambda \sum_{i=1}^{N_k} K_i' \left(\frac{\phi_i'(|K_i v|)}{|K_i v|} K_i v \right), v^{t+1}(x, 0) = x^{t+1}(x) \tag{7}$$

流函数表示为 $\phi_i(K_i v) = \frac{\phi_i'(|K_i v|)}{|K_i v|} K_i v$ ，并对时间变量差分采用显式 Euler 法得到与残差网络结构对应的迭代格式：

$$v^{t+1} = v^t - \Delta t \lambda \sum_{i=1}^{N_k} (K_i^t)^T \phi_i'(K_i^t v^t) \tag{8}$$

其中， K_i^t 和 ϕ_i^t 分别代表基于 Gabor 基函数和高斯径向基函数设计的卷积核、激活函数。 N_k 为卷积核个数， t 代表网络层数。

2.3. 基于 Gabor 基函数的卷积核设计

为克服傅里叶变换全局频域分析的缺点[18]，Gabor 提出了称之为 Gabor 变换的加窗傅里叶变换。这是一种正弦函数调谐的高斯函数，即将信号划分出多个小的时间间隔，并通过傅里叶变换逐一进行处理，来确定信号在该时间间隔存在的频率。Gabor 核函数可以在保留空间关系信息(时域)的同时，表示图像中的空间频率结构[17]。定义为：

$$\Psi_{\mu, \nu}(z) = \frac{\|k_{\mu, \nu}\|^2}{\sigma^2} e^{-\left(\|k_{\mu, \nu}\|^2 \|z\|^2 / 2\sigma^2\right)} \left[e^{ik_{\mu, \nu} z} - e^{-\sigma^2/2} \right] \tag{9}$$

$$z = (x, y), k_{\mu, \nu} = k_\nu e^{ik_\mu}, k_\nu = (\pi/2) / \sqrt{2}^{(\nu-1)}, k_\mu = \mu \frac{\pi}{U}, \nu = 0, 1, \dots, V; \mu = 0, 1, \dots, U$$

其中， μ 、 ν 分别代表方向和尺度参数，分别控制核信息的旋转角度与频率， $\|\cdot\|^2$ 为 \cdot 的 L_2 范数， $\sigma = 2\pi$ 。

为保证不同通道的基函数之间具有足够差异性，在卷积时能够获取更多图像信息，本文将 ν 和 μ 嵌入到同一层网络的基函数中。实验针对不同方向尺度的组合设计了 48 组基函数，其中，选择 n 个不同的尺度，方向在 $0 \sim \pi$ 中等距取值，个数为 $48/n$ 。

若所选择的 Gabor 基函数组为 $\Psi_j, j = 1, 2, \dots, 48$ ，则不同网络层的各通道设计的卷积核为

$$K_i^t = \sum_{j=1}^{48} \frac{\alpha_{i,j}^t}{\|\alpha_i^t\|_2} \Psi_j \tag{10}$$

其中， $\alpha_{i,j}^t$ 为训练参数， $\|\alpha_i^t\|_2 = \sqrt{\sum_{j=1}^{48} (\alpha_{i,j}^t)^2}$ 。

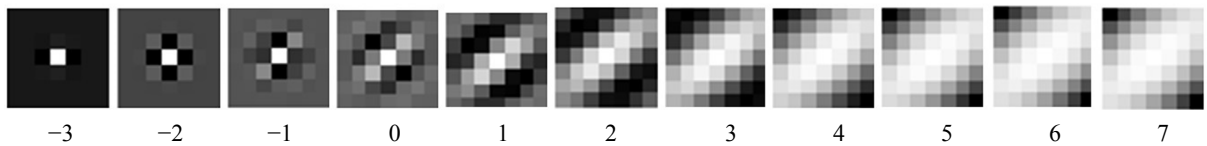


Figure 1. Gabor basis functions corresponding to different scales (-3~7)

图 1. 不同尺度(-3~7)对应 Gabor 基函数

为选择适当的 Gabor 基函数组, 暂固定角度, 并将不同尺度对应的基展示如图 1。直观的观察可发现: 当尺度在-2~4 范围内信息变化较为明显, 而随着尺度的增加, 它们间的差异逐渐变小; 当尺度为-3 时, 核含信息较少, 无法携带足够的方向等信息。所以, 下面的实验将在-2 到 4 的范围附近选择尺度。以上的可视化思想在 Zhang 的研究中针对图像分类设计的神经网络中应用, 能够提取更多图像特征[17]。

为调节 Gabor 基函数所涉及的权重以自适应网络, 在编码卷积时, 使用相同维度的学习滤波器与 Gabor 函数组通过凸组合生成特征卷积核。学习滤波器作为参数矩阵在反向传播时独立学习, 基函数之间存在信息差异以及函数组的网络共享使得在相同网络层之间, 不同通道的卷积核互不相同; 不同网络层之间, 卷积核特征相似但相互独立。如图 2 所示。

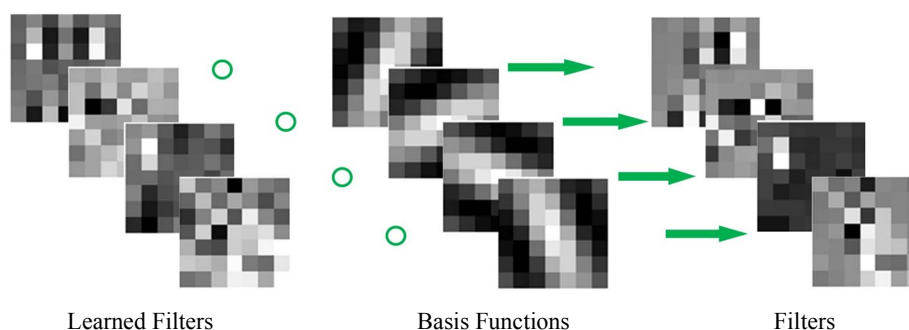


Figure 2. Design process based on Gabor-based convolution kernel
图 2. 基于 Gabor 基卷积核的设计过程

2.4. 基于高斯径向基的激活函数近似

考虑到高斯径向基函数(RBFs: Radial Basis Functions)的简单、通用, 且在学习过程收敛速度快、泛化能力强等特点[2], 本文用高斯径向基函数的组合近似激活函数, 其公式为:

$$\phi_i^t(z) = \sum_{j=1}^M \omega_j^t \varphi\left(\frac{|z - \mu_j|}{\gamma_j}\right) \quad (11)$$

$$\varphi\left(\frac{|z - \mu|}{\gamma}\right) = \exp\left(-\frac{(z - \mu)^2}{2\gamma^2}\right) \quad (12)$$

