



带噪声水平评估的快速灵活盲深度降噪模型

于海雯, 易昕炜, 徐少平, 林珍玉

(南昌大学 信息工程学院, 南昌 330031)

摘 要: 为提高快速灵活降噪网络(FFDNet)模型的降噪性能, 建立一种噪声水平估计(NLE)模型, 将其预测的噪声水平估计值输入 FFDNet 模型中, 并以 NLE 模型作为 FFDNet 深度降噪模型的前置模块, 使 FFDNet 转换为盲降噪模型。采用浅层卷积神经网络模型将噪声信号从噪声图像中分离出以得到噪声映射图, 将噪声映射图的标准差作为噪声水平的初估值。鉴于噪声水平初估值与真值之间具有强相关性的特性, 应用 BP 神经网络模型对噪声水平初估值进行修正。实验结果表明, 该 NLE 模型与 FFDNet 降噪模型相结合后, 降噪效果总体上与使用真实噪声水平值的 FFDNet 降噪模型接近, 在多数噪声水平值下, 两者的 PSNR 值相差小于 0.1 dB, NLE 模型的估计值可以达到与真实噪声水平值近似的效果, 能够充分发挥 FFDNet 降噪模型快速和灵活特性。

关键词: 快速灵活降噪; 盲降噪; 卷积神经网络; 初估值; 修正值

开放科学(资源服务)标志码(OSID):



中文引用格式: 于海雯, 易昕炜, 徐少平, 等. 带噪声水平评估的快速灵活盲深度降噪模型[J]. 计算机工程, 2020, 46(12): 222-230, 237.

英文引用格式: YU Haiwen, YI Xinwei, XU Shaoping, et al. Fast and flexible blind deep denoising model with noise level estimation[J]. Computer Engineering, 2020, 46(12): 222-230, 237.

Fast and Flexible Blind Deep Denoising Model with Noise Level Estimation

YU Haiwen, YI Xinwei, XU Shaoping, LIN Zhenyu

(School of Information Engineering, Nanchang University, Nanchang 330031, China)

[Abstract] To improve the denoising performance of the Fast and Flexible Denoising Convolutional Neural Network (FFDNet), this paper proposes a Noise Level Estimation (NLE) model that estimates the level of noise. The estimation result is input into the FFDNet model, and the NLE model is taken as the preceding module of the FFDNet deep denoising model to transform it into a blind denoising model. Then the shallow convolutional neural network model is used to separate noise signals from noisy images to obtain the noise map, the standard deviation of which is taken as the initial estimated value of the noise level. Considering the fact that there exists strong correlation between the initial estimated value and ground-truths of the noise level, a Back-Propagation (BP) neural network model is used to correct the initial estimated value of noise level. Experimental results show that when the proposed NLE model works with the FFDNet model, its denoising performance is close to that of the FFDNet denoising model which uses the ground-truths of noise level. For most of the noise level values, the difference of Peak Signal to Noise Ratio (PSNR) values between the two models is within 0.1 dB, which means the estimation results of the proposed NLE model are similar to the ground-truths of noise level, bringing the fast and flexible characteristics of the FFDNet model into full play.

[Key words] fast and flexible denoising; blind denoising; Convolutional Neural Network (CNN); initial estimated value; corrected value

DOI: 10.19678/j.issn.1000-3428.0058327

基金项目: 国家自然科学基金(61662044, 61163023); 江西省自然科学基金(20171BAB202017)。

作者简介: 于海雯(1972—), 女, 讲师、硕士, 主研方向为图形图像处理技术、机器视觉; 易昕炜, 本科生; 徐少平(通信作者), 教授、博士生导师; 林珍玉, 硕士研究生。

收稿日期: 2020-05-14

修回日期: 2020-06-15

E-mail: xushaoping@ncu.edu.cn

0 概述

近年来,受图形处理单元硬件支持的深度卷积神经网络(Deep Convolutional Neural Network,DCNN)被证明具有强大的特征学习和非线性映射能力,将其应用于图像降噪领域后,在执行效率和降噪效果 2 个关键评价指标上相对于传统降噪算法具有明显优势^[1-3]。

目前,IRCNN(Image Restoration CNN)^[4]、DnCNN(Denoising CNN)^[5]、UDNet(Universal Denoising Network)^[6]和 FFDNet(Fast and Flexible Denoising Network)^[7]等基于 DCNN 的主流降噪模型,克服了传统降噪算法依赖人工提取图像先验知识的缺点,在改善降噪效果的同时大幅提高了算法的执行效率。上述模型在本质上都是基于数据驱动(data-driven)策略实现的,通过在大量数据集上进行训练,端到端学习噪声图像与无失真图像或残差图像(噪声图像与无失真图像之间的差图)之间复杂的映射关系,从而实现噪声图像的高质量复原。

以 DnCNN 降噪模型为例,该模型以噪声图像及其对应的残差图像作为深度卷积神经网络的输入-输出训练对(Training Pairs,TP),通过训练实现一个降噪效果好且执行效率高的高斯降噪模型。然而,受技术路线的内在制约,基于 DCNN 网络构建的 DnCNN 等降噪模型的降噪性能在很大程度上依赖于训练图像集,只有在待降噪图像的噪声水平值与训练降噪模型所用图像集的噪声水平值相一致或近似的条件下,才能获得最佳的降噪效果。DnCNN 模型需要针对给定噪声图像的噪声水平值,调用专门针对这一噪声水平值训练的降噪模型,才能获得最佳的降噪效果,这意味着在噪声水平值变化范围内需要训练多个降噪模型。为解决该问题,FFDNet 降噪模型^[7]调整了 DnCNN 网络模型的输入结构,除输入噪声图像外,还增加了噪声水平映射图(Noise Level Map,NLM)作为辅助输入通道。引入的 NLM 映射图与噪声图像大小相同,但噪声水平图中各个点的亮度值为噪声水平值。经过训练后所获得的 FFDNet 降噪模型中的网络参数不再依赖图像噪声水平值,这意味着只需训练一个降噪模型就可处理任意噪声图像的降噪问题。此外,FFDNet 降噪模型在网络模型的输入端和输出端分别设置下采样和上采样模块,这种结构有助于提高模型的降噪效果和执行效率。与 DnCNN 等深度降噪模型相比,FFDNet 模型具有更好的降噪能力、更快的执行速度以及更高的灵活性,但是发挥这些优越性能的前提是 FFDNet 降噪模型需要以待降噪图像噪声水平值生成的 NLM 映射图作为网络模型的输入数据,模型性能取决于能否准确地对待降噪图像中的噪声水平值进行测定。

