

基于 CNN 噪声分离模型的噪声水平估计算法

徐少平 刘婷云 李崇禧 唐伟玲 胡凌燕
(南昌大学信息工程学院 南昌 330031)
(xushaoping@ncu.edu.cn)

Noise Level Estimation Algorithm Using Convolutional Neural Network-Based Noise Separation Model

Xu Shaoping, Liu Tingyun, Li Chongxi, Tang Yiling, and Hu Lingyan
(School of Information Engineering, Nanchang University, Nanchang 330031)

Abstract The existing noise level estimation (NLE) algorithms usually adopt the strategy that separates the noise signal from the content of an image to estimate its noise level. Since only a single noisy image can be exploited, these algorithms usually design a variety of complex processes to ensure the accuracy of noise separation, resulting in low execution efficiency. To this end, a novel NLE algorithm using convolutional neural network (CNN)-based noise separation model is proposed in this paper. Specifically, we first add Gaussian noise with different levels to a great amount of representative undistorted images to obtain a training database. Then, we train a CNN-based noise separation model on the training database to obtain the noise mapping from a given noisy image. Considering the fact that the coefficients of the noise mapping show Gaussian distribution behavior, we utilize the generalized Gaussian distribution (GGD) to model the coefficients of the noise mapping, and use two parameters (scale and shape) of the model as the noise level-aware features (NLAF) to describe the level of a noisy image. Finally, an improved back propagation (BP) neural network is used to map the NLAF features to the final noise level. Extensive experiments demonstrate that our method outperforms the most existing classical NLE algorithms in terms of both computational efficiency and estimation accuracy, which makes it more practical to use.

Key words noise level estimation (NLE); noise separation; convolutional neural network (CNN); generalized Gaussian distribution (GGD); noise level-aware feature (NLAF); noise level mapping

摘 要 现有的噪声水平估计(noise level estimation, NLE)算法通常采取先将图像内容信号与噪声信号分离,然后基于分离出的噪声信号估计出图像噪声水平值的实现策略.由于仅有噪声图像本身的信息可以利用,这些算法为保证噪声分离的准确性设计了各种复杂的处理过程,导致其执行效率偏低.为此,提出一种新的基于卷积神经网络噪声分离模型的 NLE 算法.首先,对大量原始无失真图像施加不同噪声水平的高斯噪声获得噪声图像集合,然后利用卷积神经网络构建一个专门从噪声图像中分离噪声信号获得噪声映射图(noise mapping)的预测模型.考虑到噪声映射图的系数值具有类高斯分布特性,利

用广义高斯分布(generalized Gaussian distribution, GGD)模型对噪声映射图建模并以模型参数值作为反映图像噪声水平高低的特征值.最后,利用改进的 BP 神经网络将该特征值映射为最终的噪声水平预测值.大量实验数据表明:所提出的 NLE 算法在预测准确度和执行效率 2 个方面的综合性能优于现有的 NLE 算法,更具实用价值.

关键词 噪声水平估计;噪声分离;卷积神经网络;广义高斯模型;噪声水平感知特征;噪声水平值映射

中图法分类号 TP391

对于图像超分辨率重建^[1-2]、图像去模糊^[3]、图像降噪^[4-5]等各类图像复原算法来说,噪声水平值是一个至关重要的输入参数,不准确的噪声水平值会在很大程度上影响这些算法的性能.然而,大多数已提出的图像复原算法^[6]都属于非盲类型的(non-blind),即需要人工手动设置并调整相应的噪声水平值才能获得最佳的图像复原效果,这严重限制了它们的实际应用范围.因此,研究自动化的图像噪声水平评估(noise level estimation, NLE)算法是很有必要的.

现有的大多数 NLE 算法的核心技术路线是先将图像内容信号和噪声信号分离,然后根据分离出的噪声信号提取特征进而估计噪声水平值.例如 Immerkær 利用 1 个 3×3 的矩阵掩模,对图像局部邻域内的加性 0 均值高斯噪声求和,再根据方差估算出噪声水平值^[7].虽然该算法的执行效率很高,但其噪声水平估计准确率较低;Zoran 等人^[8]认为图像噪声是引起边缘带通滤波器响应分布峰度值(kurtosis value)变化的主要因素,通过构建优化模型对噪声水平值进行估计.该算法在中低水平噪声条件下的估计准确度较高,但是在噪声水平较高的情况下其估计效果不理想,执行时间较长.近期 Katase 等人^[9]提出了一种基于离散余弦变换(discrete cosine transform, DCT)系数特征的 NLE 算法.该算法根据 DCT 系数中高频部分的标准方差选取平滑的图块,利用噪声图像中的高频系数计算出噪声水平值,它在噪声水平较低的情况下性能良好,对于高水平噪声,预测效果不够理想. Rakhshanfar 等人^[10]根据亮度值和方差对图像分解的图块进行分类,基于噪声统计特性和图像同质区域,将与噪声模型相似程度最高的图块群(patch clusters)用于估计噪声方差.该方法可以处理多种类型图像噪声,在预测准确性表现相对良好,但是其执行效率与图像内容的复杂度密切相关.当噪声水平值较高时,执行时间显著增加. Chen 等人^[11]通过研究发现无失真图像中提取的图块通常位于所谓的

低维子空间(low-dimensional subspace),可以利用主成分分析(principal component analysis, PCA)方法获得冗余维度上的特征值,将其用于估计噪声方差.该算法用循环穷举的方式来确定冗余维度上的特征值,导致算法效率不高.

总体来说,上述算法都属于所谓的基于单图像的噪声水平估计(single-image based noise level estimation, SNLE)算法.在估计噪声水平值时仅有待评价图像本身的信息可以利用,因此需要设计复杂的处理流程将噪声信号从图像中分离(提取)出来后才能进一步基于噪声信号部分进行估计,在执行效率方面并不是很理想. NLE 算法作为众多图像处理算法的预处理模块,除了预测准确性外,执行效率也是重要的评价指标.近年来,笔者试图通过基于训练的方法,充分利用蕴藏在已知多幅图像中的规律性的信息,从图像中提取反映噪声水平值高低的特征值后实现快速的噪声水平映射(预测),这种方法被称为基于多图像的噪声水平估计(multiple-image based noise level estimation, MNLE)算法.例如在文献[12]中,提出了用梯度幅值和拉普拉斯变换 2 种算子提取噪声水平感知特征(noise level-aware feature, NLAf),然后基于在大量训练数据上获得的 NLAf 特征值以及相应的噪声水平值利用支持向量回归(support vector regression, SVR)完成预测模型的训练.预测模型所预测的噪声水平值能使 BM3D(block-matching and 3D filtering)降噪算法获得最佳降噪效果.基于训练的 MNLE 算法的执行效率比现有经典算法具有明显优势,但是预测精度水平较主流算法尚有一定差距.提高基于 MNLE 算法的预测准确性需要从 2 个方面入手,即提取高质量的能反映噪声水平高低的 NLAf 特征和利用机器学习算法训练非线性映射能力更强的预测模型完成从 NLAf 特征值到噪声水平值的映射^[12-13].