其中, M 是总的插值点个数, j 为插值点, i 代表通道数, 标准差为 $\gamma = \frac{2I_{\max}}{M-1}$, μ 对应每个基函数的中心点, 以等距的方式分布在 $[-I_{\max}, I_{\max}]$ 中。

2.5. 损失函数

在有监督的方式下, 通过输出图像与真实图像之间的差异设计损失函数来训练网络。给定 x^T 为网络输出, y^{st} 为真实清晰图像, 得到损失函数:

$$L(x^T, y^{st}) = \frac{1}{2} \|x^T - y^{st}\|_2^2 \quad (13)$$

2.6. 网络设计

图 3 为公式(6)所对应整个网络层的算法流程。针对其中的扩散模块, 即公式(9), 其子网络如图 4 所示。

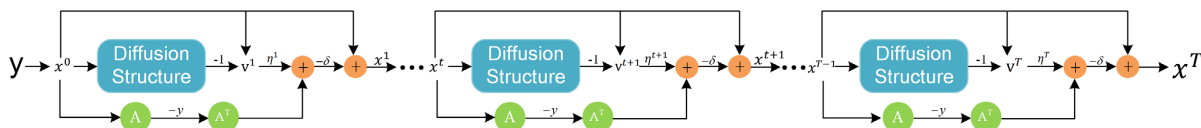


Figure 3. The entire feedforward network process
图 3. 整个网络算法流程

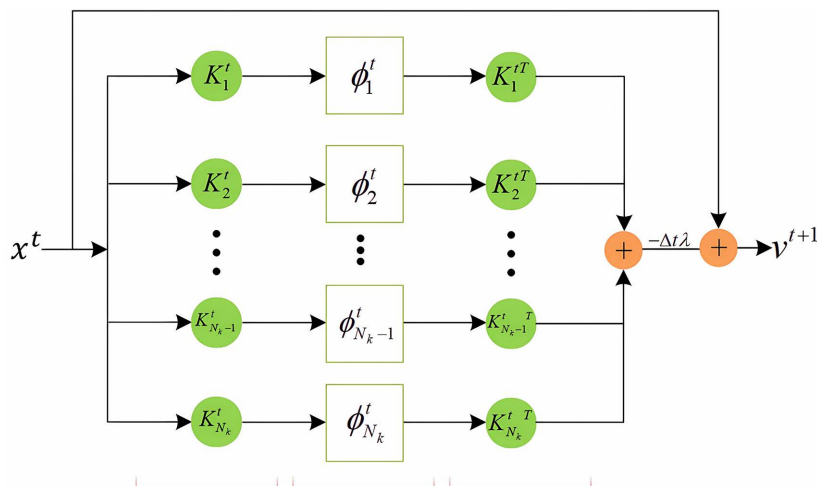


Figure 4. Diffusion structure based on Gabor basis function
图 4. 基于 Gabor 基函数的扩散结构

3. 数值实验中网络超参数分析

本文模型训练与测试均在 NVIDIA GeForce TEX 2080 Ti 版本的 GPU 上进行。由于卷积操作会缩小图像，为保证输出图像大小与原图像一致，实验选取对称填充的方法复制边缘像素将图像长宽分别扩大 12 像素。

对于其它参数的设置，相邻两个 Epoch 选择较高峰值信噪比(PSNR: Peak Signal-to-Noise Ratio)进行参数保存；每三个 Epoch 降低一半的学习率，初始学习率为 $lr = 1e-3$ 。

3.1. 层数与卷积核大小

在应用图像处理任务之前，以 Set12 为训练测试集来选择合适的网络层数与卷积核大小[12]。下图 5 是相关参数的 PSNR 折线图。为控制变量，当比较卷积核时，固定层数为 5 层，而选择层数时，固定卷积核大小为 7×7 。针对卷积核的大小，由图(a)看出，感知野越大，网络的性能越好，但当卷积核大于 7×7 时，性能无法再得到有效提升。对于网络层数的选择，根据图(b)发现，随着层数的不断加深，PSNR 持续提高，而对应层数为 5 时，PSNR 提升幅度最大，故根据性能与效率的综合衡量，以下实验将设置卷积核尺寸为 7×7 ，网络层数为 5 层。

3.2. 基函数设置

对于一个 DCT 特征的卷积核，其能获取图像在某个频率邻域响应情况，Gabor 作为基于 DCT 的加窗傅里叶函数，在频率信息的基础上，加入了尺度、方向等信息，执行卷积操作时能够提取更多图像特征，实现良好的先验形式。图 6 中的(a)与(b)分别为两种基函数对应训练后的卷积核。由观察看出，(a)存在一定数量的小作用域卷积核，对应图(b)较少，(b)中更多的是高阶求导的卷积核，并且具有充足的方向、尺度等信息。因此 Gabor 基在信息提取方面有着更高的使用价值。

