

DOI:10.16356/j.1005-2615.2020.05.015

基于样本错配训练的图像 PRNU 噪声提纯方法

郝昕泽¹, 肖延辉¹, 田华伟¹, 张明旺²

(1. 中国人民公安大学国家安全学院, 北京, 100038; 2. 四川警察学院科研所, 泸州, 646000)

摘要: 基于传感器模式噪声的图像来源鉴别算法的关键在于如何获取高质量的光响应非均匀性(Photo-response non-uniformity, PRNU)噪声, 目前大多数增强 PRNU 噪声质量的算法以及出于实际应用的目的对其压缩的算法几乎是在人工假设模型的基础上实现的。本文提出了一种基于样本错配训练的图像 PRNU 噪声提纯方法, 该方法使用了深度堆叠自编码器网络, 并设计了一种样本错配的技术对其进行训练。这种基于样本错配训练的端到端神经网络的使用有效避免了人工设计算法的局限性, 对图像 PRNU 噪声进行了有效提纯, 进而提升了其用于图像来源鉴别时的性能。在 Dresden 图像数据集上进行的比较实验结果表明, 提纯后的 PRNU 噪声具有更好的性能。

关键词: 数字图像; 传感器模式噪声; 自动编码器; 光响应非均匀性

中图分类号: TP391 **文献标志码:** A **文章编号:** 1005-2615(2020)05-0783-09

Purification Scheme of Image PRNU Noise Based on Cross-Matching Sample Training

HAO Xinze¹, XIAO Yanhui¹, TIAN Huawei¹, ZHANG Mingwang²

(1. School of National Security, People's Public Security University of China, Beijing, 100038, China; 2. Institute of Research, Sichuan Police College, Luzhou, 646000, China)

Abstract: The key to the image source identification algorithm based on sensor pattern noise is how to obtain high-quality PRNU noise. At present, most of the algorithms include to improve the quality of PRNU noise and compress them for purposes of practical application based on artificially assumed models. This paper proposes a purification scheme of image PRNU noise based on cross-matching sample training. The scheme uses a deep stacked auto-encoder network and designs a technology of sample cross-matching to train it. The use of this end-to-end deep neural network based on cross-matching sample training effectively avoids the limitations of manual design algorithms and purifies the image PRNU noise. Then the performance of PRNU noise is improved for image source identification. The results of the comparative experiment on the Dresden image dataset show that the purified PRNU noise has better performance.

Key words: digital image; sensor pattern noise; auto-encoder; photo-response non-uniformity (PRNU)

当前, 每天通过社交媒体生成和分享的视频和图像数量正在迅速增长, 这主要得益于便捷的视听

记录技术和移动互联网技术。而这些技术也为不法分子拍摄、传播含有非法信息的图像、视频提供

基金项目: 国家自然科学基金(61772539, 61972405)资助项目; 四川省科技计划(2018JY0521)资助项目; 公安部技术研究中心(2017JSYJC01)资助项目。

收稿日期: 2020-06-30; **修订日期:** 2020-08-01

通信作者: 田华伟, 副教授, E-mail: hwtian@live.cn。

引用格式: 郝昕泽, 肖延辉, 田华伟, 等. 基于样本错配训练的图像 PRNU 噪声提纯方法[J]. 南京航空航天大学学报, 2020, 52(5): 783-791. HAO Xinze, XIAO Yanhui, TIAN Huawei, et al. Purification scheme of image PRNU noise based on cross-matching sample training[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2020, 52(5): 783-791.

了便利。在数字图像取证中,如何评估图像的真实性和完整性至关重要,司法鉴定部门面临的一个常见问题就是确定图像来源的归属,即判断图像是由哪一个相机拍摄。

在过去的十几年中,学者们提出了许多方法用于图像来源鉴别和图像伪造检测,其中基于传感器模式噪声(Sensor pattern noise, SPN)^[1]的方法取得了较大的成功。图像传感器在成像时会不可避免地产生一种 SPN, SPN 主要由光响应非均匀性(Photo-response non-uniformity, PRNU)噪声组成,并且 PRNU 噪声是由图像传感器自身制造工艺的缺陷和硅晶片的非均匀性引起的,使得 PRNU 噪声对于每一个成像设备是唯一的,且不会受外界环境的影响,故该模式噪声非常适合作为成像设备的固有“指纹”。Lukas 等^[1]首次提出将 PRNU 噪声应用于相机源鉴别,PRNU 噪声对于每个成像设备的唯一性是其作为设备指纹的先天优势,因此随后吸引了大量的关注和研究。基于 SPN 的图像来源鉴别算法的关键在于如何获取高质量的 PRNU 噪声,常用的去噪算法有基于小波变换(Wavelet-based filter, WB)^[2]的方法、三维块匹配滤波的方法(Block matching and 3D filtering, BM3D)^[3]、三方加权稀疏编码模型(Trilateral weighted sparse coding model, TMSM)^[4]的方法等,这些算法可以去掉图像中的大部分噪声,有利于从图像的噪声残差中进一步提取 PRNU 噪声。但是从噪声残差中提取 PRNU 噪声的过程还是容易受到包括背景噪声、CFA 内插噪声和 JPEG 压缩噪声等其他噪声的干扰,针对上述问题,学者们研究如何通过加强从噪声残差中提取的 PRNU 噪声质量以达到提升图像来源鉴别精确度的目的,提出零均值化法(Zero-mean, ZM)^[5]、频谱均衡算法(Spectrum equalization algorithm, SEA)^[6]等方法。同时出于将 SPN 实际应用于图像来源鉴别的目的,一些学者们致力于在保证图像来源鉴别精确度的同时,探索如何降低 PRNU 噪声的计算和储存成本的方法,提出了构造紧凑型 PRNU 噪声的方法,包括重要成分保留法^[7]、基于主成分分析的方法(Principal components analysis, PCA)^[8]等。