对给定噪声图像的噪声水平值进行准确测定的方

法称为噪声水平估计(Noise Level Estimation,NLE)算法^[8-10]。对于 NLE 算法而言,估计准确性是衡量其性能的重要评价指标。为获得较准确的估计结果,ZORAN 等人^[11]利用自然图像经 DCT(Discrete Cosine Transform)变换后所得到的边缘带通系数的峰度(Kurtosis)统计值具有尺度不变的特性,通过构建目标优化函数并迭代搜索目标函数最优值来估计噪声图像的噪声水平值。ZORAN 所提算法在中、低噪声水平下具有较好的预测准确性,但是在高水平噪声条件下预测性能下降。此外,该算法执行时间较长,在目标寻优过程中可能失败。为此,DONG 等人^[12]综合利用带通域上 Kurtosis 值的分段平稳性和正则性,将噪声水平估计问题转化为迭代优化拟合 Kurtosis 模型的问题。DONG 所提算法能够准确地估计图像噪声水平值,但其执行效率仍然较低。

GUPTA 等人^[13]利用 DCT 滤波响应的局部对比度归一化系数分布统计值的尺度不变性,实现了一种 NLE 算法,其具有较好的预测准确性,但计算复杂度较高。除了利用图像变换后的统计规律构建 NLE 算法之外,利用图像中同质图块的协方差矩阵对噪声水平值进行估计也是一种常见的 NLE 方法。FANG 等人^[14]利用组合系数对协方差矩阵最小特征值和多个冗余维度的特征值进行线性组合,克服了采用最小特征值估计噪声水平值时的不足,在各个噪声水平下都具有较高的估计准确性和鲁棒性,但是,其仍然需要采用迭代的方式来选择同质图块,执行时间较长。综上,现有的多数 NLE 算法为保证预测准确性而构造复杂的推演过程,导致算法执行效率普遍较低,而作为各种非盲降噪算法的前置模块,NLE 算法的执行效率很重要,其效率较低将影响整个降噪过程的执行效率。

为了使得 FFDNet 降噪模型在实际应用中能获得最佳的降噪效果并保持较高的执行效率,在分析其网络模型结构的基础上,本文建立一种由浅层卷积神经网络(Shallow Convolutional Neural Network,SCNN)和 BP(Back-Propagation)神经网络 2 种串行网络模块构成的 NLE 模型,并将其作为 FFDNet 降噪模型的前置模块。采用与 FFDNet 模型相同的网络架构,但以 SCNN 网络方式构建噪声分离模型将噪声信号从噪声图像中分离出来,得到粗略的噪声映射图(Noise Mapping,NM)。NM 的标准差与真实噪声水平值之间有一定偏差,但鉴于它们之间具有极高的相关性,本文利用预训练的 BP 神经网络修正模型,将初估值修正为最终的噪声水平值,从而获得一种准确而快速的 NLE 模型。将该 NLE 模型与 FFDNet 降噪模型相结合,使得 FFDNet 转换为盲降噪模型 BFFDNet。

1 FFDNet 降噪模型

1.1 模型架构

与传统降噪算法以及近年来兴起的 DCNN 降噪模型相比,FFDNet 降噪模型能够取得更好的降噪效果,且在执行效率方面具有优势。如图 1 所示,FFDNet 降噪网络的隐含层共有 d 层,每一层都由卷积(Conv)、非线性激活(Rectified Linear Units, ReLU)^[15]和批归一化(Batch Normalization, BN)^[16]3种操作的特定组合构成,第1层为 Conv + ReLU,第2层~第 $(d-1)$ 层为 Conv + BN + ReLU,最后一层为 Conv。其中,每个卷积层的卷积核大小均为 3×3 ,ReLU 激活函数能够加快整个网络的收敛速度,BN 操作用于减少内部协变量转移现象^[16]对网络参数选取的影响。为了进一步加快网络模型的执行速度,FFDNet 模型在网络输入端采用下采样算子,将大小为 $W \times H \times C$ ($C=1$ 表示灰度图像, $C=3$ 表示彩色图像)的噪声图像 y 生成对应的4张大小为

$\frac{W}{2} \times \frac{H}{2} \times 4C$ 的低分辨率子图像。在网络输出端,则

采用上采样算子,将降噪后的4张大小为 $\frac{W}{2} \times \frac{H}{2} \times 4C$ 的低分辨率子图像重构为 $W \times H \times C$ 大小的估计图像 $\hat{x}$ 。此外,FFDNet 降噪模型以噪声图像的噪声水平值 σ 构成1张大小为 $\frac{W}{2} \times \frac{H}{2}$ 的 NLM 映射图 N 。映射图 N 中各个像素点位置上的噪声水平值均为 σ ,将其与下采样子图像连接成一个大小为 $\frac{W}{2} \times \frac{H}{2} \times (4C+1)$ 的张量作为模型的输入。NLM 映射图的引入使得降噪模型中的网络参数可以摆脱对噪声水平值的依赖,这意味着只需在噪声变化范围内($\sigma \in [0, 75]$)训练一个 FFDNet 降噪模型,就能够灵活处理具有不同噪声水平值的噪声图像,而且能在模型的降噪效果和细节保护之间取得较好的平衡。