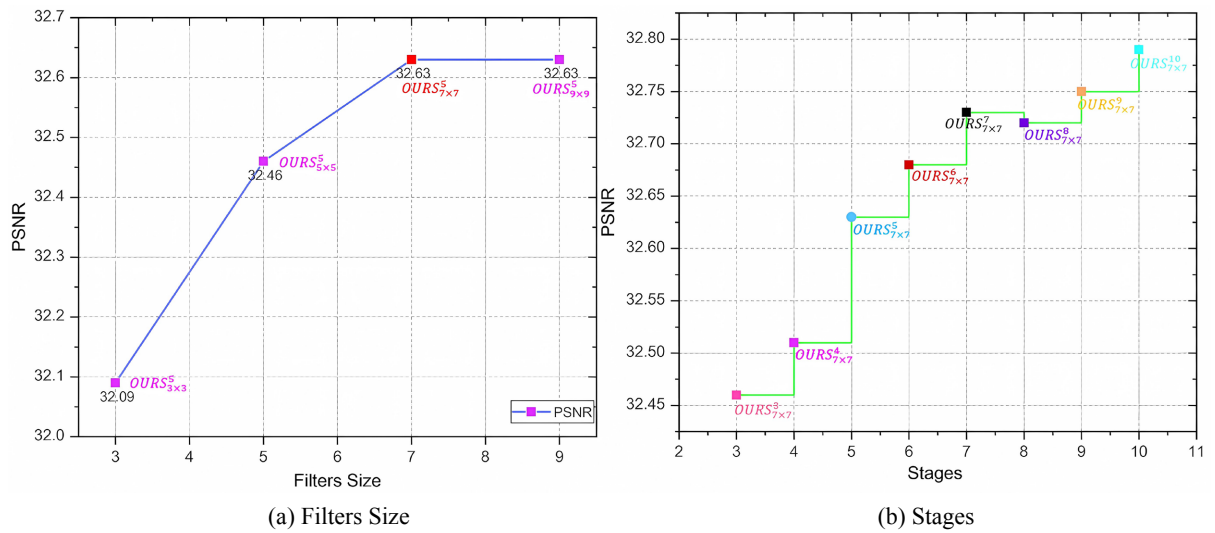
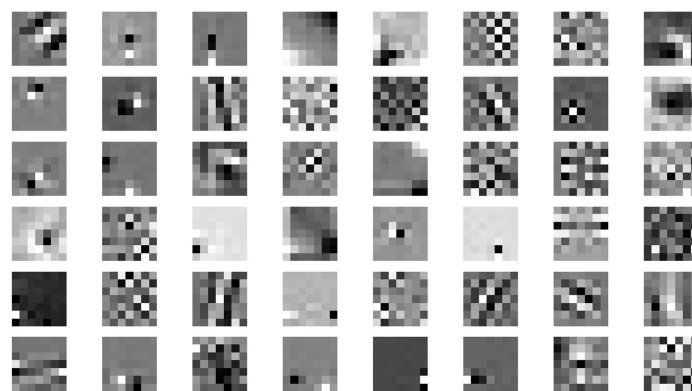
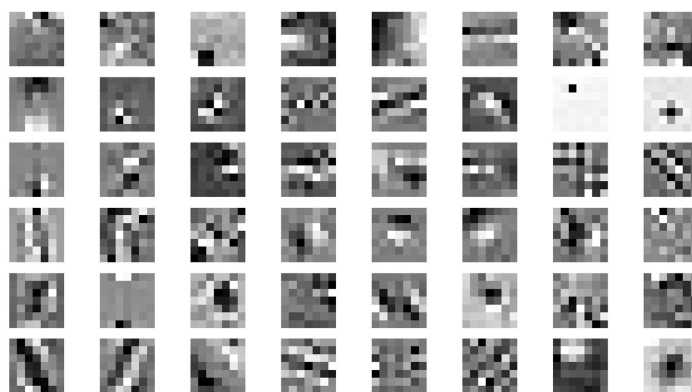


Figure 5. Performance comparison of different size convolution kernels and different network layers
 图 5. 不同大小卷积核与不同网络层数的性能比较



(a) 基于 DCT 的卷积核



(b) 基于 Gabor 的卷积核

Figure 6. Display of convolution kernel of DCT and Gabor in the first stage
 图 6. DCT 与 Gabor 基的卷积核在第一层框架的图像展示

为比较两者性能，以 TNRD 为框架，用 Set12 作为训练集，BSD68 为测试集进行实验验证[13]。由表 1 的实验数据可知，由 Gabor 基得到的 PSNR 值基本高于 DCT 函数。

Table 1. The performance comparison of Gabor and DCT basis functions and Gabor basis functions corresponding to different parameters at noise level $\sigma = 15$ **表 1.** Gabor 与 DCT 基函数以及不同参数对应的 Gabor 基函数在噪声级 $\sigma = 15$ 的性能比较

Methods	Filters	RBFs	Scales	Directions	BSD68	
Gabor	48	31	8	[0, 1, 7]	6	31.13
	48	31	8	[1, 1, 8]	6	31.14
	48	31	8	[-1, 1, 6]	6	31.12
	48	31	8	[-2, 1, 5]	6	31.16
	48	31	6	[-2, 1, 3]	8	31.13
	48	31	6	[-1, 1, 4]	8	31.11
	48	31	6	[0, 1, 5]	8	31.14
	48	31	6	[1, 1, 6]	8	31.16
	48	31	6	[2, 1, 7]	8	31.14
	48	31	12	[-3, 1, 8]	4	31.10
	48	31	4	[-1, 1, 2]	12	31.06
	48	31	4	[0, 1, 3]	12	31.13
	48	31	4	[1, 1, 4]	12	31.13
	48	31	4	[2, 1, 5]	12	31.11
	48	31	4	[-1, 2, 5]	12	31.14
	48	31	4	[0, 2, 6]	12	31.14
	48	31	4	[1, 2, 7]	12	31.13
DCT	48	31	-	-	-	31.07

此外, 本文还比较了不同参数对应的 Gabor 核的去噪能力。为充分筛选参数, 将尺度、方向个数分别设置为 8、6, 6、8 和 4、12。其中, 尺度大体范围设置在 $-1 \sim 5$, 视其个数决定间隔大小。方向参数在 0 到 π 范围内等距选择。设置 $[a, b, c]$, 代表在 $a \sim c$ 范围中, 每间隔 b 选择一个尺度。据表 1 的实验数据, 当尺度范围为 $[1, 1, 6]$ 、方向为 8 与尺度为 $[-2, 1, 5]$ 、方向为 6 时, PSNR 值达到最高。

为防止选择参数过程中易出现相同的基函数, 本文所有实验的参数均建立在尺度与频率呈正比的关系上。对于不同的恢复任务, 为调节出最优参数, 可根据性能在一定范围内改变两者的正比关系以确保获得更多频率信息。

4. 实验结果与分析

为保证参数量一致以及排除其它因素的影响, 凸显 TNRD 与 GBAUCNN 即本文模型的可比性, 以下关于 TNRD 的实验数据均为自己训练得出, 其中设置参数量与 GBAUCNN 相等, 包括径向基函数 31 个、卷积核 48 个。

4.1. 图像去噪

本文使用文献[2]中 400 张 180×180 的图像作为训练集, 进行裁剪、随机旋转等预处理操作, 生成 40×40 像素图像块作为输入样本。其中, 旋转包括上下翻转、旋转 180 度、270 度以及逆向旋转 90 度。为评估模型的去噪性能, 采用广泛用于高斯去噪测试的数据集 Set12 和 Berkeley 分割数据集 BSD68 作为