为了提取对噪声水平值更为敏感的 NLAf 特征,高质量地将图像内容与噪声信号分离是基础性的工作.为此,将近年来出现的非线性逼近能力强大

的卷积神经网络(convolutional neural network, CNN)^[14-16]技术引入到 MNLE 算法中来,提出一种基于 CNN 噪声分离模型的噪声水平估计算法.该算法利用 CNN 网络训练一个专门的预测模型完成将噪声信号从噪声图像中剥离出来的任务,然后利用通用高斯模型(generalized Gaussian distribution, GGD)对噪声映射图(noise mapping)进行建模并利用模型参数值作为反映图像噪声水平高低的 NLAF 特征值.最后利用改进的 BP 映射模型完成从 NLAF 特征值到噪声水平值的预测任务.相对于已有的各类 NLE 算法,所提出的 NLE 算法的优势主要体现在 3 个方面:

1) 噪声信号分离效果好.充分利用了 CNN 强大的非线性逼近能力,在已知噪声水平下的样本图像集合中学习先验知识,高质量地将噪声信号从噪声图像中分离出来,为构建高敏感度的 NLAF 特征打下了很好的基础.

2) NLAF 特征到噪声水平的映射能力强.采用改进的 BP 神经网络技术获得了比常规 SVR 支持向量回归映射更高的预测正确率.

3) 执行效率高.改进算法的特征值提取和噪声水平值的映射工作是由 CNN 和改进 BP(back propagation)神经网络来完成的.一旦训练完成,其执行效率非常高,尤其是 CNN 卷积网络在 GPU(graphics processing unit)硬件支持下运行时间很短,几乎可以忽略.

1 CNN 简介

CNN 卷积神经网络本质上可以看作是一种特殊的映射函数,它能够从大量原始图像(图块)数据中挖掘隐藏的特征信息,从而学习输入与输出值之间复杂的映射关系^[17].典型的多层卷积神经网络由输入层、多个隐藏层和输出层组成.隐藏层中关键的组成结构是卷积(convolutional, Conv)结构 $\tilde{\mathbf{X}} = C_w(\mathbf{X})$.对于输入的图块 $\mathbf{X}(u) = (X_1(u), X_2(u), \dots, X_p(u))$,通过 1 个滤波器 $\mathbf{W} = (w_{l,l'}), l = 1, 2, \dots, q, l' = 1, 2, \dots, p$ 和逐点非线性函数(point-wise nonlinearity function) $\psi(\cdot)$ 进行卷积^[18]:

$$\tilde{X}_l(u) = \psi\left(\sum_{l'=1}^p (X_{l'} * w_{l,l'})(u)\right). \quad (1)$$

在卷积层中,用滤波器与上一层若干个特征映射图(feature map)进行卷积,计算出新的特征映射

图,通过激活函数将特征保留并输出到下一层结构.近年来研究者们主要采用修正线性单元(rectified linear unit, ReLU)^[19]作为激活函数,其数学形式表述为 $\psi(x) = \max(x, 0)$,将负值置 0. ReLU 能够加快整个网络的收敛速度,而不影响卷积层的感受野(receptive field),大大提高了激活函数的性能^[20].卷积后产生 q 维的特征映射图(feature map)

$$\tilde{\mathbf{X}}(u) = (\tilde{X}_1(u), \tilde{X}_2(u), \dots, \tilde{X}_q(u)). \quad (2)$$

此外,池化层 $\tilde{\mathbf{X}} = P(\mathbf{X})$ 也是 CNN 网络中常用的组成构件.对于每一个输入 $\tilde{X}_l(u)$,有:

$$\tilde{X}_l(u) = P(\{X_l(u') : u' \in N(u)\}), \quad (3)$$

其中, $l = 1, 2, \dots, q, N(u)$ 是 u 的邻域, P 是某种类型的池化函数(可以是最大值、均值等类型).基本的 CNN 卷积神经网络可以表示为卷积层和可选的池化层的多次组合,即

$$\Phi_H(\mathbf{X}) = (C_H^{(K)} \cdots P \cdots C_H^{(2)} \circ C_H^{(1)})(\mathbf{X}), \quad (4)$$

其中, P 为可缺省的池化操作, $\mathbf{H} = (\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \dots, \mathbf{H}^{(K)})$ 是网络参数(卷积核的系数值)的超矢量(hyper-vector).通常,在卷积神经网络中的卷积层或者池化层之后可以连接全连接层(fully connected layer),全连接层在整个 CNN 中用来获得特征,为后面的分类或者回归任务做准备.另外,为了减少内部协变量转移对网络参数的影响,批归一化层(batch normalization, BN)^[21]常被用来对卷积神经网络隐藏层的输出数据进行批归一化处理,从而加快神经网络的收敛速度并获得稳定性.

2 基于噪声分离模型的 NLE 算法

2.1 NLE 算法的基本思想

如图 1 所示,改进算法技术路线的关键之处在于:首先高质量地将噪声图像的内容信号与噪声信号进行分离获得噪声映射图;进而基于噪声映射图提取 NLAF 特征;最后通过训练学习获得映射能力强的预测模型将 NLAF 特征准确地映射为噪声水平值.具体地,在训练阶段:1)选择大量具有代表性的自然图像添加不同噪声水平的高斯噪声获得噪声图像集合,噪声图像、无失真图像、噪声图像与无失真图像相减获得的噪声映射图以及对应的噪声水平值构成训练数据集合;2)通过训练 CNN 网络在噪声图像和噪声映射图之间构建 CNN 噪声分离模型,利用该模型即可完成将噪声信号从噪声图像中剥离出来的任务;3)分离噪声信号后,采用 GGD 模型对噪声映射图进行建模并利用模型参数值作为反

映图像噪声水平高低的 NLAF 特征值;4)基于从训练集中的噪声图像上提取出的特征值及相应的真实噪声水平值,训练改进的 BP 映射模型完成从 NLAF 特征值到噪声水平值的预测任务.在预测阶

段:1)将噪声图像输入到 CNN 噪声分离模型分离出噪声映射图;2)用 GGD 特征提取模型提取其 NLAF 特征并由改进的 BP 映射模型最终映射为噪声水平值.