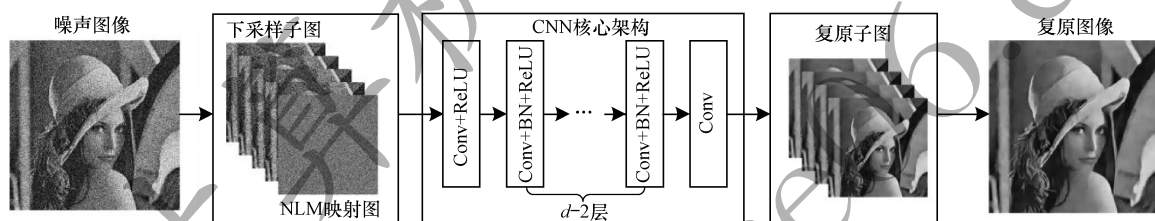


图1 FFDNet 降噪模型的网络架构

Fig.1 Network architecture of FFDNet denoising model

FFDNet 降噪模型以噪声图像 y 的下采样子图像和映射图 N 共同作为网络的输入,对应的无失真子图像作为输出进行训练,降噪模型可形式化描述为 $\hat{x} = F(y, N; \Theta)$,网络参数对噪声水平值的变化不再敏感。通过训练学习网络参数 Θ ,在给定图像准确噪声水平值的情况下,执行上述处理过程,即可根据噪声图像 y 快速复原清晰图像 x 的最佳估计图像 $\hat{x}$ 。

1.2 模型缺陷

FFDNet 降噪模型使用 NLM 映射图作为辅助信息,有助于提高模型的降噪性能,但仍然需要用户准确地给定 σ 值的大小才能获得最佳的降噪效果。当设置的 σ 值与噪声图像中真实噪声水平值不一致时,模型的降噪效果将下降。如图 2 所示,对原始无失真的 Lena 图像施加噪声水平值 $\sigma_{\text{true}} = 40$ 的高斯

噪声,使用 $\sigma_{\text{test}} = 10, 20, 30, 40, 50, 60, 70$ 不同的噪声水平值作为 FFDNet 降噪模型的输入参数分别对该噪声图像进行降噪。从图 2 可以看出,当设置的噪声水平值 σ_{test} 小于待降噪图像的真实噪声水平值 σ_{true} 时,复原图像中仍然残留了部分噪声,如图 2(c)~图 2(e) 所示,且模型的降噪效果随 σ_{test} 与 σ_{true} 之间差值的增大而降低;当 σ_{test} 大于 σ_{true} 时,复原图像的边缘区域过于平滑,纹理细节丢失,如图 2(g)~图 2(i) 所示,且 σ_{test} 与 σ_{true} 之间的差值越大,模型的边缘保持能力越差;当且仅当 σ_{test} 与 σ_{true} 相等时,模型能够取得最佳的降噪效果,复原图像的 PSNR (Peak Signal to Noise Ratio) 值为 30.25 dB。因此,在实际应用中,采用一个预测准确性好且执行效率高的 NLE 算法作为 FFDNet 降噪模型的前置模块,对于保障 FFDNet 模型的降噪效果十分重要。

图 2 不同 σ_{test} 值情况下 FFDNet 模型的降噪效果对比Fig. 2 Comparison of denoising effect of FFDNet model under different σ_{test} values

2 基于神经网络的两阶段 NLE 模型

2.1 问题分析

对于一张原始无失真图像 x , 受高斯噪声 n 污染后得到的噪声图像可以表示为 $y = x + n$ 。由上文可知, 基于噪声图像 y , 使用 FFDNet 深度卷积神经网络可获得估计无失真图像 $\hat{x} = F(y, N; \Theta)$ 。FFDNet 降噪模型同样也可以用于估计残差图像 $\hat{n} = y - \hat{x}$, $\hat{x}$ 和 $\hat{n}$ 两者的求解本质上是同一问题 (dual problem), 因此, 理论上可以使用 FFDNet 降噪模型实现噪声水平值 σ 值估计。具体地, 首先基于待降噪图像给出一个噪声水平初始估计值 σ_0 , 其可设置为噪声图像的标准差, 该值一般大于真实值 σ_{true} , 然后输入到 FFDNet 降噪模型中初步估计出复原图像 $\hat{x}_0$, 接着用噪声图像 y 减去该复原图像 $\hat{x}_0$ 得到对应的残差图像 $\hat{n}_0$, 残差图像 $\hat{n}_0$ 的标准差即为噪声水平值 σ_1 。比较 σ_1 与 σ_0 的大小, 若两者差距较大, 则将 σ_1 再次作为 FFDNet 模型的输入参数, 同样地, 依次得到复原图像 $\hat{x}_1$ 、残差图像 $\hat{n}_1$ 和噪声水平值 σ_2 。重复执行上述过程 k 次, 直到估

计的噪声水平值趋于稳定, 即 $|\sigma_k - \sigma_{k-1}| < \varepsilon$, ε 为预测精度阈值。最终, 将噪声图像 y 以及 σ_k (其作为噪声水平估计值 $\hat{\sigma}$) 输入到 FFDNet 模型中, 以保证噪声图像的复原效果达到最佳。上述过程的原理类似于文献[11]中的逼近估计策略, 使用该方法可以获得准确的噪声水平估计值 $\hat{\sigma}$, 但整体估计时间较长, 从而降低了后续降噪模型的执行效率。尽管 FFDNet 降噪模型执行一次的时间很短, 但是多次调用也会大幅延长执行时间。

