

基于有损压缩编码的降噪自编码器

袁振, 刘进锋

引用本文

袁振, 刘进锋. 基于有损压缩编码的降噪自编码器[J]. 计算机科学, 2024, 51(6A): 230400172-7.

YUAN Zhen, LIU Jinfeng. Denoising Autoencoders Based on Lossy Compress Coding[J]. Computer Science, 2024, 51(6A): 230400172-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[结合图卷积神经网络和集成方法的推荐系统恶意攻击检测](#)

Malicious Attack Detection in Recommendation Systems Combining Graph Convolutional Neural Networks and Ensemble Methods

计算机科学, 2024, 51(6A): 230700003-9. <https://doi.org/10.11896/jsjcx.230700003>

[一种基于谱图SIFT的同源频谱监测数据判定方法](#)

Method for Homologous Spectrum Monitoring Data Identification Based on Spectrum SIFT

计算机科学, 2024, 51(6A): 230300177-7. <https://doi.org/10.11896/jsjcx.230300177>

[基于图自编码器和GRU网络的分层交通流预测模型](#)

Hierarchical Traffic Flow Prediction Model Based on Graph Autoencoder and GRU Network

计算机科学, 2024, 51(6A): 230400148-6. <https://doi.org/10.11896/jsjcx.230400148>

[融合三维人脸动态信息和光流信息的人脸表情识别](#)

Facial Expression Recognition Integrating 3D Facial Dynamic Information and Optical Flow Information

计算机科学, 2024, 51(6A): 230700210-7. <https://doi.org/10.11896/jsjcx.230700210>

[基于残差密集卷积自编码的高噪声图像去噪方法](#)

Residual Dense Convolutional Autoencoder for High Noise Image Denoising

计算机科学, 2024, 51(6A): 230400073-7. <https://doi.org/10.11896/jsjcx.230400073>

基于有损压缩编码的降噪自编码器

袁 振 刘进锋

宁夏大学信息工程学院 银川 750021

(yuanzhen1999@stu.nxu.edu.cn)

摘 要 图像预处理算法的优劣程度直接关系到图像后置处理的效果,如图像分割、目标检测、边缘提取等。为了获取高质量的数字图像,对图像进行降噪处理成了必不可少的前置步骤。图像降噪旨在尽可能地保持原始信息完整性(即主要特征)的同时,又能够去除信号中无用的信息。为此,提出了一种基于有损压缩编码的卷积自编码器(AutoEncoders,AE)去噪模型;并根据最大编码率下降原则(the principle of Maximal Coding Rate Reduction,MCR²)设计了新的损失函数代替主流深度学习算法中常用的均方误差(Mean Squared Error,MSE)损失,以提高模型的鲁棒性和适应性。模型首先通过编码器处理带噪图像,得到隐变量,然后使用解码器进行解码,消除噪声并得到重构图像。接下来,保持编码器不变,将重构图像输入编码器,使编码器继续学习并得到重构隐变量。最后,通过计算隐变量与重构隐变量的距离来间接衡量重构图像与原始图像的误差,并将其作为收敛代价进行模型训练。在 thumbnails128×128 和 CBSD68 数据集上对所提模型进行了大量实验验证。实验结果表明,该自编码器框架(AE-MCR²)在不同类型的噪声(高斯噪声、伯努利噪声和泊松噪声)下均表现出良好的性能,并具有一定的可解释性。

关键词: 计算机视觉;图像去噪;自编码器;卷积神经网络;压缩编码

中图分类号 TP391

Denoising Autoencoders Based on Lossy Compress Coding

YUAN Zhen and LIU Jinfeng

Department of Information Engineering, Ningxia University, Yinchuan 750021, China

Abstract The performance of image preprocessing algorithms is directly related to the effect of image post-processing, such as image segmentation, target detection, edge extraction, etc. In order to obtain high-quality digital images, image noise reduction has become an essential pre-step. Image noise reduction aims to maintain the integrity of the original information (i. e., the main features) as much as possible, while being able to remove the useless information in the signal. To this end, this paper proposes a lossy compression coding based convolutional auto-encoders (AutoEncoders, AE) denoising model. According to the principle of maximal coding rate reduction (MCR²), a new loss function is designed to replace the mean squared error (MSE) loss commonly used in mainstream deep learning algorithms to improve the robustness and adaptability of the model. The model first processes the noisy image through an encoder to obtain the hidden variables, and then decodes it using a decoder to remove the noise and obtain the reconstructed image. Next, keeping the encoder unchanged, the reconstructed image is fed into the encoder so that the encoder continues to learn and obtains the reconstructed hidden variables. Finally, the error between the reconstructed image and the original image is indirectly measured by calculating the distance between the hidden variable and the reconstructed hidden variable, which is used as the convergence cost for model training. The proposed model is validated extensively on thumbnails128x128 and CBSD68 datasets, and the experimental results show that the self-encoder framework (AE-MCR²) exhibits good performance under different types of noise (Gaussian, Bernoulli, and Poisson) and has some interpretability.