尽管上述算法从不同角度出发,提升了 PRNU 噪声在图像来源鉴别中的识别准确率,但是上述方法几乎都是依靠人工建模的方法来实现其适用性,鉴于近些年来深度学习技术在图像特征提取、图像

内容识别等领域取得的革命性突破,并且已经成功解决了一些复杂的应用问题,因此本文提出了一种基于样本错配训练的图像 PRNU 噪声提纯方法。本文首先设计了一种深度 SAE 网络,自编码器(Auto-encoder, AE)是深度学习中无监督学习领域中的一种人工神经网络,其可以通过有效地压缩和编码数据,然后学习如何从压缩简化后的编码表示重新构造数据,使其尽可能接近原始数据。其次提出了与其对应的“样本错配策略”的训练方法,即改变传统自编码器输入与自身对应训练的方式,采用输入与其同类相似样本交叉训练的方式。通过自编码器模型对图像中的 PRNU 噪声进行压缩-解压缩,使得重构后的 PRNU 噪声在保持自身完整性的同时,去除无关噪声,达到提升 PRNU 噪声性能的目的。通过在 Dresden 图像数据集上进行比较实验表明,本文提出的图像 PRNU 噪声提纯方法可以有效提升其在图像来源鉴别任务中的精确度,并针对不同的经典去噪算法从图像中提取的 PRNU 噪声均具有较好的提纯效果。

1 自编码器与深度堆叠自编码器

1.1 自编码器模型

自编码器是一种典型的通过无监督的方式来学习数据表示的模型,最简单的自编码器是只有一个隐藏层的人工神经网络,其结构可以分为两个部分,即编码器与解码器。编码器用数学公式表示为

$$X = f(x) = \sigma_e(W_e \cdot x + b_e) \quad (1)$$

式中:编码器的输入为 x ;隐藏层的激活(非线性)函数为 $\sigma_e(\bullet)$; W_e 和 b_e 分别为连接编码器输入层与隐藏层的权重与偏置参数; X 为编码器的输出。

解码器表示为

$$\hat{x} = g(X) = \sigma_d(W_d \cdot X + b_d) \quad (2)$$

式中:解码器的输入为 X ;输出层的激活(非线性)函数为 $\sigma_d(\bullet)$; W_d 和 b_d 分别为连接隐藏层和解码器输出层的权重与偏置参数; \hat{x} 为解码器的输出。

一般来说,若隐藏层的维度 p 小于原始输入的维度 q ,编码器学习如何减少输入数据的维度,并将输入数据压缩为编码后的数据表示,解码器学习如何重构数据,使其解码后的表示形式与原始输入尽可能接近。因此自编码器训练的学习目标就是最小化编码器输入与解码器输出的重构误差,其损失函数可以表示为

$$J(W,b) = \frac{1}{N} \sum_{n=1}^N \|x^{(n)} - \hat{x}^{(n)}\|_2^2 \quad (3)$$

式中: n 为训练样本的个数; $x^{(n)}$ 为第 n 个输入样本; $\hat{x}^{(n)}$ 为第 n 个重构样本; $\|\cdot\|_2^2$ 为均方误差^[9]。在自编码器的训练过程中,通过优化以上损失函数,自编码器可以将输入自身 x 表示为低维的数据 X ,并且可以将 X 重构为与输入 x 相近的 \hat{x} 。模型结构如图 1 所示。

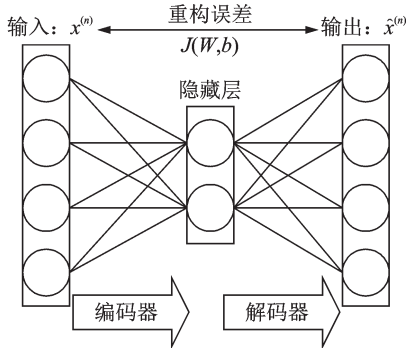


图 1 自编码器模型

Fig.1 Auto-encoder model

1.2 深度堆叠自编码器模型

对于很多数据,仅使用具有 3 层神经网络的自编码器还不足以获取一种好的数据表示。Bengio 等^[10]提出使用逐层堆叠的方式来训练一个深层的自编码器,称为堆叠自编码器(Stacked auto-encoder, SAE)。SAE 可以看作是许多个基本的自编码器模块的组合,第一个编码器训练完成后,将第一个自编码器的隐藏层输出看作下一个自编码器的编码输入,然后训练下一个自编码器,依次循环,最后堆叠形成深层的自编码器网络,此时每层的参数可以用作进行优化的深层自编码器网络的参数初始化。这种贪婪的逐层训练过程已经被证明比深层网络的随机初始化会产生更好的局部极小值,实现更好的泛化。