标准测试数据集[13]。

实验主要针对高斯噪声的去除。对于单级别去噪网络,将图像分别加入 $\sigma = 15, 25$ 和 50 的高斯噪声。为验证网络有效性,对比实验选择几种较为先进的去噪算法,如 BM3D、EPLL、CSF、MLP 和 TNRD [2] [19] [20] [21] [22], 并且以峰值信噪比(PSNR)作为客观评价标准。

由表 2、表 3 数据可得, TNRD 模型在各个噪声水平下的性能明显优于 BM3D、EPLL、CSF 传统模型。在 Set12 中, MLP 与 TNRD 基本持平。本文模型 GBAUCNN 能得到比 TNRD 更高的 PSNR。根据表 2 的数据, 当噪声级为 15、25 和 50 时, GBAUCNN 的 PSNR 分别提高 0.14 dB、0.07 dB 和 0.07 dB。图 7、图 8 和图 9 为算法去噪结果的展示。 $\sigma = 15$ 时, 根据图 7 框中图像的放大效果, (b)没有保留住城堡的整体轮廓, (a)、(c)相比(b)虽保留了图像信息, 但没有实现良好的去噪。(d)的去噪效果优于(a)、(b)和(c), 但依旧有一些噪声残余, 而(e)能够在保留图像信息的基础上更好的去除噪声, 获得更平滑的图像。对于 $\sigma = 25$, 由图 8 看出, EPLL、BM3D、CSF、WNNM 等传统模型能实现一定的去噪效果, 但得到的边缘模糊。TNRD 的去噪图像虽细节清晰, 但是信息保留较少。GBAUCNN 在去噪的同时边缘细节保留得更加细致充足, 达到了更好的视觉效果。噪声级为 50 时, 根据图 9, TNRD 获得的图像(c)相对于(a)和(b)而言更加清晰, 但框中的人物面部细节依旧不足, (d)得到的面部表情则更加清楚自然。由此可知, 以 Gabor 基为卷积先验的基础, GBAUCNN 模型在实现去噪的同时能够获得更清晰的边缘信息。

Table 2. Comparison for PSNR of noise removal for Set12 by different algorithms

表 2. 不同算法对 Set12 消除噪声的 PSNR 比较

Method	BM3D	EPLL	CSF	MLP	TNRD	GBAUCNN
$\sigma = 15$	32.37	32.13	32.31	-	32.40	32.54
$\sigma = 25$	29.96	29.69	29.83	30.02	30.00	30.07
$\sigma = 50$	26.72	26.47	-	26.78	26.77	26.84

Table 3. PSNR of different models on dataset BSD68 at different noise levels

表 3. 不同模型对数据集 BSD68 在不同噪声级别上的 PSNR

Method	BM3D	CSF	EPLL	WNNM	TNRD	GBAUCNN
$\sigma = 15$	31.08	31.24	31.19	31.37	31.36	31.46
$\sigma = 25$	28.56	28.74	28.68	28.83	28.90	28.95
$\sigma = 50$	25.62	-	25.68	25.87	25.99	26.02

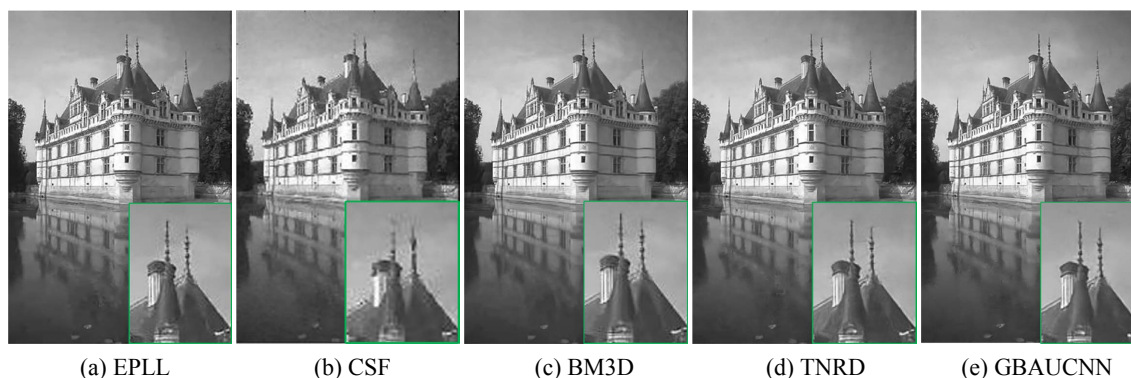


Figure 7. Comparison of denoising with different algorithms at noise level $\sigma = 15$