Keywords Computer vision, Image denoising, Autoencoder, Convolutional neural network, Compression coding

1 概述

近年来,深度学习以其强大的建模能力引领了各个领域的革新热潮。在图像处理方面,深度学习广泛应用于图像表示、图像分割、目标检测与分类、动作识别等领域^[1]。然而,在一些图像中,像素值主要会因为以下两点原因发生变化:(1)光子或其他信号的有限计数统计、在芯片内的电子偏移中引入的损耗;(2)放大器或电缆中出现了电子噪声。像素值的变化

通常会引入无用的信息形式,扰乱图像的可观测信息,称为图像噪声。在图像信号处理成像过程的各个流程中,均有可能生成噪声。按照物理成因,可分为沟道热噪声、散粒噪声、放大器噪声和量化噪声等;按照其视觉效果和满足的分布特性,可分为椒盐噪声、泊松噪声、高斯白噪声和脉冲噪声等。

噪声的存在会破坏图像信息,引起颜色或亮度值的随机变化。此外,噪声还可能导致不合理的结果,如不真实的边缘、被忽略的线条、角落、伪影和模糊的物体,从而妨碍后续图

基金项目:宁夏自然科学基金(2023AAC03126)

This work was supported by the Ningxia Natural Science Foundation(2023AAC03126).

通信作者:刘进锋(jfliu@nxu.edu.cn)

像相关操作的进行。因此,去除噪声成为增强和检索图像中有价值隐藏细节的基本要求^[2]。在现代计算机视觉领域,图像去噪技术被广泛用于实现这一目的^[3-4]。

图像去噪技术旨在通过消除噪声,从噪声图像中生成干净图像,在保留图像基本细节的同时抑制噪声^[5]。按照处理原理差异,主要分为基于信号处理原理的图像去噪技术(传统图像降噪方法)和基于深度学习的图像去噪技术。传统图像降噪方法主要包括稀疏编码建模、基于块的方法和低秩结构假设等。本文采用了自编码器与有损压缩编码相结合的深度学习方法来去除退化图像中的噪声,并通过最大编码率下降原则设计了新的损失函数,以衡量重构误差,从而训练模型。

自编码器是一种人工神经网络,旨在学习数据的表示(编码)。在本文中,自编码器从训练图像中学习噪声特征,并尝试生成与原始输入非常接近的干净图像。有损压缩编码通过将信号压缩为少量的线性组合测量,然后使用特定算法在计算机中进行解压缩,以恢复原始信号^[6-9]。最大编码率下降原则是一种基于有损压缩编码的表示学习原则,旨在使用较少的位数编码数据,同时尽可能多地保留输入数据中包含的信息。所提出的 AE-MCR² 自编码器模型主要采用卷积层、池化层和反卷积层来消除噪声。最后,使用峰值信噪比(Peak Signal to Noise Ratio, PSNR)定量衡量模型的性能,并对各个模型的去噪效果进行定性对比。

2 相关工作

图像去噪方法可以分为传统方法和基于深度学习的方法,其中自编码器去噪是基于深度学习的主要分支之一。目前已有许多的基于自编码器的深度学习架构用以图像的去噪任务,并成功推广到工程领域。Zhang 等提出的 DnCNN 是一种深度卷积神经网络架构,通过残差学习来去除各种强度和类型的图像噪声,具有比传统高斯去噪器更好的性能^[10]。然而,该方法需要训练大量的深度卷积神经网络模型,计算资源和时间开销较大,并且对特定类型的噪声可能表现不佳。接着,他们提出了 FFDNet,这是一种快速而灵活的基于卷积神经网络模型的图像去噪算法,能够高效地处理各种图像噪声,并在不同的噪声水平下展现出良好的去噪效果,但在处理过多噪声的图像时可能存在过拟合问题^[11]。随后, Tian 等提出了一种双重去噪网络 DudeNet,通过全局和局部两个子网络提取全局和局部特征,并融合这些特征来恢复复杂噪声图像的细节^[12]。然而, DudeNet 的双重卷积神经网络模型需要训练两个不同的神经网络模型,训练时间和复杂度较高,参数数量较多。最近, Bajaj 等提出了基于自编码器的深度学习模型(Deep Learner AutoEncoders, DL-AE),其利用自编码器来对有噪声的图像进行编码和解码,从而去除图像中的噪声,实验证明了其有效性和提高图像质量的能力^[2]。此外,随着 transformer 架构的兴起,一些研究者也提出了与之相关的图像去噪模型,如 SwinIR^[13] 等。