2 基于样本错配训练的深度 SAE

自编码器是一种经过训练尝试从输入复制到输出的神经网络,主要应用于数据降维、图像降噪和信息检索任务^[11-12]。自编码器同样也可以看作一个生成模型,由于其不是简单地使输入到输出完全相等,而是在输入经过编码器和解码器后保留有用的信息或者根据需要产生新的表示,所以本文将其应用至图像 PRNU 噪声的提纯,旨在通过对图像 PRNU 噪声进行有损压缩,使重构后的 PRNU 噪声能够进一步去除含有的无关噪声。

2.1 基于样本错配训练的单个自编码器

如果对自编码器没有任何额外的约束,那么它将学习输入到输出的恒等映射,即一个恒等函数。为了使自编码器学习到更为鲁棒的特征,Vincent 等^[13]提出了一种降噪自编码器,此时自编码器网络的输入变成添加了随机噪声的数据,而损失函数的目标是最小化原始无噪声的数据与重构数据之间的误差,即期望输出无噪声的数据。通过强加噪声这种约束,使得降噪自编码器学习更为鲁棒的数据表示,实验证明降噪自编码器泛化能力要强于一般的自编码器模型。

为了实现对 PRNU 噪声本身的提纯处理,本文基于降噪编码器的思想,对自编码器模型的训练方式进行了改进,提出了一种样本错配的训练方式。由于 PRNU 图像指纹本身是一种噪声,研究者们设计算法提取的 PRNU 噪声中还会含有一些无关噪声,因此无法获取真实的 PRNU 噪声,而通过同一相机拍摄的不同照片所提取的 PRNU 噪声尽管具有一定的区别,但其本质上仍然都是该相机的图像指纹,为了挖掘某一相机各个 PRNU 图像指纹内在的联系,获取某一相机更为泛化的 PRNU 图像指纹,本文对自编码器的期望输出做出了改变。传统自编码器的目标函数是最小化输入与其自身的重构误差 $\|x^{(n)} - \hat{x}^{(n)}\|_2^2$,即训练方式是输入与其自身对应匹配,本文改进后的自编码器的目标函数是最小化输入与其同类相似样本的重构误差 $\|x^{(m)} - \hat{x}^{(n)}\|_2^2$, $x^{(m)}$ 表示同一相机的 PRNU 噪声样本,即训练方式是输入与其同类相似样本交叉匹配,如图 2 所示。通过采取样本错配的训练方式,可以强制自编码器学习除自身以外的特征,使得经过自编码器模型的 PRNU 噪声样本更泛化,获取的 PRNU 噪声特征更具鲁棒性,从而实现提纯处理。

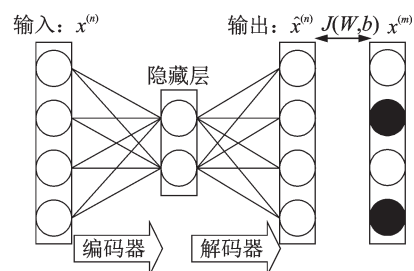


图 2 基于样本错配训练的自编码器模型

Fig.2 Auto-encoder model based on cross-matching sample training

2.2 深度SAE及其训练方法

为了达到对PRNU噪声更进一步的提纯处理,可以使用更深层的自编码器。基于本文提出的样本错配的训练方式,因此本文在经典SAE的基础上进行了改进,提出适合本文任务的堆叠方式以及训练方式。

本文将两个基本自编码器的模型作为1个模块,模块内的每个训练模型都是独立的,但将模块依次堆叠后整个的训练过程是相互关联的。1个模块内两个自编码器各层的神经元个数相同,第1个自编码器的输入与期望输出保持一致,通过最小化重构误差 $\|x^{(n)} - \hat{x}^{(n)}\|^2$ 进行无监督学习,训练完成后,提取第1个自编码器的输出作为第2个自编码器的输入,并且将原始输入打乱顺序,形成第2个自编码器的期望输出,通过最小化重构误差进行无监督学习,其中 $\hat{x}^{(n)} = g(f(\hat{x}^{(n)}))$,表示第2个自编码器的真实输出,此时第1个模块训练完成,如

图3(a)所示。接下来进行第2个模块的训练,首先取第1个模块中第1个自编码器的隐藏层输出(即 $X^{(n)}$)作为第2个模块内第1个自编码器的输入,再一次与输入自身进行无监督学习,训练完成后,同样提取该自编码器的输出作为该模块内第2个自编码器的输入,此时取出第1个模块中第2个自编码器的隐藏层输出(即 $X^{(m)}$)作为第2个模块内第2个自编码器的期望输出,同样通过无监督学习的方式完成第2个模块的训练,如图3(b)所示。以此类推,在完成第*i*个模块的训练后,其中*i*=1, ..., *N*,依次取各个模块中的第1个自编码器的中间层组成深度SAE模型的编码器部分,之后依次取各个模块中的第2个自编码器的输出层组成深度SAE模型的解码器部分,深度SAE各层网络参数初始化使用预训练过程中的参数,之后再次训练完整的深度SAE网络,完成网络参数的微调,如图4所示。