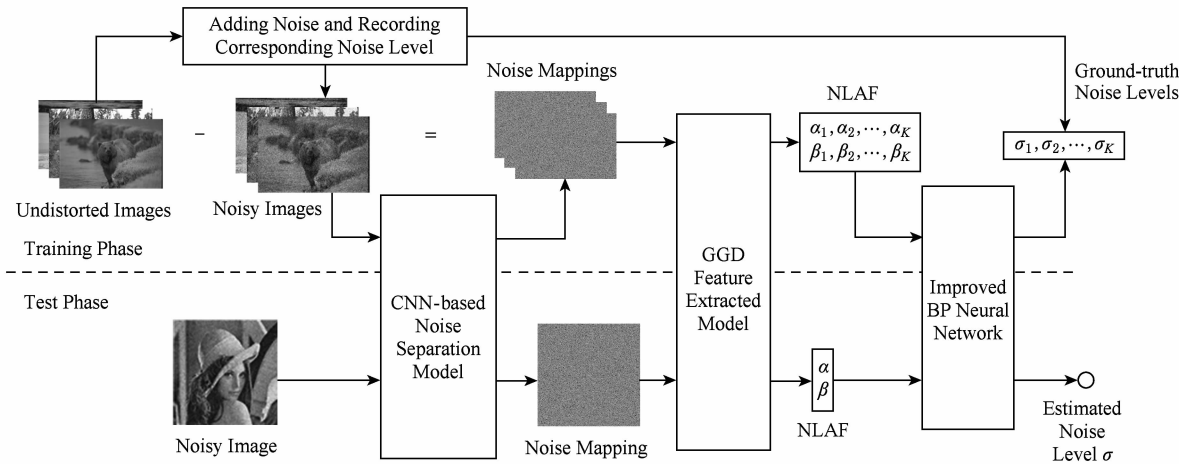


Fig. 1 The overall framework of the proposed method

图 1 改进算法流程框图

2.2 分离噪声信号

本文主要工作是利用 CNN 网络强大的学习能力,达到从噪声图像中分离出噪声信号的目的.一般来讲,受到噪声干扰的图像的数据模型可以定义为

$$y=x+n, \tag{5}$$

其中, y 是噪声图像, x 是原始图像, n 是高斯噪声.笔者利用 CNN 预测模型 $\hat{n}=f_{\text{CNN}}(y)$ 获得噪声映射图,进而基于噪声映射图估计噪声图像的噪声水平值.

基于上述设计需求,与一般利用 CNN 卷积神经网络完成分类或者回归任务不同,所提出的算法需要基于输入的噪声图像各个像素点的亮度值得到对应像素点的噪声值,故输入模型的图像大小和输出图像的大小相同.因此,所设计的 CNN 卷积神经网络中不使用池化层和全连接层结构,仅包括卷积层、BN 和 ReLU 共 3 种类型的网络部件,具体结构如图 2 所示.从图 2 中可以看出,网络从左到右共分 3 段:

- 1) Conv+ReLU. 第 1 段用 64 个大小为 3×3 的卷积核生成 64 个特征图像.
- 2) Conv+BN+ReLU. 第 2 段与第 1 段的区别是使用了 BN 批归一化层,使用 BN 层的目的是为了保证每一层网络的输入具有相同的分布、加快训练进程并提高噪声水平预测准确率.
- 3) Conv. 最后 1 段采用 1 个大小为 $3\times3\times64$ 的卷积核来输出噪声映射图.

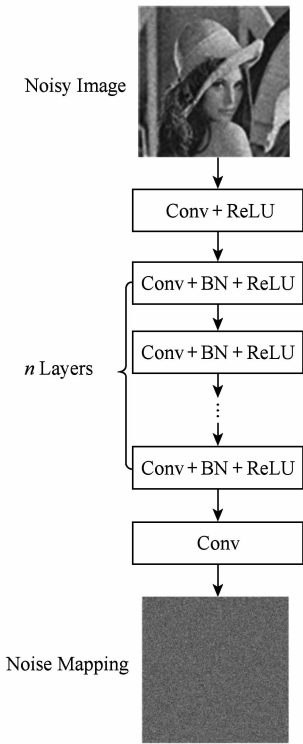


Fig. 2 The network architecture of CNN-based noise separation model

图 2 CNN 噪声分离模型的网络体系架构图

2.3 噪声映射图与 GGD 建模

为了验证卷积神经网络对噪声图像中的噪声信号部分的实际分离效果,以 1 幅 512×512 大小的 Lena 图像为例,对其分别添加噪声级别为 5,10,20,

40,50 的高斯噪声,然后利用所提出的 CNN 分离模型将噪声信号分离出来得到噪声映射图,计算噪声映射图中系数值分布情况,系数分布直方图如图 3 所示(为了便于显示和比较,图 3 中的横坐标值所代表的残差系数值(噪声值)被归一化到 $[-1,1]$ 之间,而纵坐标则是像素点个数除以 1 000 以后的结果).由图 3 可知,噪声映射图的系数值分布具有显著的一类高斯分布特征,可用 GGD 广义高斯分布对噪声映射图进行建模.

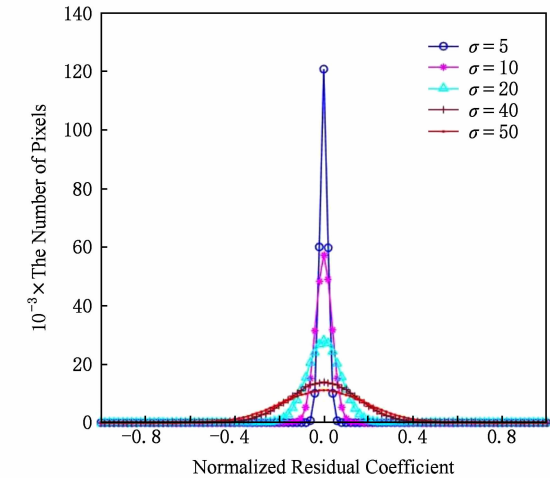


Fig. 3 The distribution histogram of coefficient values
图 3 系数值分布直方图

GGD 模型^[22]是一种应用非常广泛的分布模型,一般定义为

$$p(x;\alpha,\beta)=\frac{\beta}{2\alpha\Gamma(1/\beta)}\exp\left(-\left(\frac{|x|}{\alpha}\right)^\beta\right), \quad (6)$$