这些工作在图像去噪问题上提供了有效的解决方案和方法,具有创新性和实用价值。它们都采用了基于深度学习的方法,通过构建适当的网络结构和损失函数,实现了对图像噪声的准确去除。这些主流方法大多使用均方误差来计算原始图像与重构图像之间的误差。均方误差是一种像素点级别的

运算方式,不能很好地兼顾数据的空间结构。这使得部分方法的鲁棒性和适应性有待进一步提高。

如何衡量原始图像与重建图像的差异一直是深度学习降噪自编码器的设计重点。本文根据有损压缩编码领域的相关研究,设计了新的损失函数,以间接计算重构误差。Ma 等提出了一种基于有损数据编码和压缩的多元混合数据分割方法,并将其用于处理混合数据的分割问题。该方法将数据转换为编码后的位流,并利用信息熵和互信息进行分割,适用于包含不同数据类型的混合数据^[14]。在此方法基础上, Yu 等提出了一种新的表示学习原则——最大编码率下降(MCR²),旨在学习多样化和有辨识度的数据表示。MCR² 通过优化互信息下界和最小化表示的编码长度来鼓励学习表示使用较少的位数编码输入数据,并尽可能多地编码输入数据的信息^[15-16]。MCR² 提出了一种计算输入数据编码长度的方法,为衡量原始图像与重建图像的差异提供了新的方向与思路。

综上,本文模型根据 MCR² 设计了新的损失函数来代替传统的均方损失函数,通过自编码器将原始图像与重建图像映射到同一特征空间内,然后通过计算隐变量与重构隐变量的距离来间接衡量重构图像与原始图像的误差,将误差作为收敛代价进行模型训练。

3 模型架构

3.1 降噪自编码器

在图像去噪领域内,使用深度学习自编码器去噪可以取得不错的效果。自编码器旨在同时学习从 X 到 Z 的编码映射 f 和从 Z 到 \hat{X} 的解码(逆)映射 $g(z, \eta)$ ^[17]:

$$X \xrightarrow{f(x, \theta)} Z \xrightarrow{g(z, \eta)} \hat{X} \quad (1)$$

在本文中,为了方便表示,使用 $X = [x^1, \dots, x^n] \in \mathbb{R}^{D \times n}$ 表示一个有限样本 x 的矩阵, $Z = [z^1, \dots, z^n] \in \mathbb{R}^{d \times n}$ 表示映射的特征矩阵。自编码器的目标是使得解码样本 \hat{X} 与输入样本 X 尽可能地相似。因此,如何衡量 \hat{X} 与 X 之间的差异是自编码器设计的重点。由于使用 KL 散度(Kullback-Leibler Divergence)等方法衡量 \hat{X} 与 X 之间的距离是十分困难的,因为 \hat{X} 与 X 的分布较为复杂,通常无法预估^[18-19],因此,在实践中,不少研究通常选择最小化某些近似边界或者代理解决上述问题^[20]。

本文通过设计标准自编码器新的学习路径,使用最大编码率降低原则设计了新的损失函数,以间接衡量 \hat{X} 与 X 之间的距离以替代传统的均方误差损失函数,在计算重构误差的同时兼顾了数据之间的空间结构。

3.2 最大编码率降低原则

Yu 等最近提出了一种深度学习的新目标,旨在学习多类数据的线性判别性表示(Linear Discriminant Representation, LDR)。其基本思想是将真实数据的分布(可能是在多个非线性子流形上 $\bigcup_{j=1}^k M_j \subset \mathbb{R}^D$, 见图 1 左)映射到一系列标准模型上,这些模型由多个独立(或正交)线性子空间组成,表示为 $\bigcup_{j=1}^k S_j \subset \mathbb{R}^d, d \ll D$ (见图 1 右)。具体而言,考虑一组来自 k 个不同类