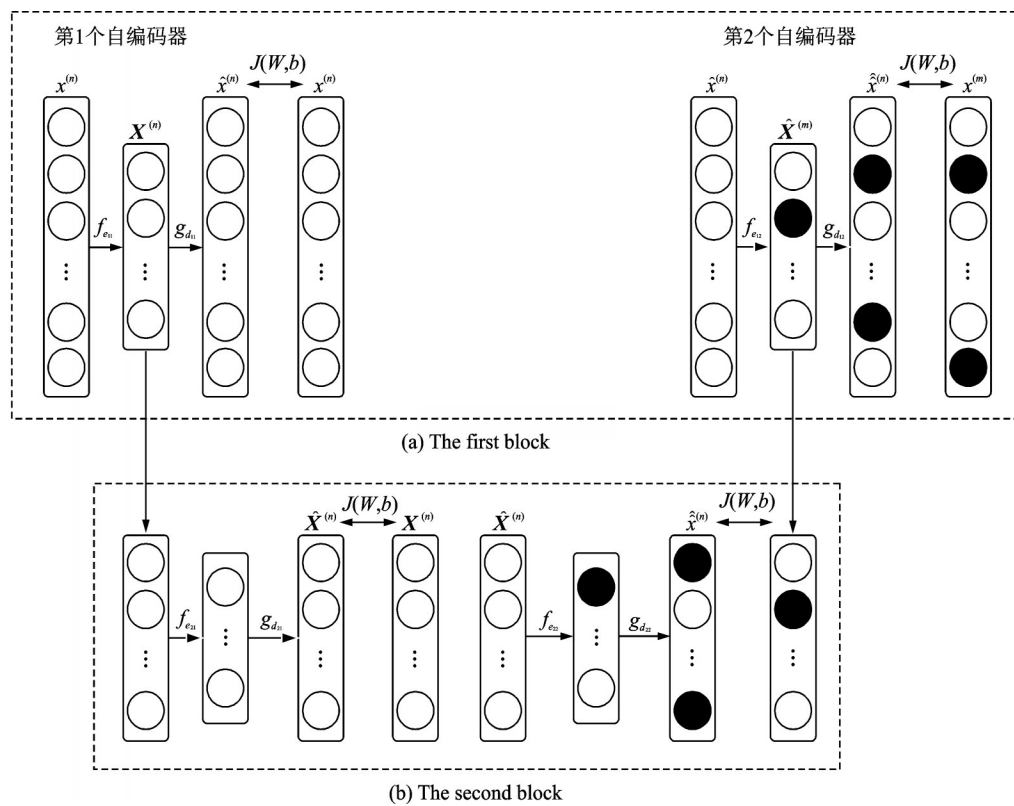


图3 基于样本错配的深度SAE预训练过程

Fig.3 Pre-training process of deep stack auto-encoder based on cross-matching sample

由图4可知,深度SAE模型的编码器部分实现了对输入数据的逐层压缩,进一步完善了对深层抽象特征的提取与表示,其中编码器的输出还可用于后续压缩等任务的输入,解码器部分趋于重构与输入相似的同类样本,使得重

构样本除了含有输入本身的重要特征外,将通过解码过程进一步结合其同类样本的特征,以克服传统自编码器输入映射至输出的特征表达单一性,实现重构样本的强泛化性,以满足本文的任务要求。

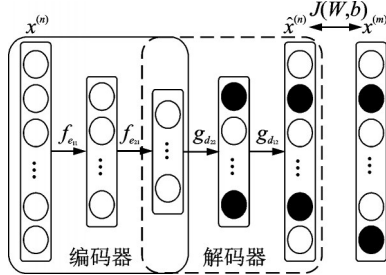


图 4 基于样本错配的深度 SAE 微调过程及模型结构
Fig.4 Fine-tuning process of deep stack auto-encoder based on cross-matching sample and model structure

3 算法步骤

将基于样本错配训练的图像 PRNU 噪声提纯方法应用于图像来源鉴别任务中, 具体算法步骤如下:

(1) 获取相机 PRNU 噪声。对由同一相机拍摄的多幅图像 I^i , 其中 $i=1, \dots, N$, 进行降噪处理, 之后计算得到多幅图像的噪声残差 W , 最后对多幅图像的噪声残差 W 进行最大似然估计, 估算出该相机的一个 PRNU 噪声 $\hat{K}^{[14]}$, 即

$$W^i = I^i - F(I^i) \quad i = 1, \dots, N \quad (4)$$

$$\hat{K} = \frac{\sum_{i=1}^N W^i I^i}{\sum_{i=1}^N (I^i)^2} \quad (5)$$

式中: $F(I^i)$ 为对第 i 幅图像进行降噪处理; W^i 为第 i 幅图像的噪声残差。

(2) 生成自编码器训练集。重复步骤(1), 获取同一相机的多个 PRNU 噪声图像, 形成深度 SAE 的训练集。

(3) 训练深度 SAE 模型。训练深度自编码器, 不断迭代优化, 形成可对 PRNU 噪声提纯的模型。

(4) 获取降噪后的 PRNU 噪声。重复步骤(1), 再次提取一个 PRNU 噪声图像 \hat{K} , 将其输入深度 SAE 模型中, 得到重构的 PRNU 噪声。