2.2 本文模型实现策略

为了避免直接使用深度 FFDNet 降噪模型来估计噪声水平值并实现一种准确而快速的图像噪声水平值估计模型, 本文提出利用人工神经网络技术预训练 NLE 的网络预测模型, 采用粗、精两阶段实现策略完成 NLE 任务。图 3 所示为本文所提出的 NLE 模型的整体流程框图。为了获得较高的执行效率, 利用 SCNN 网络将噪声信号从噪声图像中分离出来获得 NM 映射图, 其中, SCNN 网络是 FFDNet 降噪模型的浅层简化版本, 仅含有 1 层 Conv + BN + ReLU 结构。上述过程所获得的 NM 映射图的标准

差 $\bar{\sigma}$ 即为初步估计的噪声水平值。由于 SCNN 网络所获得的 NM 映射图不够准确,对应的标准差 $\bar{\sigma}$ 与噪声水平真实值 σ_{true} 之间存在一定差距,不能直接作为噪声水平估计值 $\hat{\sigma}$,但是鉴于两者之间具有

非常高的相关性,可以使用常规神经网络模型(本文采用 BP 神经网络)学习两者之间的非线性映射关系,以修正噪声水平初步估计值 $\bar{\sigma}$,从而获得较为准确的噪声水平估计值 $\hat{\sigma}$ 。

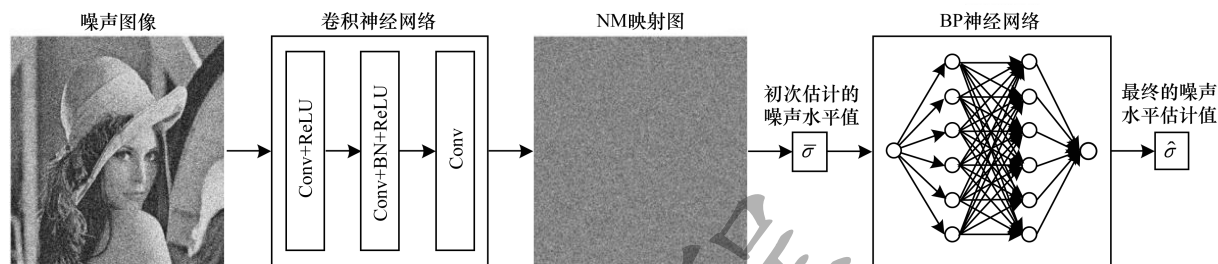


图 3 本文 NLE 模型流程

Fig. 3 Procedure of NLE model in this paper

2.3 噪声水平值初估

本文所设计的 SCNN 网络模型由 3 层卷积层构成,其中,第 1 层为 Conv + ReLU,第 2 层为 Conv + BN + ReLU,第 3 层为 Conv,每一层的卷积核大小均为 $3 \times 3 \times 64$ 。在最后一层卷积层输出 NM 映射图之后,计算该 NM 映射图的标准差 $\bar{\sigma}$,将其作为图像的噪声水平估计值 $\bar{\sigma}$ 。为训练基于浅层卷积神经网络的噪声分离模型,从 BSD 数据库^[17]中随机选取 100 张原始无失真图像,并对这 100 张图像分别施加噪声水平值为 1~100、步长为 1 的高斯噪声生成 10 000 张噪声图像,然后将噪声图像减去对应的无失真图像得到相应的 NM 映射图,以噪声图像集和残差图像集作为所构建网络模型的 TP 训练对进行训练。SCNN 网络仅具有 3 层卷积层,通过 SCNN 来完成噪声水平值估计任务时具有较高的执行效率,但其噪声图像与噪声信号的分离精度低于 DCNN 网络,导致 SCNN 所获得的 NM 映射图像的标准差只能作为噪声水平初步估计

值 $\bar{\sigma}$ 使用。

为了分析 SCNN 浅层网络结构对噪声分离性能的影响,分别使用 1 层 Conv + BN + ReLU(即浅层结构,记为 layer = 1)和 15 层 Conv + BN + ReLU(即深层结构,记为 layer = 15)来构建初步降噪网络。以大小为 512 像素 × 512 像素、噪声水平为 60 的 Lena 噪声图像为例,分别使用上述 2 种网络模型对噪声图像中的噪声进行分离,得到 NM 映射图。将噪声图像减去降噪网络的输出(即残差图像)可获得降噪后的图像,计算降噪后图像的 PSNR 值和相应的执行时间,噪声分离以及降噪效果如图 4 所示,其中,Time 为执行时间。从图 4 可以看出,使用上述 2 种网络结构所得到的降噪、噪声分离效果相差不大,这意味着基于 NM 映射图像计算的标准差 $\bar{\sigma}$ 与噪声图像中噪声水平真实值具有很高的一致性。然而,深层模型的执行时间是浅层模型的 5 倍,因此,本文后续使用 1 层 Conv + BN + ReLU 网络来构建噪声分离网络模型。

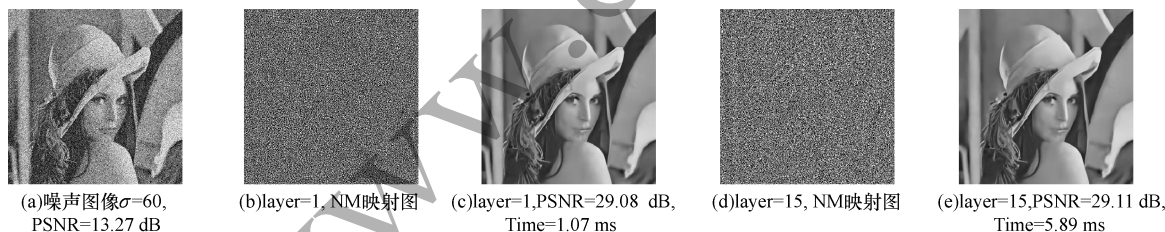


图 4 使用不同网络层数构建的初步降噪网络降噪效果对比

Fig. 4 Comparison of denoising effect of preliminary denoising network constructed with different network layers