的 n 个数据样本 $X = [x_1, \dots, x_n] \in R^{D \times n}$ 。MCR² 假设 $X = \bigcup_{j=1}^k X_j$, 其中每个子集 X_j 都分布在一个低维子流形上, 即 $X_j \subset M_j, j=1, \dots, k$ 。矩阵 $\Pi^i(i, i)=1$ 表示样本 i 属于类 j ($\Pi^i=0$ 表示相反)。MCR² 寻求一个连续的映射 $f(\cdot, \theta): x \rightarrow z$, 即 X 到一个最佳表征 $Z = [z^1, \dots, z^n] \subset R^{d \times n}$:

$$X \xrightarrow{f(x, \theta)} Z \quad (2)$$

最大编码率降低原则具体表示为^[15]:

$$\max_z \Delta R(Z | \Pi, \epsilon) \doteq \frac{1}{2} \log \det(I + \alpha Z Z^T) - \underbrace{R(Z | \epsilon)}_{\sum_{j=1}^k \frac{\gamma_j}{2} \log \det(I + \alpha_j Z \Pi^j Z^T)} \quad (3)$$

其中, $\alpha = \frac{d}{n\epsilon^2}, \alpha_j = \frac{d}{\text{tr}(\Pi^j)\epsilon^2}, \gamma_j = \frac{\text{tr}(\Pi^j)}{n}, j=1, \dots, k$ 。

在本文中, 为了简化研究, 模型假定 Π 和 ϵ 是已知并且固定的。第一项简写为 $R(Z)$, 表示整个特征集 Z 的编码率 (以高斯源编码), 相应的精度为 ϵ ; 第二项简写为 $R_c(Z)$, 表示 k 个特征子集 $Z_j = f(X_j)$ (每个以高斯编码) 的平均编码率。

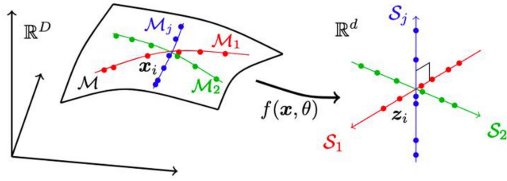


图 1 MCR² 的学习过程^[15]

Fig. 1 Learning process of MCR²

3.3 AE-MCR² 模型

正如 3.1 节所言, 在深度学习自编码器的体系结构中, 直接衡量原始数据与重构数据之间的差异是十分困难的。针对此问题, 本文提出了以下学习过程, 如图 2 所示。

模型的具体学习过程表达为式(4)^[20]:

$$X \xrightarrow{f(x, \theta)} Z \xrightarrow{g(z, \eta)} \hat{X} \xrightarrow{f(x, \theta)} \hat{Z} \quad (4)$$

其中, 模型的整体参数为 $\Theta = \{\theta, \eta\}$, $X \rightarrow Z \rightarrow \hat{X}$ 为标准自编码器的工作流程, 即解码器从样本 X 学习到特征表示 Z , 随后解码器将特征表示 Z 解码, 最后生成 \hat{X} 。同时, 假设编码器

$f(x, \theta)$ 不变, 使编码器继续学习解码得到的样本 \hat{X} , 随后生成特征表示 \hat{Z} 。最后通过计算 Z 与 \hat{Z} 之间的“距离”, 从而间接衡量 \hat{X} 与 X 之间的差异。理想状态下, 对于训练数据 $x \in X$ 的所有特征表示 $z = f(x) \in Z$, 与 \hat{X} 编码后的样本 $\hat{z} = f(g(z, \eta), \theta) \in \hat{Z}$ 接近。

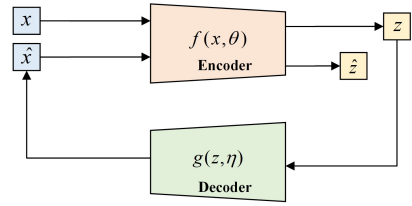


图 2 AE-MCR² 模型学习过程

Fig. 2 Learning process of AE-MCR²

对于所提出的学习过程, 由于特征空间 z 是通过设计线性子空间或(退化)高斯来表示的, 因此对特征空间 z 中的样本和分布可以做出有几何或统计意义的度量。由于每个类的特征 Z_j 和 \hat{Z}_j 类似于子空间/高斯分布, 因此它们的“距离”可以通过码率降低来衡量。在式(3)的限制下, Z_j 和 \hat{Z}_j 为两个相同大小的集合, 二者的距离可表达为:

$$\Delta R(Z_j, \hat{Z}_j) \doteq R(Z_j \cup \hat{Z}_j) - \frac{1}{2} (R(Z_j) + R(\hat{Z}_j)) \quad (5)$$