$$\hat{K}_{ae} = g(f(\hat{K})) \quad (6)$$

式中 \hat{K}_{ae} 是自编码器重构后的 PRNU 噪声。

(5) 提取待检测图像的噪声残差。根据步骤(1)中式(4)获取待检测图像 I 的噪声残差 W 。

(6) 分别计算原始 PRNU 噪声 \hat{K} , 重构的 PRNU 噪声 \hat{K}_{ae} 与待检测图像的噪声残差 W 之间的峰值相关能量比 (Peak to correlation energy ratio, PCE) 值^[15]。

文献[15]将 PCE 值作为检验待检图片是否来

源于特定相机的依据, 在计算得出 PCE 值后, 对该值进行阈值判断, 若大于等于阈值, 则认为待检测图片来源于具有 PRNU 噪声 \hat{K} 的相机。本文所使用 PCE 值计算公式为

$$PCE = \frac{\text{sign}(\rho(s_{\text{peak}}; X, Y)) \rho(s_{\text{peak}}; X, Y)^2}{\frac{1}{mn - |N|} \sum_{s, s \in N} \rho(s; X, Y)^2} \quad (7)$$

式中: $X = I\hat{K}$, $Y = W$, 二者维度为 $m \times n$; $\rho(s; X, Y)$ 表示 X, Y 之间的相关性函数 (维度同 X, Y); $\rho(s_{\text{peak}}; X, Y)$ 为相关性函数的最大值; $\text{sign}(\rho(s_{\text{peak}}; X, Y))$ 为函数取最大值时的符号; N 为以 s_{peak} 为中心的小块区域 (原文中取 11×11 的正方形区域)。

4 实验结果与分析

本节将在 Dresden 图像数据集上测试本文提出的深度 SAE 模型在提升 PRNU 噪声性能方面的表现。将本文所提出的 PRNU 提纯算法分别应用于基于小波变换去噪^[2]的 PRNU 提取算法和基于 BM3D 去噪^[3]的 PRNU 提取算法提取到的 PRNU 噪声, 以验证该模型的提纯能力。另外, 还与频谱均衡提纯算法^[6]进行了比较, 以验证本文算法的先进性。

4.1 训练数据集生成

为了达到提升 PRNU 噪声性能的目的, 需要大量的图像 PRNU 噪声样本以训练深度 SAE 模型。本节实验使用公开的 Dresden 图像数据集以用来生成 PRNU 噪声训练数据, 包含 74 个相机, 每个相机内随机挑选 100 张图片。对于每个相机内的图片, 随机划分两个子数据集, 分别是训练集和验证集, 训练集内含有 50 张图片, 用于生成训练深度自编码器的 PRNU 噪声数据集, 验证集内含有 50 张图片, 用于生成待检测图片的噪声残差。由于原始数据集中的图片分辨率不尽相同, 其次为了提升训练数据集的生成效率, 本文统一将其裁剪为 48×48 、 64×64 和 96×96 这 3 种尺寸分别进行实验。对于每个相机设备的训练集, 从中随机选取 40 张图片提取 1 张 PRNU 噪声图像, 根据不同的排列组合方式, 一共提取 20 000 张 PRNU 噪声图像, 最终生成容量为 $20\,000 \times 74$ 的训练数据集。

4.2 深度 SAE 模型参数配置

本节实验采用 2.2 节提出的深度 SAE 模型训练方法训练完成 3 个模块的训练, 并依次堆叠形成

一个7层的深度自编码器网络,由于PRNU噪声图像的大小与原始图片相同,因此基于3种不同尺寸实验的深度自编码器网络的输入、输出大小分别为 $48 \times 48=2\ 304$ 、 $64 \times 64=4\ 096$ 、 $96 \times 96=9\ 216$ 。中间每层的神经元个数相同,设置为Input $\rightarrow[2\ 048 \rightarrow 1\ 024 \rightarrow 512 \rightarrow 1\ 024 \rightarrow 2\ 048] \rightarrow$ Output,各层之间采用全连接的方式,模块内的各个自编码器隐藏层的激活函数采用缩放指数型线性单元(Scaled exponential linear units, SELU)^[16],该单元具有自归一化的性质,能够使每层的神经元激励值逼近于零均值和单位方差,保证训练过程中梯度不会爆炸或消失,因此可以训练深层的神经网络。为了加速神经网络的收敛,对训练集进行标准化处理,并且神经网络的权值也采用了均值为0、标准差为 $1/m$ 的初始化方法,其中 m 为每层神经元的个数,其次自编码器训练的过程一般采用误差反向传播(Back-propagation, BP)^[17]的方式迭代更新神经网络的参数以达到最小化损失函数的目的,因此本文的优化算法采用了Adabound^[18]算法,笔者在论文中介绍Adabound是Adam^[19]算法的变体,可以使网络在训练前期保持较高的迭代速度,同时在训练后期可以避免Adam自适应学习率算法可能无法收敛的问题,保证了在测试数据上的强泛化能力。