其中, $\Gamma(\cdot)$ 是 Gamma 函数,即:

$$\Gamma(t)=\int_0^\infty u^{t-1}\exp(-u)du, t>0. \quad (7)$$

式(6)中的 α 为尺度(scale)参数,控制着 GGD 密度函数的扩散程度. β 为形状(shape)参数,用于修正峰度值的衰减速度. 参数 α 和 β 与 GGD 分布的标准差 σ 满足关系式:

$$\alpha=\sigma\sqrt{\frac{\Gamma(1/\beta)}{\Gamma(3/\beta)}}, \quad (8)$$

其中, $\beta>0$. 当 $\beta=1$ 时, $\alpha=\sqrt{2}\sigma/2$,对应于拉普拉斯分布;当 $\beta=2$ 时, $\alpha=\sqrt{2}\sigma$,对应于标准高斯分布;当 $\beta\rightarrow0$ 时,分布趋于 δ 函数;当 $\beta\rightarrow+\infty$ 时,分布趋于均匀分布.

2.4 噪声水平感知特征选取

用 GGD 模型为噪声映射图系数值分布建模后,模型中共有 2 个参数,即 α 和 β . 为了测试模型中的 2 个参数是否能够作为 NLAF 特征值使用,用 GGD 模型为图 2 中 5 个级别的噪声映射图建模,所

提取的模型参数如表 1 所列. 从表 1 中可以看出: α 参数值与相应噪声水平值之间的 Pearson 线性相关系数(Pearson linear correlation coefficient, PLCC)为 1,表明这 2 者之间具有极强的相关性; β 参数值与相应噪声水平值之间的 PLCC 系数为 0.649 6,表明这 2 者之间的相关性处于中等水平. 为了全面验证该参数是否与噪声值具有强相关性,在 100 幅图像上执行与图 2 相同的处理过程并提取 GGD 模型的 2 个参数值,再计算它们分别与真实噪声水平值之间的 PLCC 相关系数. 经计算, α 值与真实噪声水平值的线性相关系数为 0.998 7,继续保持非常高的相关性. 而 β 值与真实噪声水平值的线性相关系数为 0.668 3,有所提高但仍然处于中等水平. 为了获得鲁棒的噪声水平值,将参数值 α 和 β 组合作为 NLAF 噪声水平感知特征矢量 $F=(\alpha,\beta)$ 来使用.

Table 1 Correlation Coefficients Between the Noise Levels and Parameters of GGD Model

Parameters	Noise Levels					PLCC
	5	10	20	40	50	
α	0.017 1	0.036 5	0.074 8	0.151 7	0.190 1	1.000 0
β	1.754 0	1.961 0	1.990 0	1.996 0	1.989 0	0.649 6

2.5 改进的 BP 神经网络

用 GGD 模型在大量训练图像上提取噪声映射图的特征值后,理论上即可利用任何回归分析技术将特征值(构成矢量形式)映射为最终的噪声水平预测值(标量). 由于所使用的特征值只有 2 维,维数较少,利用常规的 SVR 构建预测模型的预测准确性不是很理想. 如图 4 所示,为了进一步提升预测准确

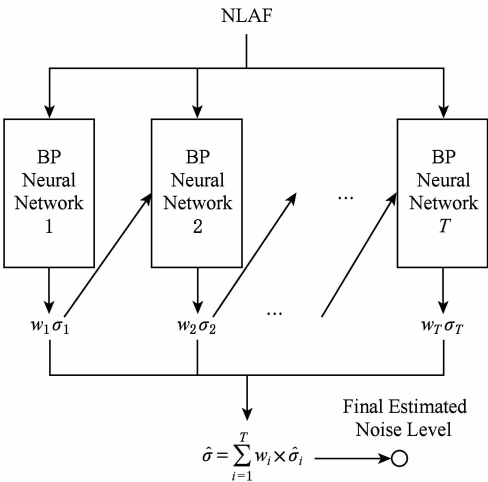


Fig. 4 Pipeline of the improved BP neural network
图 4 改进 BP 神经网络工作原理

性,首先利用若干个 BP 网络构建预测能力相对较弱的弱预测器,然后利用 AdaBoost 算法^[23]在多个弱预测器的基础上构建比较健壮、准确的强预测器。AdaBoost 算法原本的设计思想是利用若干个弱分类器构建强分类器的分类问题,这里将其运用到回归问题中来,所用的预测模型称为改进的 BP 映射模型。

选取 K 幅图像进行训练,对每幅图像添加已知噪声水平值为 σ_i 的高斯噪声,用 GGD 模型提取对应的特征矢量 $\mathbf{F}_i$ (其中 $i=1,2,\cdots,K$) 构成训练集合 $\{(\mathbf{F}_1,\sigma_1),(\mathbf{F}_2,\sigma_2),\cdots,(\mathbf{F}_K,\sigma_K)\}\subset\mathbb{R}^2\times\mathbb{R}$, 训练出预测函数 $\varphi(\cdot)$ 以实现特征矢量到噪声水平值之间的映射。改进 BP 映射算法的具体执行过程如下:首先,确定弱预测器的数量(根据实验数据 $T=5$ 时,已能满足精度要求且同时兼顾执行效率)。从第 1 个弱预测器开始,利用 BP 神经网络建立特征矢量 $\mathbf{F}_i$ 和噪声水平值 σ_i 之间的映射关系,第 j 条训练数据对预测结果的权重更新方法为

$$D_{i,j}=\begin{cases} \frac{1}{K}, & i=1, \\ D_{i-1,j}\times(1+C\times l(\sigma_j-\hat{\sigma}_{j-1,j})), & i=2,3,\cdots,T, \end{cases} \quad (9)$$

其中, $D_{i,j}$ 代表在第 i 个弱预测器中第 j 条训练数据对预测误差的权重, C 为 $[0,1]$ 之间的常数。对于第 1 个弱预测器来说,各个训练数据的权重初始化为 $1/K$,从第 2 个弱预测器开始,则根据前一个预测器预测结果的误差利用二值函数(binary function)

$$l(x)=\begin{cases} 1, & x>\text{threshold} \\ 0, & x\leq\text{threshold} \end{cases} \quad (10)$$