2.4 $\bar{\sigma}$ 与 σ_{true} 之间的相关性分析

为验证 $\bar{\sigma}$ 与 σ_{true} 之间是否具有较高的相关性,如图 5 所示,对 Barbara、Boat、Cameraman、Couple、Hill、House、Lena、Man、Monarch 和 Peppers 这 10 张常用测试图像,分别施加噪声水平值为 10、20、40、60、80 和 100 的高斯噪声,使用预先训练好的 SCNN 噪声分离模型获取对应的残差映射图,计算每个残差映射图的标准差并作为噪声水平初步估计值 $\bar{\sigma}$ 。各噪声图像上的噪声水平真实值 σ_{true} 与相对应的噪声水平初步估计值 $\bar{\sigma}$ 之间

的 PLCC(Pearson Linear Correlation Coefficient)、SROCC(Spearman Rank-Order Correlation Coefficient)和 KROCC(Kendall Rank-Order Correlation Coefficient)3 种相关系数^[18]如表 1 所示。从表 1 可以看出,通过 SCNN 网络获得的噪声水平初步估计值 $\bar{\sigma}$ 与真实噪声水平值存在一定的差距,总体上 $\bar{\sigma} < \sigma_{\text{true}}$,且随着噪声水平值的增加,噪声水平初估值与真实噪声水平值之间的差值也同步增加,但是两者之间的 PLCC 系数值达到 0.990 0 以上,SROCC 和 KROCC 系数值均为 1.000 0。

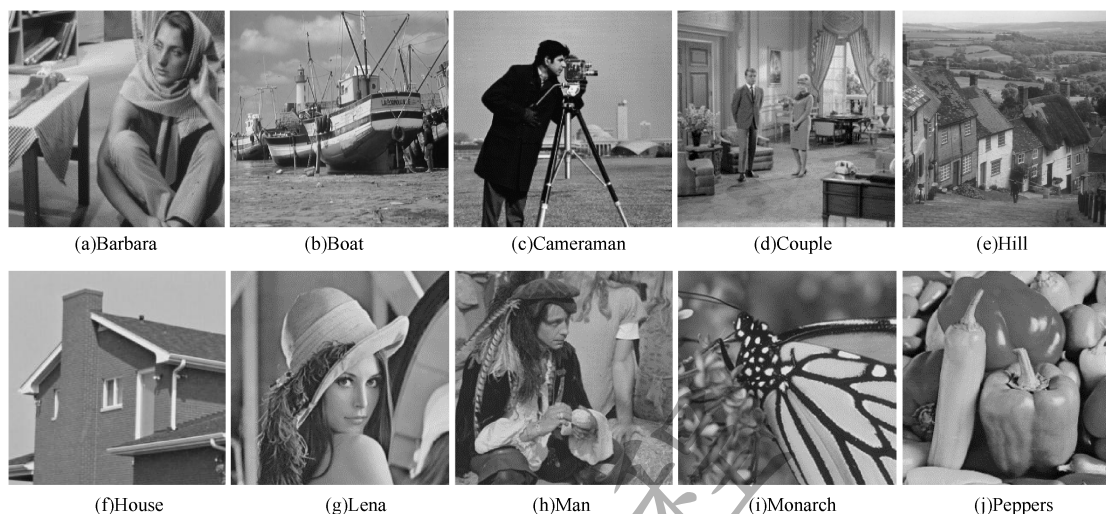


图 5 10 张常用测试图像
Fig. 5 Ten common test images

表 1 单张图像噪声水平初步估计值与噪声水平真值之间的相关性

Table 1 Correlation between the initial estimated noise level and the true noise level of a single image

测试图像	噪声水平初估值						PLCC	SROCC	KROCC
	噪声水平 值为 10	噪声水平 值为 20	噪声水平 值为 40	噪声水平 值为 60	噪声水平 值为 80	噪声水平 值为 100			
Barbara	11.58	19.07	36.50	51.95	64.29	73.86	0.996 0	1.000 0	1.000 0
Boat	7.39	16.49	35.32	51.93	64.85	74.42	0.994 9	1.000 0	1.000 0
Cameraman	9.68	17.56	34.81	50.70	63.32	72.82	0.996 2	1.000 0	1.000 0
Couple	7.32	16.41	35.37	52.21	65.32	74.93	0.995 0	1.000 0	1.000 0
Hill	7.29	16.62	35.25	51.45	64.49	74.10	0.995 1	1.000 0	1.000 0
House	7.28	16.54	35.16	51.22	64.77	74.19	0.995 3	1.000 0	1.000 0
Lena	7.07	16.43	35.25	51.58	64.65	74.27	0.994 9	1.000 0	1.000 0
Man	7.38	16.66	35.22	51.48	64.40	74.20	0.995 3	1.000 0	1.000 0
Monarch	7.56	15.97	34.21	49.99	63.03	72.86	0.996 0	1.000 0	1.000 0
Peppers	8.34	16.89	34.93	51.10	63.65	73.21	0.995 5	1.000 0	1.000 0