4.3 实验结果分析

为了评价本文提出的模型在提升PRNU噪声性能上的表现,本节实验分别采用受试者工作特征曲线(Receiver operating characteristic curve, ROC)、ROC曲线下面积(Area under ROC curve, AUC)和Kappa统计系数等3个评测指标进行实验结果的分析与比较。

4.3.1 ROC曲线与AUC值

对于每部相机,使用训练集中所有的50张图片生成新的PRNU图像噪声,之后使用深度自编码器模型对其进行重建,得到重建后的PRNU噪声图像,该部相机验证集中的50张图片作为同类样本,其余73部相机验证集中的3 650张图片作为异类样本。最后,在所有的相机上得到 50×74 个同类样本的PCE值和 $3\ 650 \times 74$ 个异类样本的PCE值,对这些PCE值进行阈值判断,对于每一个阈值可计算得到一个真正率值(True positive rate, TPR)和一个假正率值(False positive rate, FPR),从而确定ROC曲线中的一点,因此进行不同阈值

下的判断,可得到坐标系中的多个点,将这些点连起来即可得到完整的ROC曲线,其中

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (9)$$

式中:TP、FN、FP和TN分别表示将同类样本预测为同类样本、将同类样本预测为异类样本、将异类样本预测为同类样本以及将异类样本预测为异类样本的数量。

一般ROC曲线在保持FPR值较小的同时,TPR值越大,即ROC曲线离左上角(0,1)点越近,其代表的算法性能越好。为了更直观评价模型分类性能的优劣,还采用AUC值,即ROC曲线与坐标轴围成的下面积来比较各算法的性能。ROC曲线下面积即AUC值越大,则其对应的算法性能越好。

4.3.2 Kappa统计系数

Kappa统计系数的计算基于混淆矩阵。为了得到混淆矩阵,对于一部相机中的每张待检测图片,分别使用74部相机的PRNU图像噪声与其计算PCE值,通过比较74个PCE值,从而找到最大值对应的相机,则认为待检测图片来源于该部相机。最后将74部相机中的 50×74 张待检测图片依次进行计算,得到混淆矩阵 A 后计算Kappa统计系数数值为

$$\text{Kappa} = \frac{o - e}{T - e} \quad (10)$$

式中: o 为实际被正确分类的样本数,对应于混淆矩阵的对角线和; T 为总样本数; e 为期望被正确分类的样本数,其公式为

$$e = \frac{\sum_{n=1}^N \sum_{i=1}^N A(n, i) \sum_{j=1}^N A(j, n)}{T} \quad (11)$$

式中 N 为相机的个数。Kappa统计系数用来衡量实际预测结果与随机预测结果之间的差别是否具有显著性,Kappa系数值越大,说明模型预测性能越好。

本节实验分别使用基于小波变换的去噪算法和基于BM3D的去噪算法初步从图像中提取PRNU噪声,并同时经过零均值化^[5]处理,作为比较实验的基准PRNU噪声。在3种不同图像尺寸下进行PRNU图像噪声实验,并将本文算法与频谱均衡算法^[5]进行比较实验,得出ROC曲线和

Kappa 统计系数, 实验结果分别如图 5, 6 和表 1—4 所示。可以看出, 在相同图像尺寸下, 经自编码器重构的 PRNU 噪声的性能均要优于原始 PRNU 噪声和经过频谱均衡算法处理的 PRNU 噪声, AUC 值均提升大概 1%, 而且经过频谱均衡算法处理的 PRNU 噪声性能略有下降。从 Kappa 系数表中可以得知在 3 种图像尺寸下, 经过自编码器重构的 PRNU 噪声在原始 PRNU 噪声的基础上精度全部有了提升, 效果最好的是在 64 像素 \times 64 像素图像分辨率下, 使用小波去噪算法提取的 PRNU 噪声的基础上 Kappa 系数提高了 0.007 7。