图 7. 不同算法在噪声级 $\sigma = 15$ 中去噪对比图

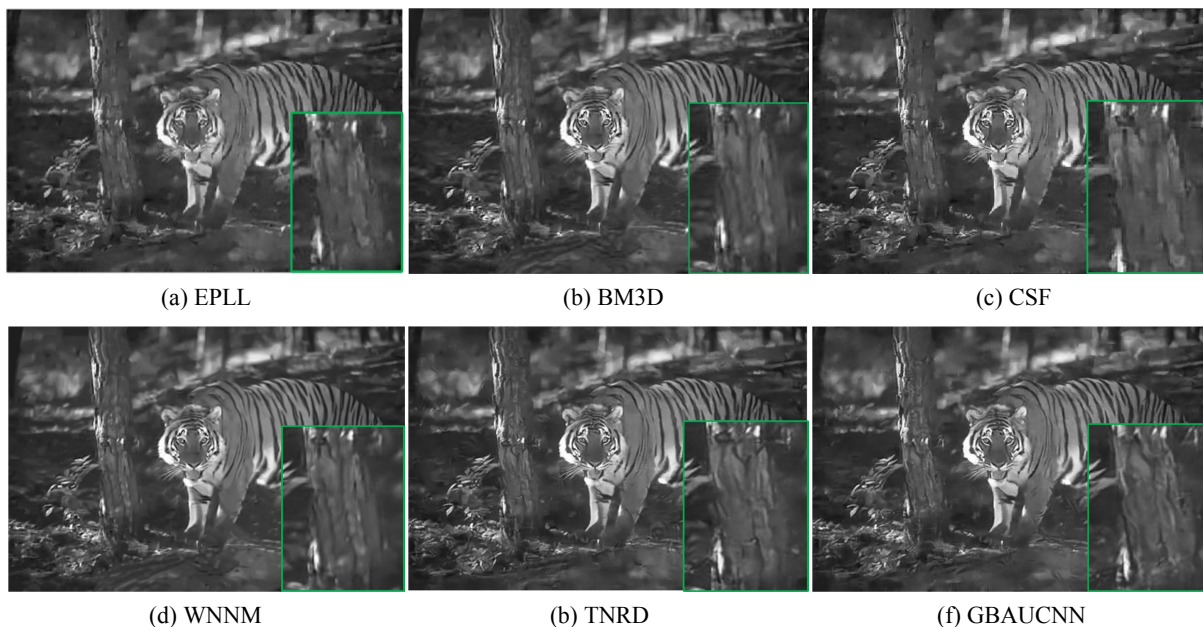


Figure 8. Comparison of denoising with different algorithms at noise level $\sigma = 25$
图 8. 不同算法在噪声级 $\sigma = 25$ 中去噪对比图

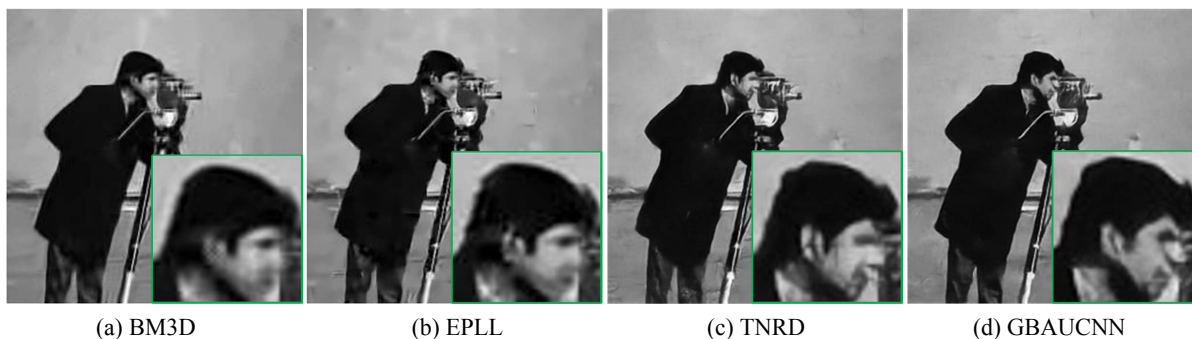


Figure 9. Comparison of denoising at noise level $\sigma = 50$ with different algorithms
图 9. 不同算法在噪声级 $\sigma = 50$ 中去噪对比图

由于所提出框架的广泛适用性以及以 Gabor 为基础的扩散过程具有强化边缘、清晰图像的效果，作为扩展，本文模型将应用到图像超分辨率、图像去模糊与彩色去噪任务中以验证性能。

4.2. 图像超分辨率

对于退化矩阵 A ，引入双三次插值的方法实现[23]。考虑到各邻点间灰度值变化率的影响，该算法利用待求像素坐标反变换得到的浮点坐标周围 16 个邻近像素的权重，与对应像素点的灰度值作加权线性求和得到。双三次插值是一种较为复杂的插值方式，它能创造出比双线性插值更平滑的图像边缘，得到更接近高分辨率的图像效果。

在训练模型时，选用开源超分辨率重建数据集 DIV2K 中的 800 张训练数据与对应比例因子分别为 2、3 的低分辨率数据作为 HR/LR 训练集。本次实验只针对图像的亮度分量，因此，在预处理时，选择图像由 RGB 模式转化成 YUV 模式的第一维作为输入。实验采用[12]中的数据 Set5、Set14、Urban100 和 BSD100 对模型进行评估。

表 4 为四种算法在超分辨率问题上性能的比较，包括 PSNR 与结构相似性(SSIM: Structural Similarity)。

通过分析可以看出, Bicubic 与 SR-CNN 模型得到的 PSNR 值较低[24], 而 TNRD 略有优势, GBAUCNN 模型在 TNRD 的基础上能得到更高的 PSNR 值。特别的, 当比例因子为 2 时, 表中各个数据集对应的 PSNR 指标分别高出 0.08 dB、0.05 dB、0.15 dB 以及 0.06 dB。对于 SSIM 指标, 几种算法生成的图像结构与原图像相似度大体相同, 这意味着输出图像的结构在没有失真基础上都达到了较好的图像处理效果。图 10 为四种算法的图像展示, 根据整体图像与框中放大部位的鸟喙处, SR-CNN 算法对应图像比 Bicubic 要更加清晰, 但与 TNRD 和 GBAUCNN 相比则较为模糊。本文模型依旧保留了更多的细节纹理, 图像效果更加清晰真实。

Table 4. Comparison of the average performance (PSNR, SSIM) of different methods on the test set Set5, Set14, Urban100 and BSD100

表 4. 不同方法在测试集 Set5、Set14、Urban100 以及 BSD100 平均性能(PSNR, SSIM)的比较

Dataset	Scaling factor	SR-CNN		Bicubic		TNRD		GBAUCNN	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Set5	2	36.66	0.95	33.97	0.94	37.20	0.96	37.28	0.96
Set14		32.42	0.91	30.31	0.89	32.65	0.92	32.70	0.92
Urban100		29.50	0.90	26.85	0.86	29.60	0.91	29.75	0.92
BSD100		31.36	0.89	29.74	0.86	31.60	0.90	31.66	0.90
Set5	3	32.75	0.91	30.62	0.88	33.35	0.93	33.39	0.93
Set14		29.28	0.82	27.56	0.80	29.38	0.84	29.43	0.85
BSD100		28.41	0.79	27.31	0.77	28.63	0.81	28.66	0.81