在表 1 中,相关性系数均是在包括 6 个数据点的 1 张图像上获得的,为了排除图像内容对相关性指标的影响,选用 BSD 数据库^[17]中不同于训练集的 50 张图像,并对其施加噪声水平值为 1~100、间隔为 1 的高斯噪声构成包含各种图像内容的测试集,该测试集包含 5 000 个数据点,在测试集上再次计算各噪声图像的噪声水平初步估计值 $\bar{\sigma}$ 与噪声水平真值之间的 3 种相关系数,结果如表 2 所示。从表 2 可以看出,SROCC 相关性指标值有所下降,但是仍然达到了 0.91 以上,PLCC 和 KROCC 指标值很高,体现出 $\bar{\sigma}$ 与 σ_{true} 两者间的强相关性。由表 1、表 2 可知,噪声水平初步估计值 $\bar{\sigma}$ 与噪声水平真实值 σ_{true} 之间具有强相关性,这意味着它们之间存在某种形式的非线性映射关系,且不受图像内容的影响。因此,可以考虑采用某种修正方法,如训练某种非线性映射模型来对噪声水平初步估计值 $\bar{\sigma}$ 进行修正处理,从而实现准确的噪声水平估计。

表 2 多张不同噪声图像噪声水平初步估计值与噪声水平真值之间的相关性

Table 2 Correlation between the initial estimated noise level and the true noise level of multiple images with different noises

测试集	PLCC	SROCC	KROCC
50 张 BSD 图像	0.992 3	0.914 4	0.986 0

2.5 $\bar{\sigma}$ 值修正

表 1、表 2 中的数据表明,初步估计的噪声水平值 $\bar{\sigma}$ 与噪声水平真实值 σ_{true} 之间具有强相关性。因此,可以利用非线性映射函数对 $\bar{\sigma}$ 值进行修正。为构建合适的映射函数,对 50 张 BSD 图像分别施加噪声水平值为 5、15、25、35、45、55、65 和 75 的高斯噪声,使用预先训练好的 SCNN 噪声分离模型获取对应的残差映射图,计算每个残差映射图的标准差并作为噪声水平初步估计值,再以噪声水平估计值 $\bar{\sigma}$ 作为输入,对应的噪声水平真实值 σ_{true} 作为输出,分别利用具有 2 层隐含层、每个隐含层的神经元个数

为 6 的 BP 神经网络以及 3 种常规多项式拟合函数 $p(x) = p_1x^n + p_2x^{n-1} + \cdots + p_nx + p_{n+1}$, 在大量训练数据集上学习 $\bar{\sigma}$ 与 σ_{true} 之间的非线性映射函数 f , 并计算在各个噪声水平值下预测值与真值之间的均方根误差 (Root Mean Squared Error, RMSE), 结果如表 3 所示, 最优结果加粗表示。从表 3 可以看出, 虽然 BP 神经网络的映射时间比多项式拟合函数长, 但

在各个噪声水平值下使用 BP 神经网络所获得的均方根误差最小, 表明其修正效果最好。综合考虑执行时间和预测性能, 在图 3 所示的第 2 阶段, 本文使用 BP 神经网络对噪声水平初步估计值进行修正。该模型训练完成后即可获得最终的噪声水平估计值 $\hat{\sigma} = f(\bar{\sigma})$, $\hat{\sigma}$ 能够尽可能地逼近真实的噪声水平值 σ_{true} 。

表 3 50 张 BSD 图像上噪声水平预测值与真值之间的均方根误差
Table 3 RMSE between predicted value and true value of noise level on 50 BSD images

映射函数	RMSE								RMSE 均值	时间/ms
	噪声水平 值为 5	噪声水平 值为 15	噪声水平 值为 25	噪声水平 值为 35	噪声水平 值为 45	噪声水平 值为 55	噪声水平 值为 65	噪声水平 值为 75		
BP 神经网络	2.89	1.49	0.98	1.15	1.58	2.30	2.93	3.36	2.09	8.750
2 次多项式	3.11	1.61	1.30	1.08	1.82	2.81	3.38	3.38	2.31	0.004
3 次多项式	3.01	1.56	1.00	1.19	1.76	2.42	2.97	3.21	2.14	0.005
5 次多项式	2.91	1.63	1.10	1.23	1.52	2.20	2.94	3.55	2.13	0.004

综上, 对于一张待降噪图像, 本文所提的 NLE 模型对其噪声水平值进行估计的整个执行过程, 只需依靠 2 个事先训练好的神经网络模型而没有繁琐的迭代过程, 执行效率高, 这使得其作为 FFDNet 的前置模型具有一定优势。

2.6 优势分析

与其他 NLE 算法相比, 本文所提 NLE 模型优点如下:

1) 采用预训练的 SCNN 网络和 BP 神经网络粗、精相结合, 两阶段串行运行即可实现准确的噪声水平值估计, 没有过多迭代求解过程, 具有较高的执行效率。

2) 在模型训练完成后, 仅需利用训练得到的网络权重值就可直接将待降噪图像映射为残差映射图, 易于使用。

SCNN 网络所获得的 NM 映射图像的精度虽然低于深度 FFDNet 降噪模型, 但是 SCNN 网络能够充分利用在大量训练图像集合上学习的先验知识, 捕获图像受噪声干扰后规律性的变化特点。因此, NM 映射图像的标准差 $\bar{\sigma}$ 与噪声图像中噪声水平真值具有极高的相关性, 基于这一特性, 利用 BP 修正网络模型后可以快速、准确地获得最终的噪声水平估计值 $\hat{\sigma}$, 而无需通过冗长的迭代求解过程。综上, 与现

有的 NLE 算法相比, 基于神经网络实现的两阶段 NLE 预测模型在预测准确性和执行效率方面具有明显优势, 适合作为 FFDNet 降噪模型的前置模块使用, 促使 FFDNet 降噪模型直接转化为盲降噪模型, 使得改进后的降噪模型具有更为广阔的应用前景, 尤其适用于对执行效率要求较高的场景。