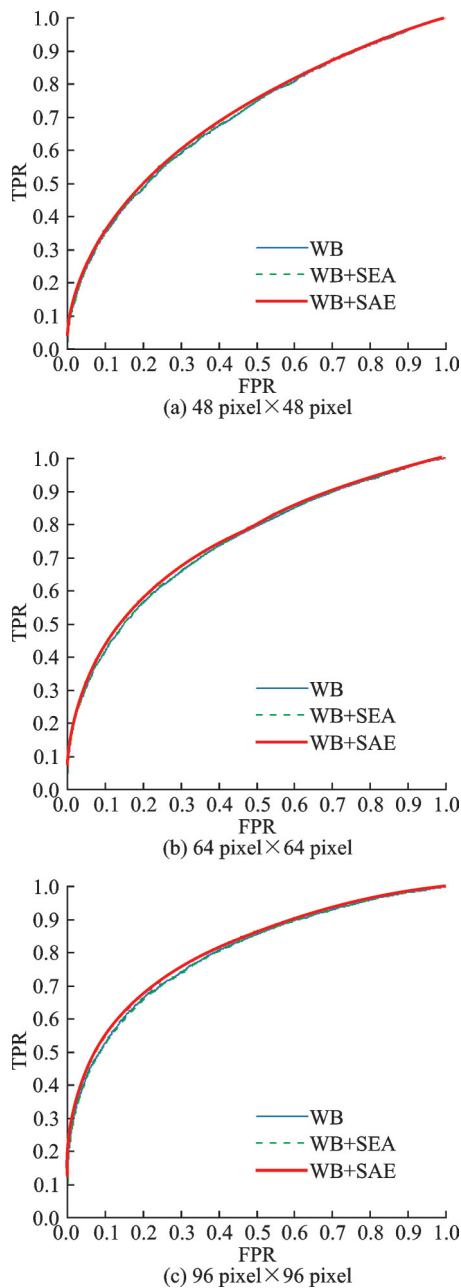


图 5 不同分辨率下基于小波变换去噪算法比较实验的 ROC 曲线图

Fig.5 ROC curves of WB denoising algorithm comparative experiments under different resolutions

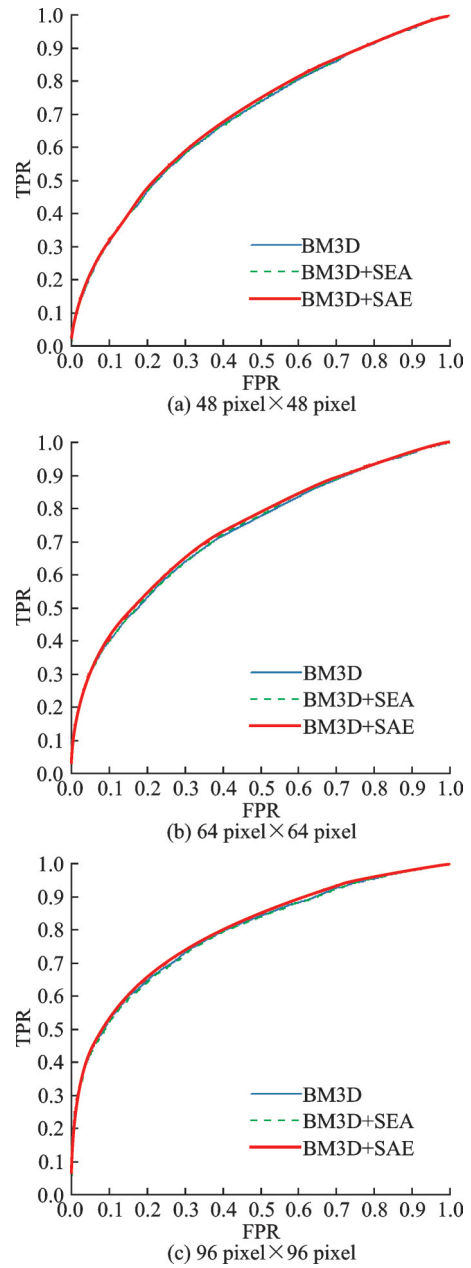


图 6 不同分辨率下基于 BM3D 去噪算法比较实验的 ROC 曲线图

Fig.6 ROC curves of BM3D denoising algorithm comparative experiments under different resolutions

表 1 基于小波变换去噪算法比较实验的 AUC 值

Table 1 AUC of WB denoising algorithm comparative experiments

算法	图像分辨率/(像素 \times 像素)		
	48 \times 48	64 \times 64	96 \times 96
WB	0.696 8	0.743 0	0.800 8
WB+SEA	0.696 8	0.742 7	0.797 8
WB+SAE	0.704 2	0.752 2	0.809 5

本文提出的 PRNU 噪声提纯方法对基于小波变换去噪算法提取的 PRNU 噪声和基于 BM3D 去噪算法提取的 PRNU 噪声都有一定的提纯作用。虽然本文实验是在 48 \times 48、64 \times 64 和 96 \times 96 这 3

表2 基于BM3D去噪算法比较实验的AUC值

Table 2 AUC of BM3D denoising algorithm comparative experiments

算法	图像分辨率/(像素×像素)		
	48×48	64×64	96×96
BM3D	0.686 7	0.730 0	0.793 6
BM3D+SEA	0.685 9	0.728 5	0.789 6
BM3D+SAE	0.693 4	0.738 4	0.799 2

表3 基于小波变换去噪算法比较实验的Kappa系数

Table 3 Kappa coefficients of WB denoising algorithm comparative experiments

算法	图像分辨率/(像素×像素)		
	48×48	64×64	96×96
WB	0.107 5	0.157 8	0.259 7
WB+SEA	0.107 5	0.157 0	0.257 8
WB+SAE	0.109 6	0.165 5	0.266 6

表4 基于BM3D去噪算法比较实验的Kappa系数

Table 4 Kappa coefficients of BM3D denoising algorithm comparative experiments

算法	图像分辨率/(像素×像素)		
	48×48	64×64	96×96
BM3D	0.090 3	0.155 9	0.260 5
BM3D+SEA	0.090 6	0.151 8	0.253 7
BM3D+SAE	0.091 3	0.156 2	0.260 8