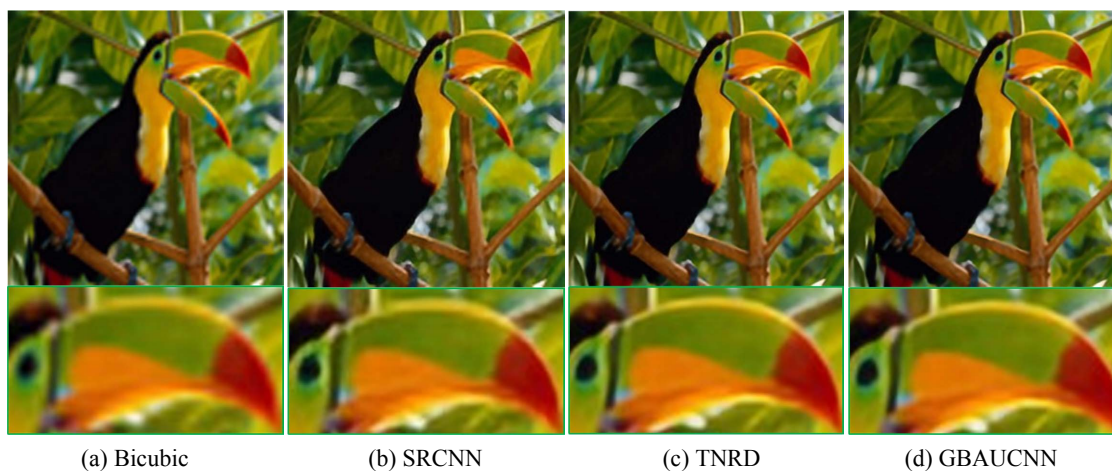


Figure 10. Results of different algorithms for Set5 images

图 10. 不同算法针对 Set5 图像的展示结果

4.3. 图像去模糊

为进一步验证模型的通用性, 将实现 TNRD 与 GBAUCNN 的非盲去模糊实验。采用标准偏差为 1.2、1.6 以及 2.0, 大小为 25×25 的高斯模糊核进行实验。对于模糊之后的图像, 加入标准差为 2 的高斯噪声得到输入图像。本文采取边训练边处理的预处理方式, 将图像随机裁剪成 64×64 的图像块并模糊。选取 DIV2K 中的 800 张灰度图像作为训练数据, Set10 作为测试数据, 以此评估模型性能。

本文将 GBAUCNN 与 IDD-BM3D [25]、EPLL、TNRD 进行性能的比较。表 5 列举了模糊核标准差

分别为 1.2、1.6 和 2.0 的实验数据，其中，固定噪声级为 2。根据表中的结果，EPLL 与其它模型相比数据最低，TNRD 比 IDD-BM3D 占优势。对于这三种不同的模糊核，GBAUCNN 的峰值信噪比高 TNRD 模型平均 0.1 dB 左右。

Table 5. Comparison of the average performance of different methods on the test dataset Set10
表 5. 不同方法在测试集 Set10 平均性能的比较

Method	Butterfly	Peppers	Parrort	Starfish	Barbara	Boats	Man	House	Leaves	Lena	Average
Gaussian Blur with standard deviation 1.2											
IDD-BM3D	32.17	31.74	34.55	32.90	31.92	33.33	30.18	35.60	33.18	33.12	32.60
EPLL	28.99	28.17	32.75	30.57	28.41	29.84	28.31	31.45	28.22	30.90	29.76
TNRD	32.77	30.65	35.32	33.75	30.96	33.34	30.39	35.64	33.02	33.81	32.97
GBAUCNN	33.00	30.75	35.47	33.90	31.20	33.39	30.42	35.68	33.29	33.90	33.10
Gaussian Blur with standard deviation 1.6											
IDD-BM3D	29.79	29.64	31.90	30.57	25.99	31.17	27.68	33.56	30.13	30.91	30.13
EPLL	25.78	26.73	31.32	28.52	24.22	28.84	26.57	31.76	25.29	29.46	27.85
TNRD	30.52	29.70	32.43	31.27	24.80	31.16	28.06	33.39	30.22	31.14	30.27
GBAUCNN	30.59	29.71	32.51	31.30	24.96	31.18	28.11	33.63	30.26	30.31	30.36
Gaussian Blur with standard deviation 2.0											
IDD-BM3D	28.21	28.38	30.35	28.90	24.36	29.72	26.30	32.39	27.93	29.50	28.60
EPLL	21.95	22.31	28.77	25.50	23.05	24.62	25.09	27.49	21.13	25.87	24.58
TNRD	29.01	28.47	30.89	29.58	24.38	29.80	26.80	32.28	28.23	29.80	28.93
GBAUCNN	29.03	28.31	31.19	29.74	24.40	29.85	26.75	32.58	28.42	29.95	29.02

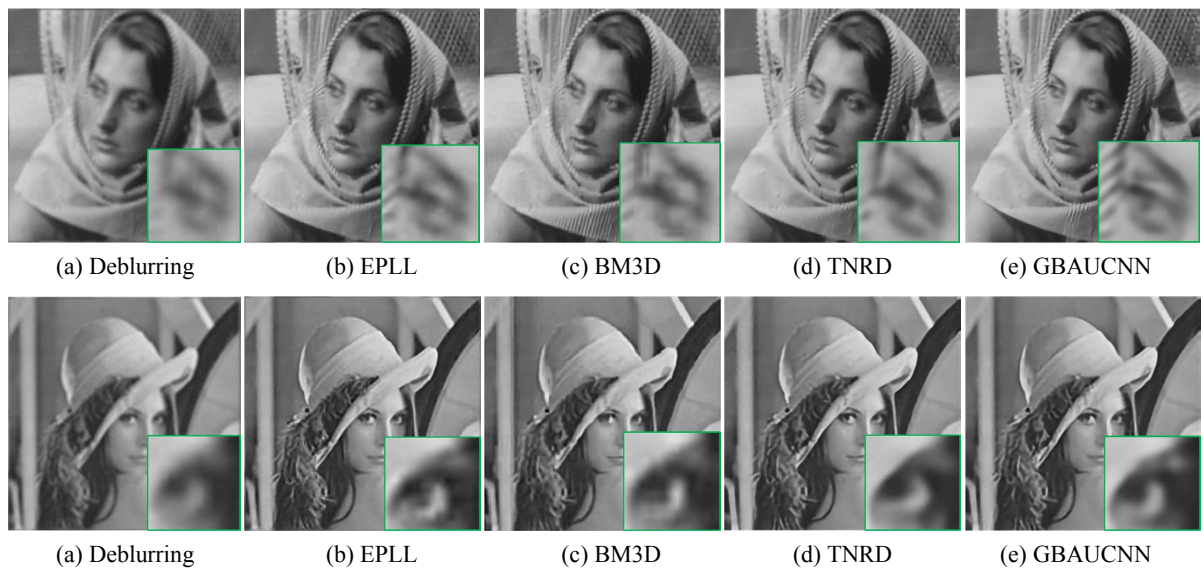


Figure 11. Effects of different algorithms on blurred images
图 11. 不同算法对于模糊图像的效果展示