进行修正。根据实验数据,设定参数 $C=0.1$, $\text{threshold}=0.1$ 。第 i 个弱预测器整体的预测误差可以定义为用第 i 个预测器所预测的噪声估计值与真实噪声水平值之间的差值和训练数据权重乘积的累积和,从数学形式上描述为

$$Err_i=\sum_{j=1}^K D_{i,j}\times l(\sigma_j-\hat{\sigma}_{i,j}). \quad (11)$$

利用凸函数将每个弱预测器的预测误差转换为弱预测器在最终预测结果中的权重。这样使得那些预测误差较小的弱预测器的权重增加,而误差较大的弱预测器的权重减小。第 i 个弱预测器的权重定义为

$$w_i=\frac{1}{\exp(-b(|Err_i|-c))}, \quad (12)$$

最后,基于第 i 个弱预测器的权重及其所预测的噪声水平值 $\hat{\sigma}_i$,计算出最终的噪声水平估计值:

$$\hat{\sigma}=\sum_{i=1}^T w_i\times\hat{\sigma}_i. \quad (13)$$

实验数据表明使用过多或者过少的弱预测器都将导致预测准确性下降。另外一方面,随着使用更多的弱预测器,执行效率也将随之下降。文中最终使用 5 个弱预测器构建改进的 BP 映射模型,在这个配置下能在预测准确性和执行效率这 2 个方面获得最佳平衡点(限于篇幅,这里不再给出具体实验数据)。综上所述,基于多个 BP 网络模型利用 AdaBoost 算法提升预测准确性的伪代码描述如下:

算法 1. 改进 BP 神经网络模型训练过程。

输入:训练图像集 $\{(\mathbf{F}_1,\sigma_1),(\mathbf{F}_2,\sigma_2),\cdots,(\mathbf{F}_K,\sigma_K)\}\subset\mathbb{R}^2\times\mathbb{R}$;

输出:各预测器模型 $\{BP_i|1\leq i\leq 5,i\in\mathbb{N}_+\}$ 及对应的权重值 $w=\{w_1,w_2,\cdots,w_5\}$ 。

- ① 初始化基于 BP_i 网络的弱预测器的分布 $D_{1,j}$ 、计算对应的估计误差 Err_1 及权值 w_1 ;
- ② for $i=2,3,\cdots,5$
- ③ 根据式(11)计算第 i 个弱预测器 BP_i 的估计误差 Err_i ;
- ④ 根据式(9)计算第 i 个弱预测器 BP_i 的分布 $D_{i,j}$;
- ⑤ 根据式(12)计算第 i 个弱预测器 BP_i 的权值 w_i ;
- ⑥ 归一化弱预测器的分布 $D_{i,j}$;
- ⑦ end for
- ⑧ 输出各个 BP 预测器并归一化它们的权值。

3 实验与分析

3.1 测试环境

从 BSD(Berkeley segmentation dataset) 数据库^[24]中随机选取了 100 幅图像,对这些图像添加 51 个级别(0,1,⋯,50)的高斯噪声作为训练图像集合,基于训练图像集合训练预测模型。测试图像集包括 2 个:第 1 个由在各类图像处理相关文献中常用的 10 幅图像组成,如图 5 所示;为了增加测试难度,在 BSD 数据库另选了与训练集合不同的 50 幅纹理图像构成第 2 个测试集合。BSD 数据库原本作为纹理分割的基准测试数据库,其图像中的纹理细节非常丰富,很多 NLE 算法的性能在该数据库上表现不佳,比较适合用来作为测试图像集合。参与比较的 NLE 算法包括 Immerkær 算法^[7]、Zoran 算法^[8]、

Liu 算法^[25]、Chen 算法^[11]、Rakhshanfar 算法^[10] 和 Xu 算法^[1], 其中 Xu 算法是笔者之前基于主成分分析和 SVR 实现的 NLE 算法. 为了验证改进的 BP 映射模型的性能, 本文对比了 4 种从 NLAf 特征到噪声水平值的映射模型. 其中, 直接利用常规 SVR 和 BP 构建映射模型的 NLE 算法分别记为 CNN+SVR 和 CNN+BP; 基于 SVR 和 BP 利用 AdaBoost

算法改进的映射模型实现的 NLE 算法分别记为 CNN+AdaSVR 和 CNN+AdaBP, 这样共有 10 个 NLE 算法参与对比实验. 通过比较以上各个算法在不同噪声水平下的噪声估计准确性以及执行时间, 来测试所提出算法的综合性能. 实验在 Matlab R2017b(运行在 Intel® Core™ i7-6700 CPU RAM 8 GB 主机上)环境下完成.

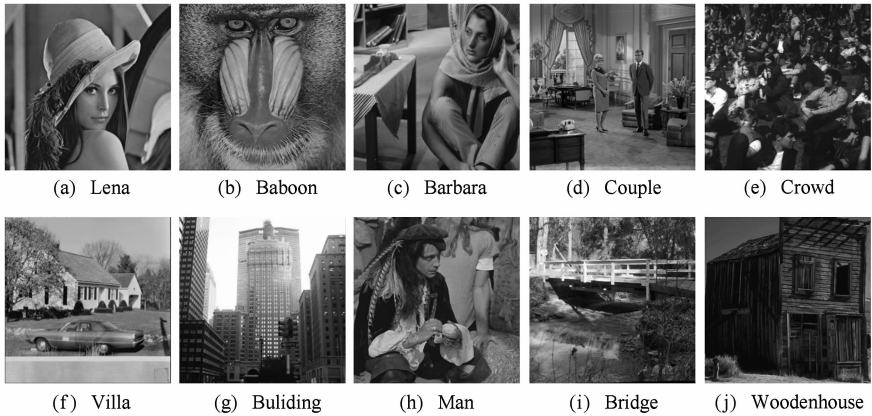


Fig. 5 Commonly used images in the literature
图 5 各类文献中常用的图像集合

3.2 预测准确性

在常用图像集合中的图像上施加噪声级别为 5~50、间隔为 5 的噪声, 统计各对比算法在每幅图像上的预测结果, 以噪声水平预测值与真实值之间的绝对偏差值作为衡量算法预测准确性的指标. 限于篇幅, 这里仅给出在 Couple 图像的预测结果, 预测结果排名前 3 的算法加粗标出, 准确性最高的另用下划线标出. 由表 2 可知, Immerkær 算法、Zoran

算法、Chen 算法、Xu 算法、CNN+SVR 算法和 CNN+AdaSVR 算法在 Couple 图像上的预测结果较其他对比算法还不够理想; Liu 算法、Rakhshanfar 算法、CNN+BP 算法和所提出的 CNN+AdaBP 算法在各个噪声级别下表现出良好的预测准确性. 从整体上看, Rakhshanfar 算法的预测准确性在不同噪声水平值上变化较大, 稳定性较差; Liu 算法在各噪声水平下尤其是低水平噪声条件下预测性能最好,