3 实验结果与分析

3.1 测试环境

为了评估本文所提 NLE 模型的性能, 将其与文献[11]算法、文献[19]算法、文献[9]算法、文献[20]算法、文献[13]算法和文献[14]算法共 6 种代表性 NLE 算法, 在预测准确性、实际应用效果和执行效率 3 个方面进行比较。测试图像集合是由 BSD 数据库^[17]中的 50 张图像 (不同于 NLE 模型训练时所使用的图像) 构成的纹理图像集, 其中, 10 张具有代表性的图像如图 6 所示, 纹理图像集中的图像内容具有更为复杂的细节特征, 适用于测试算法的鲁棒性。所有实验均在统一的环境下完成, 硬件平台为 CPU Intel® Core™ i7-6700 @ 3.40 GHz RAM 16 GB, 软件环境为 Windows 10.0 操作系统, 编程环境为 Matlab R2017b。



图 6 BSD 数据库中的 10 张代表性图像

Fig. 6 Ten representative images in BSD database

3.2 预测准确性

对 BSD 数据库中 50 张图像分别施加噪声水平值为 5、15、25、35、45、55、65 和 75 的高斯噪声,然后利用各种对比算法对每张噪声图像的噪声水平值进行估计,计算每个噪声级别下各个算法所获得的噪声水平估计值与真实值之间的均方根误差,结果如

表 4 50 张 BSD 图像上各算法噪声水平预测值与真值之间的均方根误差

Table 4 RMSE between predicted value and true value of noise level of each algorithm on 50 BSD images

算法	RMSE								RMSE 均值
	噪声水平值为 5	噪声水平值为 15	噪声水平值为 25	噪声水平值为 35	噪声水平值为 45	噪声水平值为 55	噪声水平值为 65	噪声水平值为 75	
文献[20]算法	4.08	1.83	0.91	1.16	2.39	4.44	7.23	10.76	4.10
文献[9]算法	0.20	0.36	0.92	1.86	3.39	5.55	8.48	12.11	4.11
文献[19]算法	0.77	0.36	0.72	1.70	3.36	5.42	8.19	11.13	3.96
文献[11]算法	2.28	1.81	2.18	2.78	4.74	7.11	9.66	13.29	5.48
文献[14]算法	0.59	1.68	2.55	3.89	3.55	5.52	8.17	11.45	4.68
文献[13]算法	1.30	0.39	1.15	2.39	4.15	6.49	9.51	13.26	4.83
本文模型	2.90	1.47	0.99	1.13	1.57	2.30	2.91	3.40	2.08

3.3 实际应用效果

为了验证本文所提 NLE 模型的实际应用效果,将其与 FFDNet 降噪模型相结合,构建 BFFDNet 模型,使用噪声水平预测值完成盲降噪任务,并与使用真实噪声水平值的 FFDNet 降噪模型在常用图像集和纹理图像集 2 个测试图像集合上进行对比实验。对 2 个测试图像集的图像依次施加噪声水平值为 10、20、30、40、50 和 60 的高斯噪声,计算 2 个模型在每个噪

表 4 所示。从表 4 可以看出,本文 NLE 模型在所有噪声水平下的均方根误差均值最小,且在各个噪声条件下获得的均方根误差值相差不大,说明其预测准确性更为稳定,且在处理复杂图像内容时具有较高的鲁棒性。

声比例下复原所有图像获得的 PSNR 均值,结果如表 5、表 6 所示。从表 5、表 6 可以看出,BFFDNet 降噪模型的降噪效果总体上与 FFDNet 降噪模型相差不大, $\sigma_{\text{true}} = 10$ 时差距较大,但在视觉上效果差别较小,当 $\sigma_{\text{true}} \geq 20$ 时两者几乎没有区别,从而有效地验证了本文 NLE 模型的预测准确性,其估计值可以达到与真实噪声水平值相近的效果。

表 5 FFDNet 与 BFFDNet 降噪模型在常用图像集上的 PSNR 均值比较

Table 5 Comparison of PSNR mean values of FFDNet and BFFDNet denoising model on common image sets dB

模型	PSNR 均值					
	噪声水平值为 10	噪声水平值为 20	噪声水平值为 30	噪声水平值为 40	噪声水平值为 50	噪声水平值为 60
FFDNet	34.96	31.83	30.03	28.80	27.82	27.03
BFFDNet	33.03	31.59	30.03	28.78	27.80	26.98

表 6 FFDNet 与 BFFDNet 降噪模型在纹理图像集上的 PSNR 均值比较

Table 6 Comparison of PSNR mean values of FFDNet and BFFDNet denoising model on texture image sets dB

模型	PSNR 均值					
	噪声水平值为 10	噪声水平值为 20	噪声水平值为 30	噪声水平值为 40	噪声水平值为 50	噪声水平值为 60
FFDNet	33.61	30.04	28.17	26.94	26.02	25.32
BFFDNet	32.42	29.91	28.13	26.88	25.96	25.23

3.4 执行效率

NLE 算法作为多种非盲降噪算法的前置预处理模块,其执行效率关系到图像降噪算法的整体性能。为了更全面地评估本文所提 NLE 模型的执行效率,对大小为 512 像素 $\times$ 512 像素的 Lena 图像施加噪声水平值为 30 的高斯噪声,计算各个对比算法估计该图像噪声水平值时的执行时间,结果如表 7 所示。从表 7 可以看出,在所有对比 NLE 算

法中,本文 NLE 模型具有最少的执行时间,其在预测准确性和执行效率 2 个方面的综合性能最佳,且本文所构建的 BFFDNet 模型在单张 Lena 图像上的总执行时间为 3.2 s,其中,噪声水平估计模块的执行时间为 0.01 s,几乎可以忽略不计,表明 FFDNet 降噪模型与所提 NLE 模块相结合后只需花费很少的时间代价便可实现盲降噪,且能获得较好的降噪效果,具有很强的实用性。