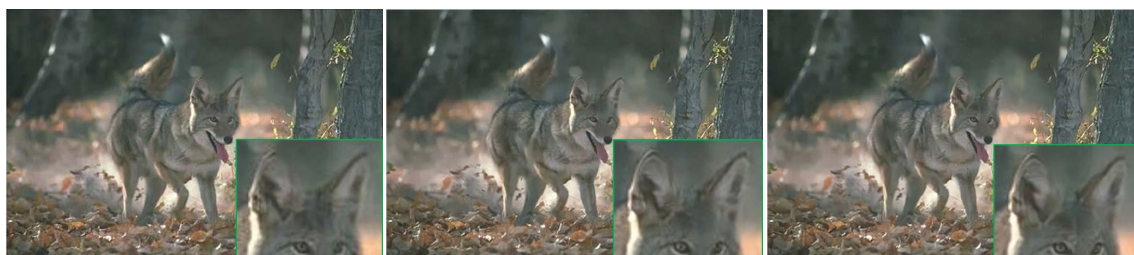
灰度图像上针对高斯模糊的效果展示如图 11。在第一幅图的比较中，EPLL 与 BM3D 在眼睛处略微

模糊，整体图像也不够平滑。TNRD 虽能得到较为光滑的图像，但框中的放大部分依旧不够清晰，GBAUCNN 能够维持图像细节，在整体良好的视觉效果基础上保持足够的边缘信息。对于 Lena 图像的去模糊结果，可以看到(c)、(d)的面部表情稍微变形，(b)、(e)能较好地维持，且(e)图像具有更加清晰真实的视觉效果。

4.4. 彩色图像去噪

与灰度图像不同，彩色图像存在 3 个通道 R、G、B，而本文执行的特征图固定通道为 48，所以在操作之前，用大小为 1×1 的卷积核，将图像从 3 通道卷积成 48 通道，同时图像大小保持不变。为与真实的清晰图像对应，在去噪后的输出操作中用相同的方法将特征图卷积成通道为 3 的输出图像。本文以 Berkeley 分割数据集中的 200 张自然图像作为训练集，采用与图像去噪同样的预处理方式得到 40×40 的 729,000 个图像块，并随机进行图像旋转、翻转等操作实现图像增强。以下数据将采用 CBSD68 数据集与 Kodak24 作为测试集进行模型之间的比较。其中，CBSD68 为伯克利分割数据集和基准的一部分，共包含 68 张图像，是用于图像去噪基准的彩色数据集，而 Kodak24 为数字图像处理常用数据集，大小为 500×500 的 24 张彩色图像。

$\sigma = 15$



(a) CBM3D

(b) TNRD

(c) GBAUCNN

$\sigma = 25$



(a) CBM3D

(b) TNRD

(c) GBAUCNN

$\sigma = 50$



(a) CBM3D

(b) TNRD

(c) GBAUCNN

Figure 12. Image effects of different models on CBSD68 dataset at different noise levels

图 12. 不同模型针对数据集 CBSD68 在不同噪声级上的图像效果

由表 6 中的数据可知, 对于以上两种数据集, GBAUCNN 对比 TNRD 有着明显的优势。在 CBS68 中, GBAUCNN 模型在不同噪声水平下高于 TNRD 模型 0.06 dB、0.13 dB 和 0.09 dB, 而对于 Kodak24 数据集能高出 0.24 dB、0.13 dB 以及 0.10 dB。根据图 12 的图像对比展示, 噪声级为 15、25 时, 从整体的视觉效果看, 三种模型都能得到较好地去噪效果, 但 CBM3D 去噪并不够彻底[19], 放大部分依旧残留少量噪声, TNRD 与 GBAUCNN 在此方面更具优势, GBAUCNN 还能获得更多的纹理信息。对于噪声级为 50 的图像, TNRD、GBAUCNN 在清晰度与亮度方面均强于 CBM3D, 两者对于豹纹具有较高的边缘保持。同时, 本文模型去噪效果更加明显, 图像背景更加光滑。

Table 6. Comparison of PSNR values of different models at different noise levels
表 6. 不同算法在不同噪声级的 PSNR 值比较

Dataset	Methods	$\sigma = 15$	$\sigma = 25$	$\sigma = 50$
CBS68	CBM3D	33.47	30.69	27.37
	TNRD	33.65	30.85	27.57
	GBAUCNN	33.71	30.98	27.66
Kodak24	TNRD	34.13	31.66	28.38
	GBAUCNN	34.37	31.79	28.48

5. 结论

本文设计了基于 Gabor 基算法展开卷积神经网络的图像恢复模型。其算法展开基于图像恢复变分模型的半二次分裂方法, 所分离出的扩散模型可用于设计高斯去噪器网络。卷积核基于 Gabor 基设计, 能有效提取更多图像特征。针对多个标准数据集的实验表明, 本文方法在图像噪声去除、超分辨率、去模糊等图像恢复方面综合性能指标超过现有相关深度学习方法。此外, 所提出的网络构架和基于 Gabor 基的特征提取方法还可方便地应用于其它图像恢复任务的深度学习网络设计。在基函数组的选取中, 本文并不能充分保证参数选择为最佳方式, 因此未来工作可致力于探索如何处理不同基函数使得在信息不被破坏的前提下达到差异最大化。

致 谢

在这里, 我要以最诚挚的心意感谢潘振宽、魏伟波教授, 感谢他们在我完成硕士论文的过程中提供的一切帮助, 无论是指导研究课题的创新点, 还是语言的润色, 都让我受益匪浅。除此之外, 还要感谢实验室的博士师哥、同级研究生和宿舍的伙伴, 无论在学习上、生活上, 还是论文课题上, 都给予了我无私的帮助和热心的照顾, 这将成为我研究生生涯最美好的风景。感谢国家自然科学基金和山东省联合基金以及本文引用的文献资料等的作者对我论文的大力支持。感恩之情难以用言语度量, 谨以最朴实的话语致以最崇高的敬意。

基金项目

国家自然科学基金“曲面上图像处理的非局部变分模型与算法”(编号: 61772294)和山东省联合基金“基于分布式 FDGA 平台的深度学习模型并行技术研究”(项目编号: ZR2019LZH002)。

参考文献

- [1] 阮秋琦, 仵冀颖. 数字图像处理中的偏微分方程方法[J]. 信号处理, 2012, 28(3): 301-314.