Table 2 Estimation Results Obtained with Different Algorithms on Couple Image at Different Noise Levels
表 2 不同噪声水平下各算法在 Couple 图像上的评估结果

Methods	Noise Levels									
	5	10	15	20	25	30	35	40	45	50
Immerkær	5.64	10.40	15.39	20.38	25.51	30.46	35.47	40.32	45.48	50.56
Zoran	4.13	9.13	14.12	19.04	23.75	38.83	33.78	38.83	43.59	48.72
Ref [25]	4.96	9.89	14.90	19.77	24.81	29.90	34.69	39.49	44.72	49.47
Ref [11]	5.49	10.43	15.29	20.41	25.26	30.29	35.59	40.40	45.26	50.61
Ref [10]	4.93	10.07	14.52	18.37	24.78	30.09	35.18	40.40	44.99	50.05
Ref [1]	4.18	9.28	14.52	19.39	27.17	30.98	35.04	40.23	44.06	50.30
CNN+SVR	8.09	12.99	17.42	21.40	26.09	30.59	35.05	38.94	43.04	46.36
CNN+AdaSVR	8.69	13.96	18.28	22.81	27.35	31.51	35.95	40.23	44.72	48.65
CNN+BP	4.74	9.63	14.70	20.12	25.05	30.32	35.15	40.13	45.09	49.38
CNN+AdaBP	4.56	9.68	14.78	19.80	25.00	29.15	34.68	40.05	44.95	49.49

Note: The top 3 results are bold, and the best values are underlined.

但 Liu 算法是以执行时间为代价才获得高准确性的, CNN+BP 算法和 CNN+AdaBP 算法的预测准确性与 Liu 算法的相差不大,但执行效率则具有显著优势,因此所提出的 CNN+AdaBP 算法综合性能更好.

在单幅图像上的噪声水平估计值的对比并不能全面地评价某个算法的鲁棒性. 因此,对于图 5 中的 10 幅测试图像,计算各个算法对它们噪声水平的估计值与真实值之间的均方根误差(root mean squared error, RMSE),结果列在表 3 中. 在各噪声水平下,前 3 个最好的算法加粗标出,最小的 RMSE 值另用下划线标出. 从表 3 中的数据可以看出:各算法性能与在单幅图像上情况近似, CNN+SVR 和 CNN+

AdaSVR 算法的预测准确性比较差,未进入前 3 名,这说明 SVR 不适合用于完成噪声水平值的映射任务,而 Adaboost 技术也不能显著提高 SVR 的映射能力;Liu 算法、CNN+BP 算法和 CNN+AdaBP 算法的均方根误差排名前 3,其中 Liu 算法在低噪声水平条件下的均方根误差非常小(以执行时间为代价),在中高噪声水平条件下的均方根误差相对于低噪声水平却要大许多,稳定性不够好;CNN+BP 算法整体上均方根误差的平均值比 CNN+AdaBP 算法稍差. CNN+AdaBP 算法在各个级别的噪声条件下的均方根误差都比较小,且预测准确性比较稳定. 这说明了 CNN+AdaBP 算法利用 Adaboost 技术提升 BP 映射模型的性能是成功的,这是因为 CNN

Table 3 RMSE Between Estimated Results and Ground Truths on Ten Commonly Used Images

表 3 各算法在 10 幅常用图像集合上预测结果的均方根误差

Methods	Noise Levels										Average Value
	5	10	15	20	25	30	35	40	45	50	
Immerkær	2.08	1.55	1.26	1.06	0.95	0.87	0.84	0.82	0.84	0.90	1.12
Zoran	2.34	1.28	1.19	1.37	1.60	1.32	1.38	1.50	1.49	1.81	1.53
Ref [25]	0.07	0.09	0.16	0.16	0.30	0.30	0.32	0.35	0.44	0.42	0.26
Ref [11]	2.90	1.91	1.48	1.23	0.97	0.74	0.93	0.59	0.64	0.75	1.21
Ref [10]	0.59	0.55	0.48	0.30	0.59	0.34	0.59	0.56	0.94	0.85	0.58
Ref [1]	0.87	1.31	1.11	0.84	2.48	1.44	0.84	1.03	0.97	0.73	1.16
CNN+SVR	2.84	2.85	2.30	1.82	1.42	0.85	0.44	0.73	1.95	3.50	1.87
CNN+AdaSVR	3.31	3.42	3.02	2.55	2.05	1.61	1.15	0.59	0.46	1.16	1.93
CNN+BP	0.60	0.30	0.41	0.36	0.37	0.36	0.35	0.31	0.43	0.74	0.42
CNN+AdaBP	0.35	0.35	0.38	0.39	0.36	0.38	0.37	0.33	0.25	0.52	0.37

Note: The top 3 results are bold, and the best values are underlined.

Table 4 RMSE Between Estimated Results and Ground Truths on Fifty Images from BSD Database

表 4 不同噪声水平下不同算法在 50 幅 BSD 图像上预测结果的均方根误差

Methods	Noise Levels										Average Value
	5	10	15	20	25	30	35	40	45	50	
Immerkær	2.81	2.05	1.65	1.42	1.28	1.17	1.18	1.09	1.11	1.19	1.50
Zoran	1.54	1.49	1.52	1.57	1.59	1.55	1.47	1.62	1.63	1.65	1.56
Ref [25]	0.14	0.22	0.22	0.22	0.25	0.27	0.38	0.38	0.46	0.54	0.31
Ref [11]	4.42	3.32	2.54	1.89	1.56	1.49	1.24	1.27	1.09	1.07	1.99
Ref [10]	1.98	1.18	0.92	0.70	0.73	1.29	0.91	1.02	0.74	1.01	1.05
Ref [1]	1.39	1.46	1.20	1.09	2.67	1.57	1.20	1.11	1.12	1.10	1.39
CNN+SVR	2.38	2.38	2.03	1.64	1.36	1.12	0.95	1.24	2.11	3.69	1.89
CNN+AdaSVR	3.22	3.24	2.94	2.55	2.19	1.82	1.35	0.97	0.81	1.22	2.03
CNN+BP	0.52	0.56	0.71	0.78	0.81	0.75	0.78	0.78	0.83	0.74	0.73
CNN+AdaBP	0.45	0.57	0.70	0.75	0.77	0.79	0.74	0.77	0.73	0.62	0.69

Note: The top 3 results are bold, and the best values are underlined.