种小尺寸的图片上进行,并且使用了自编码器模型对原始的PRNU噪声进行了有损压缩,但是重构后的PRNU噪声的性能不但没有下降,反而有了提升。因此,可证明PRNU噪声在经过自编码器模型的压缩-解压缩过程时进一步去除了无关噪声,实现了对PRNU噪声的提纯。另外,与频谱均衡提纯方法的比较实验也验证了本文方法的先进性。

5 结 论

为了提升图像PRNU噪声在图像来源鉴别方面的性能,本文提出了一种基于样本错配训练的图像PRNU噪声提纯方法。实验结果表明在相同图像尺寸的条件下,该方法对使用经典去噪算法(如基于小波变换的去噪算法、基于BM3D的去噪算法)提取的PRNU噪声具有一定的提纯作用,同时该方法也可以用于其他去噪算法提取的PRNU图像噪声性能的提升。由于受到实验所使用显卡显存大小的限制,无法在更高分辨率上进行实验与比

较。未来将进一步研究轻量级的深度神经网络算法,使其满足对更高分辨率的图像PRNU噪声的提纯。

参考文献:

- [1] LUKAS J, FRIDRICH J, GOLJAN M. Digital camera identification from sensor pattern noise [J]. IEEE Transactions on Information Forensics and Security, 2006, 1(2): 205-214.
- [2] MIHCAK M K, KOZINTSEV I, RAMCHANDRAN K. Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising [C]//Proceedings of 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. [S.l.]: IEEE, 1999: 3253-3256.
- [3] DABOV K, FOI A, KATKOVNIK V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering [J]. IEEE Transactions on Image Processing, 2007, 16(8): 2080-2095.
- [4] 张永胜, 田华伟, 肖延辉, 等. 基于三方加权稀疏编码模型的PRNU提取算法 [J]. 应用科学学报, 2019, 37(5): 704-710.
ZHANG Yongsheng, TIAN Huawei, XIAO Yanhui, et al. PRNU extraction algorithm based on trilateral weighted sparse coding model [J]. Journal of Applied Sciences, 2019, 37(5): 704-710.
- [5] CHEN M, FRIDRICH J, GOLJAN M, et al. Determining image origin and integrity using sensor noise [J]. IEEE Transactions on Information Forensics and Security, 2008, 3(1): 74-90.
- [6] LIN X, LI C T. Preprocessing reference sensor pattern noise via spectrum equalization [J]. IEEE Transactions on Information Forensics and Security, 2015, 11(1): 126-140.
- [7] HU Y, YU B, JIAN C. Source camera identification using large components of sensor pattern noise [C]//Proceedings of 2009 2nd International Conference on Computer Science and its Applications. [S.l.]: IEEE, 2009: 1-5.
- [8] LI R, GUAN Y, LI C T. PCA-based denoising of sensor pattern noise for source camera identification [C]//Proceedings of 2014 IEEE China Summit & International Conference on Signal and Information Processing. [S.l.]: IEEE, 2014: 436-440.
- [9] 张赛, 芮挺, 任桐炜, 等. 基于监督学习深度自编码器

- 的图像重构[J]. 计算机科学, 2018, 45(11): 267-271, 297.
- ZHANG Sai, RUI Ting, REN Tongwei, et al. Image reconstruction based on supervised learning deep auto-encoder [J]. Computer Science, 2018, 45(11): 267-271, 297.
- [10] BENGIO Y, LAMBLIN P, POPOVICI D, et al. Greedy layer-wise training of deep networks[C]//Proceedings of Advances in Neural Information Processing Systems. Vancouver, British Columbia, Canada: DBLP, 2007: 153-160.
- [11] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [12] VINCENT P, LAROCHELLE H, LAJOIE I, et al. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion[J]. Journal of Machine Learning Research, 2010, 11(12): 3371-3408.
- [13] VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoders[C]//Proceedings of the 25th International Conference on Machine Learning. New York: ACM, 2008: 1096-1103.
- [14] 郭浩龙,张荣,郭立君,等.数字照片图像来源检测中的 SPN 预处理方法[J].光子学报, 2017, 46(6): 135-144.
- GUO Haolong, ZHANG Rong, GUO Lijun, et al. Preprocessing method of sensor pattern noise in camera source detection of photo images[J]. Acta Photonica Sinica, 2017, 46(6): 135-144.
- [15] GOLJAN M, FRIDRICH J, FILLER T. Large scale test of sensor fingerprint camera identification[C]//Proceedings of SPIE Conference on Electronic Imaging, Security and Forensics of Multimedia Contents XI. Bellingham: SPIE, 2009: 1-12.
- [16] KLAMBAUER G, UNTERTHINER T, MAYR A, et al. Self-normalizing neural networks[C]//Proceedings of Advances in Neural Information Processing Systems. New York: ACM, 2017: 971-980.
- [17] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by backpropagation error [J]. Nature, 1986, 323: 533-536.
- [18] LUO L, XIONG Y, LIU Y, et al. Adaptive gradient methods with dynamic bound of learning rate [EB/OL]. (2019-02-26) [2020-05-27]. <https://arxiv.org/abs/1902.09843>.
- [19] KINGMA D P, BA J. Adam: A method for stochastic optimization. [EB/OL]. (2017-01-30) [2020-05-27]. <https://arxiv.org/abs/1412.6980>.

(编辑:陈珺)