- [2] Chen, Y. and Pock, T. (2017) Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1256-1272. <https://doi.org/10.1109/TPAMI.2016.2596743>
- [3] Chalasani, R. and Principe, J.C. (2014) Dynamic Sparse Coding with Smoothing Proximal Gradient Method. 2014 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, 4-9 May 2014, 7188-7192. <https://doi.org/10.1109/ICASSP.2014.6854995>
- [4] Chan, S.H., Wang, X. and Elgendy, O.A. (2017) Plug-and-Play ADMM for Image Restoration: Fixed-Point Convergence and Applications. *IEEE Transactions on Computational Imaging*, **3**, 84-98. <https://doi.org/10.1109/TCI.2016.2629286>
- [5] Ono, S. (2017) Primal-Dual Plug-and-Play Image Restoration. *IEEE Signal Processing Letters*, **24**, 1108-1112. <https://doi.org/10.1109/LSP.2017.2710233>
- [6] Monga, V., Li, Y. and Eldar, Y.C. (2021) Algorithm Unrolling: Interpretable, Efficient Deep Learning for Signal and Image Processing. *IEEE Signal Processing Magazine*, **38**, 18-44. <https://doi.org/10.1109/MSP.2020.3016905>
- [7] Gavaskar, R.G. and Chaudhury, K.N. (2020) Plug-and-Play ISTA Converges with Kernel Denoisers. *IEEE Signal Processing Letters*, **27**, 610-614. <https://doi.org/10.1109/LSP.2020.2986643>
- [8] Zhang, K., Zuo, W., Gu, S. and Zhang, L. (2017) Learning Deep CNN Denoiser Prior for Image Restoration. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 2808-2817. <https://doi.org/10.1109/CVPR.2017.300>
- [9] Zhang, K., Zuo, W., Chen, Y., Meng, D. and Zhang, L. (2017) Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. *IEEE Transactions on Image Processing*, **26**, 3142-3155. <https://doi.org/10.1109/TIP.2017.2662206>
- [10] Zhang, K., Zuo, W. and Zhang, L. (2018) FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising. *IEEE Transactions on Image Processing*, **27**, 4608-4622. <https://doi.org/10.1109/TIP.2018.2839891>
- [11] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science*, Vol. 9351, Springer, Cham, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [12] Dong, W., Wang, P., Yin, W., et al. (2019) Denoising Prior Driven Deep Neural Network for Image Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**, 2305-2318. <https://doi.org/10.1109/TPAMI.2018.2873610>
- [13] Roth, S. and Black, M.J. (2009) Fields of Experts. *International Journal of Computer Vision*, **82**, 205-229. <https://doi.org/10.1007/s11263-008-0197-6>
- [14] Mutgekar, M.B. and Bhaskar, P.C. (2019) Analysis of DCT and Fast DCT Using Soft Core Processor. 2019 *3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, 23-25 April 2019, 1128-1132. <https://doi.org/10.1109/ICOEI.2019.8862715>
- [15] Kwolek, B. (2005) Face Detection Using Convolutional Neural Networks and Gabor Filters. In: Duch, W., Kacprzyk, J., Oja, E. and Zadrozny, S., Eds., *Artificial Neural Networks: Biological Inspirations—ICANN 2005. ICANN 2005. Lecture Notes in Computer Science*, Vol. 3696, Springer, Berlin, 551-556. https://doi.org/10.1007/11550822_86
- [16] Alekseev, A. and Bobe, A. (2019) GaborNet: Gabor Filters with Learnable Parameters in Deep Convolutional Neural Network. 2019 *International Conference on Engineering and Telecommunication (EnT)*, Dolgoprudny, 20-21 November 2019, 1-4. <https://doi.org/10.1109/EnT47717.2019.9030571>
- [17] Luan, S., Chen, C., Zhang, B., Han, J. and Liu, J. (2018) Gabor Convolutional Networks. *IEEE Transactions on Image Processing*, **27**, 4357-4366. <https://doi.org/10.1109/TIP.2018.2835143>
- [18] Vargas-Rubio, J.G. and Santhanam, B. (2005) On the Multiangle Centered Discrete Fractional Fourier Transform. *IEEE Signal Processing Letters*, **12**, 273-276. <https://doi.org/10.1109/LSP.2005.843762>
- [19] Dabov, K., Foi, A., Katkovnik, V. and Egiazarian, K. (2007) Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing*, **16**, 2080-2095. <https://doi.org/10.1109/TIP.2007.901238>
- [20] Zoran, D. and Weiss, Y. (2011) From Learning Models of Natural Image Patches to Whole Image Restoration. 2011 *International Conference on Computer Vision*, Barcelona, 6-13 November 2011, 479-486. <https://doi.org/10.1109/ICCV.2011.6126278>
- [21] Schmidt, U. and Roth, S. (2014) Shrinkage Fields for Effective Image Restoration. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 23-28 June 2014, 2774-2781. <https://doi.org/10.1109/CVPR.2014.349>
- [22] Schmidt, U., Jancsary, J., Nowozin, S., Roth, S. and Rother, C. (2016) Cascades of Regression Tree Fields for Image

- Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **38**, 677-689. <https://doi.org/10.1109/TPAMI.2015.2441053>
- [23] Keys, R. (1981) Cubic Convolution Interpolation for Digital Image Processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **29**, 1153-1160. <https://doi.org/10.1109/TASSP.1981.1163711>
- [24] Dong, C., Loy, C.C., He, K. and Tang, X. (2014) Learning a Deep Convolutional Network for Image Super-Resolution. In: Fleet, D., Pajdla, T., Schiele, B. and Tuytelaars, T., Eds., *Computer Vision—ECCV 2014. ECCV 2014. Lecture Notes in Computer Science*, Vol. 8692, Springer, Cham. https://doi.org/10.1007/978-3-319-10593-2_13
- [25] Danielyan, A., Katkovnik, V. and Egiazarian, K. (2012) BM3D Frames and Variational Image Deblurring. *IEEE Transactions on Image Processing*, **21**, 1715-1728. <https://doi.org/10.1109/TIP.2011.2176954>