+AdaBP 算法中 BP 网络弱预测器本身具有一定的预测能力,且适当个数的弱预测器之间能形成互补,在 Adaboost 框架下组合利用多个 BP 弱预测器后其预测准确性得到进一步提升.

为了验证算法对测试图像的鲁棒性,随机选取了 BSD 数据库 50 幅图像作为测试集,分别给图像加入级别为 5~50、间隔为 5 的高斯噪声,并计算各个算法预测结果与真实值之间的均方根误差,列在表 4 中(排名标记方法与上文相同).由表 4 中可知:各个算法的预测准确性在难度很大的纹理图像上有所下降,但是排名情况总体与表 3 类似,所提出的 CNN+AdaBP 算法仍然具有最好的性能.

为了验证 CNN+AdaBP 算法的泛化能力,对 BSD 数据库中的 50 幅图像添加 2.5,7.5,12.5,17.5,22.5,27.5,32.5,37.5,42.5,47.5 这 10 个在训练模型时未用到的噪声水平值,并计算对应的均方根误差,结果如表 5 所示.

从表 5 可以看出:所提出的 CNN+AdaBP 算法能够很好地预测噪声级别在 0~50 之间的任意噪声水平值,且算法的预测准确性足够稳定.综上所述,CNN+AdaBP 算法具有良好且稳定的预测准确性,其原因在于该算法是根据自然图像统计规律在大量自然图像上训练学习获得的.

Table 6 PSNR of Image Denoised by BM3D Algorithm Using Ground Truth and Estimated Noise Levels

表 6 BM3D 分别使用真实噪声值和预测噪声值的降噪结果 dB

Methods	Noise Levels									
	5	10	15	20	25	30	35	40	45	50
True Noise Level	39.92	36.56	34.66	33.33	32.37	31.45	30.68	30.02	29.55	29.11
Estimated Noise Level	39.87	36.57	34.65	33.32	32.36	31.44	30.70	30.17	29.56	29.10
Absolute Deviation	0.05	0.01	0.01	0.01	0.01	0.01	0.02	0.15	0.01	0.01

从表 6 中的数据可以看出,使用 CNN+AdaBP 算法预测出的噪声水平值作为参数的降噪效果与使用真实噪声水平值作为参数的降噪效果差别很小,这充分说明了 CNN+AdaBP 算法的实用性和有效性.

3.4 执行时间的对比

对于一个图像噪声水平估计算法来说,不应仅仅根据算法的预测准确性来评判该算法性能是否良好,执行效率也是非常重要的评价指标.为了充分证明所提出的算法性能良好,笔者对各个算法的执行时间进行了比较.将各算法在 1 幅大小为 512×512 的 Lena 图像上重复执行 20 次,计算算法估计该图像噪声水平值的平均执行时间,各算法在不同噪声

Table 5 RMSE of the Estimated Results at Different Noise Levels

表 5 本文算法在 50 幅 BSD 图像上预测结果的均方根误差

Noise Levels	RMSE
2.5	0.49
7.5	0.53
12.5	0.67
17.5	0.78
22.5	0.83
27.5	0.91
32.5	0.78
37.5	0.79
42.5	0.77
47.5	0.80

3.3 BM3D 算法降噪

为了进一步验证 CNN+AdaBP 算法的实际应用效果,用经典 BM3D 降噪算法对分别加入不同噪声级别的 Lena 图像进行降噪,表 6 列出了分别使用真实噪声水平值和根据 CNN+AdaBP 算法预测出的噪声值作为 BM3D 算法的输入参数进行降噪后的图像的峰值信噪比(peak signal to noise ratio, PSNR).

级别下平均执行时间如表 7 所示.从表 7 可以看出:Zoran 算法和 Liu 算法的执行时间非常长,不适合应用于时间要求比较严格的系统中.相比而言,Chen 算法和 Rakhshanfar 算法的执行效率一般.由表 2~4 中的数据可知,这 2 个算法的预测精度也一般.Immerkaer 算法和 Xu 算法的执行时间非常快,但是从表 3~4 中的数据可知,这 2 种算法的评价准确性比较其他算法要差一些.CNN+SVR 算法、CNN+AdaSVR 算法、CNN+BP 算法和 CNN+AdaBP 算法作为基于训练的 MNLE 算法,模型一旦训练完成,执行速度处于所有参与比较算法中的前列.综上所述,所提出的 CNN+AdaBP 算法在预测准确率和执行效率这 2 个方面的综合性能较其他算法更优.

Table 7 Average Execution Time of Different NLE Methods at Different Noise Levels

表 7 各算法在不同噪声级别下的平均执行时间

s

Methods	Noise Levels									
	5	10	15	20	25	30	35	40	45	50
Immerkær	2.9	3.0	2.8	2.8	2.9	2.8	2.8	2.8	2.9	3.0
Zoran	1011.8	1008.8	1013.9	992.8	986.4	999.5	1030.0	1019.1	995.7	987.0
Ref [25]	607.9	621.5	632.4	637.8	644.3	647.3	656.8	660.4	649.4	651.9
Ref [11]	162.9	166.0	165.4	165.5	165.1	164.9	165.8	165.7	165.6	166.1
Ref [10]	17.0	20.5	22.8	24.9	27.0	28.0	28.5	30.2	30.8	31.4
Ref [1]	15.0	14.9	15.1	15.1	15.2	14.8	14.8	15.0	15.2	15.2
CNN+SVR	17.7	17.1	17.1	17.0	17.0	16.9	17.0	17.0	17.1	17.1
CNN+AdaSVR	18.03	18.09	18.05	18.05	18.03	18.05	18.07	18.06	18.06	18.09
CNN+BP	19.87	18.90	17.95	18.82	18.60	17.01	17.45	18.66	18.60	18.77
CNN+AdaBP	19.4	20.0	20.3	21.5	22.7	21.7	21.8	19.9	20.2	21.5

4 总 结

传统基于单幅图像的 SNLE 算法中将图像内容与噪声信号分离的关键模块在算法设计上比较复杂,导致执行效率比较低.为此,提出了一种利用 CNN 在多幅图像上通过训练学习实现快速分离噪声信号获得噪声映射图的新方法.然后,利用广义高斯分布 GGD 模型为噪声映射图建模,并以模型参数值作为反映噪声严重程度的 NLAf 特征值.最后,训练改进后的 BP 网络模型实现从 NLAf 特征值到噪声水平值的预测任务.与现有的算法相比,CNN+AdaBP 算法从提取高质量的 NLAf 特征和构建非线性映射能力强大的映射模型 2 个方面入手,在保持高效率的前提下,进一步提高了 MNLE 这类算法的预测准确率,获得了一种预测准确性和执行效率的综合性能更好的噪声水平评估算法.