表 7 噪声水平值估计的执行时间比较
Table 7 Comparison of execution time of noise level estimation

level estimation	ms
算法	时间
文献[20]算法	90.0
文献[9]算法	988.7
文献[19]算法	59.6
文献[11]算法	1 284.9
文献[14]算法	2 080.4
文献[13]算法	112 769.3
本文模型	10.3

4 结束语

引入 NLM 映射图后的 FFDNet 仅需训练一个降噪模型即可对具有任意噪声水平值的噪声图像进行降噪,但是其在实际应用中,必须通过准确的噪声水平估计值构建 NLM 映射图后才能获得最佳降噪效果。为此,本文建立一种基于神经网络的两阶段 NLE 模型,并将其作为 FFDNet 降噪模型的前置模块,从而构建 BFFDNet 盲降噪模型。实验结果表明,该 NLE 模型具有较高的预测准确性和执行效率,BFFDNet 盲降噪模型的降噪效果与原 FFDNet 降噪模型的最佳降噪效果相近,验证了所提 NLE 模型的有效性。本文 NLE 模型除了可以与 FFDNet 降噪模型相结合外,也可与多数非盲降噪算法兼容使用。下一步考虑将所提噪声水平估计模型中的浅层降噪网络和 BP 神经网络集成为一个网络,进一步提高模型在训练和应用时的便捷性。

参考文献

- [1] LI Mengmeng, XIA Yang, LI Xinru, et al. Dynamic hybrid recommendation algorithm based on stacked denoising autoencoder[J]. Computer Engineering, 2019, 45(8): 184-189. (in Chinese)
李梦梦, 夏阳, 李心茹, 等. 基于栈式降噪自动编码器的动态混合推荐算法[J]. 计算机工程, 2019, 45(8): 184-189.
- [2] ZHANG Guoling, WANG Xiaodan, LI Rui, et al. Extreme learning machine based on stacked denoising sparse auto-encoder[J]. Computer Engineering, 2020, 46(9): 61-67. (in Chinese)
张国令, 王晓丹, 李睿, 等. 基于栈式降噪稀疏自编码器的极限学习机[J]. 计算机工程, 2020, 46(9): 61-67.
- [3] LÜ Xiaoqi, WU Liang, GU Yu, et al. Low dose CT lung denoising model based on deep convolution neural network[J]. Journal of Electronics and Information Technology, 2018, 40(6): 87-93. (in Chinese)
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): 87-93.
- [4] ZHANG Kai, ZUO Wangmeng, GU Shuhang, et al. Learning deep CNN denoiser prior for image restoration [C]// Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 2808-2817.
- [5] 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.
- [6] LEFKIMMIATIS S. Universal denoising networks: a novel CNN architecture for image denoising[C]// Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2018: 145-168.
- [7] ZHANG Kai, ZUO Wangmeng, ZHANG Lei. FFDNet: toward a fast and flexible solution for CNN based image denoising[J]. IEEE Transactions on Image Processing, 2018, 27(9): 4608-4622.
- [8] PYATYKH S, HESSER J, ZHENG L. Image noise level estimation by principal component analysis[J]. IEEE Transactions on Image Processing, 2013, 22(2): 687-699.
- [9] LIU X H, TANAKA M, OKUTOMI M. Single-image noise level estimation for blind denoising[J]. IEEE Transactions on Image Processing, 2013, 22(12): 5226-5237.
- [10] XU Shaoping, ZENG Xiaoxia, TANG Yiling. Fast noise level estimation algorithm based on two-stage support vector regression[J]. Journal of Computer-Aided Design and Computer Graphics, 2018, 30(3): 447-458. (in Chinese)
徐少平, 曾小霞, 唐祎玲. 基于两阶段支持向量回归的快速噪声水平估计算法[J]. 计算机辅助设计与图形学学报, 2018, 30(3): 447-458.
- [11] DANIEL Z, YAIR W. Scale invariance and noise in natural images [C]// Proceedings of 2009 IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2009: 2209-2216.
- [12] DONG Li, ZHOU Jiantao, TANG Yuanyan. Noise level estimation for natural images based on scale-invariant kurtosis and piecewise stationarity[J]. IEEE Transactions on Image Processing, 2017, 26(2): 1017-1030.
- [13] GUPTA P, BAMPIS C G, JIN Y, et al. Natural scene statistics for noise estimation[C]// Proceedings of 2018 IEEE Southwest Symposium on Image Analysis and Interpretation. Washington D. C., USA: IEEE Press, 2018: 85-88.
- [14] FANG Zhuang, YI Xuming. A novel natural image noise level estimation based on flat patches and local statistics [EB/OL]. [2020-03-10]. <https://doi.org/10.1007/s11042-018-7137-4>.
- [15] RAWAT W, WANG Z H. Deep convolutional neural networks for image classification: a comprehensive review[J]. Neural Computation, 2017, 29(9): 2352-2449.
- [16] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. [2020-03-10]. <https://arxiv.org/abs/1502.03167>.

(下转第 237 页)

(上接第 230 页)

- [17] PABLO A, MICHAEL M, CHARLESS F, et al. Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.
- [18] ZHANG Lin, ZHANG Lei, MOU Xuanqin, et al. A comprehensive evaluation of full reference image quality assessment algorithms [C]//Proceedings of the 19th IEEE International Conference on Image Processing. Washington D. C., USA: IEEE Press, 2012: 1477-1480.
- [19] YANG S M, TAI S C. Fast and reliable image-noise estimation using a hybrid approach [J]. Journal of Electronic Imaging, 2010, 19(3): 033007.
- [20] CHEN G, ZHU F, HENG P A. An efficient statistical method for image noise level estimation [C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2015: 477-485.

编辑 吴云芳