参 考 文 献

[1] Xu Shaoping, Zeng Xiaoxia, Tang Yiling. Fast noise level estimation algorithm based on two-stage support vector regression [J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(3): 447-458 (in Chinese)
(徐少平,曾小霞,唐玮玲.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)

[2] Wang Lingfeng, Huang Zehao, Gong Yongchao, et al. Ensemble based deep networks for image super-resolution [J]. Pattern Recognition, 2017, 68(8):191-198

[3] Ren Jingjing, Fang Xianyong, Chen Shangwen, et al. Image deblurring based on fast convolutional neural networks [J]. Journal of Computer-Aided Design & Computer Graphics, 2017, 29(8): 1444-1456 (in Chinese)
(任静静,方贤勇,陈尚文,等.基于快速卷积神经网络的图像去模糊[J].计算机辅助设计与图形学学报,2017,29(8):1444-1456)

[4] Dabov K, Foi A, Katkovnik V, et al. Image denoising by sparse 3D transform-domain collaborative filtering [J]. IEEE Transactions on Image Processing, 2007, 16(8): 2080-2095

[5] Huang Dongmei, Dai Liang, Wei Lifei, et al. A secure outsourced fusion denoising scheme in multiple encrypted remote sensing images [J]. Journal of Computer Research and Development, 2017, 54(10): 2378-2389 (in Chinese)
(黄冬梅,戴亮,魏立斐,等.一种安全的多帧遥感图像的外包融合去噪方案[J].计算机研究与发展,2017,54(10):2378-2389)

[6] Dong Weisheng, Zhang Lei, Shi Guangming, et al. Nonlocally centralized sparse representation for image restoration [J]. IEEE Transactions on Image Processing, 2013, 22(4): 1620-1630

[7] Immerkær J. Fast noise variance estimation [J]. Computer Vision and Image Understanding, 1996, 64(2): 300-302

[8] Zoran D, Weiss Y. Scale invariance and noise in natural images [C] //Proc of the 12th IEEE Int Conf on Computer Vision (ICCV 2009). Piscataway, NJ: IEEE, 2009: 2209-2216

[9] Katase H, Yamaguchi T, Fujisawa T, et al. Image noise level estimation by searching for smooth patches with discrete cosine transform [C] //Proc of the 59th IEEE Int Midwest Symp on Circuits and Systems. Piscataway, NJ: IEEE, 2016. DOI:10.1109/MWSCAS.2016.7870110

[10] Rakhshanfar M, Amer M A. Estimation of Gaussian, Poissonian-Gaussian, and processed visual noise and its level function [J]. IEEE Transactions on Image Processing, 2016, 25(9): 4172-4185

- [11] Chen Guangyong, Zhu Fengyuan, Heng P A. An efficient statistical method for image noise level estimation [C] //Proc of the 18th IEEE Int Conf on Computer Vision (ICCV 2015). Piscataway, NJ: IEEE, 2015: 477-485
- [12] Xu Shaoping, Hu Lingyan, Yang Xiaohui. Quality-aware features-based noise level estimator for block matching and three-dimensional filtering algorithm [J]. Journal of Electronic Imaging, 2016, 25(1): 013029
- [13] Xu Shaoping, Zeng Xiaoxia, Jiang Yinnan, et al. A multiple image-based noise level estimation algorithm [J]. IEEE Signal Processing Letters, 2017, 24(11): 1701-1705
- [14] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] //Proc of the 27th IEEE Conf on Computer Vision and Pattern Recognition (CVPR 2014). Piscataway, NJ: IEEE, 2014: 580-587
- [15] Taigman Y, Yang Ming, Ranzato M, et al. Web-scale training for face identification [C] //Proc of the 18th IEEE Conf on Computer Vision and Pattern Recognition (CVPR 2015). Piscataway, NJ: IEEE, 2015: 2746-2754
- [16] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651
- [17] Chang Liang, Deng Xiaoming, Zhou Mingquan, et al. Convolutional neural networks in image understanding [J]. Acta Automatica Sinica, 2016, 42(9): 1300-1312 (in Chinese)
(常亮, 邓小明, 周明全, 等. 图像理解中的卷积神经网络 [J]. 自动化学报, 2016, 42(9): 1300-1312)
- [18] Zhou Feiyan, Jin Linpeng, Dong Jun. Review of convolutional neural network [J]. Chinese Journal of Computers, 2017, 40(6): 1229-1251 (in Chinese)
(周飞燕, 金林鹏, 董军. 卷积神经网络研究综述 [J]. 计算机学报, 2017, 40(6): 1229-1251)
- [19] Nair V, Hinton G E. Rectified linear units improve restricted Boltzmann machines [C] //Proc of the 27th Int Conf on Int Conf on Machine Learning. New York: ACM, 2010: 807-814
- [20] Rawat W, Wang Zenghui. Deep convolutional neural networks for image classification: A comprehensive review [J]. Neural Computation, 2017, 29(9): 2352-2449
- [21] Li Yanghao, Wang Naiyan, Shi Jianping, et al. Adaptive batch normalization for practical domain adaptation [J]. Pattern Recognition, 2016, 80: 109-117
- [22] Mallat S G. A theory for multiresolution signal decomposition: The wavelet representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(7): 674-693
- [23] Sun Bo, Chen Songcan, Wang Jiandong, et al. A robust multi-class AdaBoost algorithm for mislabeled noisy data [J]. Knowledge-Based Systems, 2016, 102(5): 87-102
- [24] Arbelaez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916
- [25] Liu Xinhao, Tanaka M, Okutomi M. Single-image noise level estimation for blind denoising [J]. IEEE Transactions on Image Processing, 2013, 22(12): 5226-5237



Xu Shaoping, born in 1976. PhD, professor and PhD supervisor. His main research interests include digital image processing machine vision, and virtual surgery simulation.



Liu Tingyun, born in 1996. Master candidate. Her main research interests include digital image processing, machine vision.



Li Chongxi, born in 1994. Master candidate. His main research interests include digital image processing, machine vision.



Tang Yiling, born in 1977. PhD candidate, lecturer. Her main research interests include digital image processing and machine vision.



Hu Lingyan, born in 1978. PhD, professor and PhD supervisor of Nanchang University. Her main research interests include digital image processing machine vision, and virtual surgery simulation.