综上所述, 本文所提出学习过程的目标函数可表达为:

$$d(Z, \hat{Z}) \doteq \min_{\eta} \sum_{j=1}^k \Delta R(Z_j, \hat{Z}_j) = \min_{\theta, \eta} \sum_{j=1}^k \Delta R(Z_j, f(g(\hat{Z}_j, \eta), \theta)) \quad (6)$$

由于自编码器是无监督学习的范式之一, 其监督信号为解码器解码生成的图片, 所以 $k=2$, 表示输入图片以及解码图片两个类别(分别为 0 和 1)。

AE-MCR² 模型的总体架构如图 3 所示, 模型分为编码器(Encoder)和解码器(Decoder)两个部分。编码器主要由 3 个卷积层与 3 个最大池化层组成, 并使用 ReLU 函数作为激活函数。解码器部分主要由 4 个反卷积层组成, 前 3 层使用 ReLU 函数作为激活函数, 最后一层使用 Sigmoid 函数作为激活函数。值得一提的是, AE-MCR² 模型的整体参数量约为 30000 个。

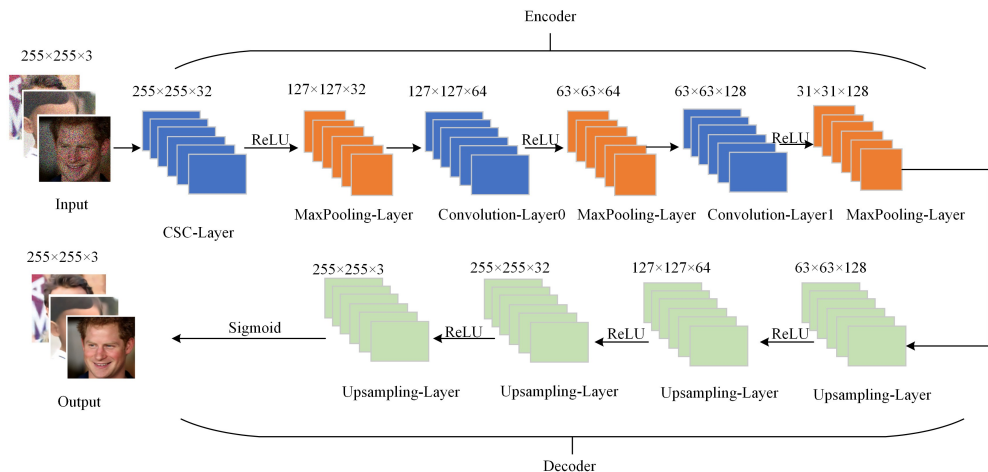


图 3 AE-MCR² 模型架构

Fig. 3 Model architecture of AE-MCR²

4 实验结果与分析

4.1 实验数据集

本文使用的 thumbnails 128×128 数据集由 70 000 张分辨率为 128×128 的高质量 PNG 图像组成,并且在年龄、种族、和背景方面包含巨大的差异。数据集还对眼镜、太阳镜、帽子等细节配件有很好的覆盖。图 4 显示了 thumbnails 128×128 数据集中的一组示例图像。在本文所涉及的实验中,训练集与测试集的比例为 7:3。



图 4 Thumbnails 128×128 数据集图像示例

Fig. 4 Example image of thumbnails 128×128 data

为了测试模型的泛化性能,本文在 CBSD68 数据集上进行了相关实验。CBSD68 数据集被广泛用于衡量图像去噪算法的性能。它包括 68 张彩色 jpg 图片,数据集涵盖的图像类别较为丰富,包括动物、人物、建筑以及风景等。每张图片的分辨率为 321×481,原始数据集包含了不同程度的高斯噪声图片。为了更好地评估本文模型,本文就伯努利噪声以及泊松噪声扩充了数据集。另外,在模型训练过程中,CBSD68 数据集只作为测试集,训练集中未包含任何 CBSD68 数据集中的相关图片。图 5 展示了 CBSD68 数据集相关示例图像。



图 5 CBSD68 数据集图像示例

Fig. 5 Example of images in CBSD68 dataset

4.2 实验设置

实验均在 Inter Core i7 11th CPU 16RAM 以及 Nvidia GeForce RTX3070 8GB GPU 上实现。采用深度学习的开源框架 Pytorch。实验均未采用任何数据增强方法。

在模型训练过程中,设置 batch_size = 16,学习率为 0.001,使用 Adam 算法^[21]作为优化器,epoch 设置为 50,ε² 设置为 0.5,d=128 为编码器的输出维度。本文使用 PNSR 衡量模型性能。PSNR 计算公式如下:

$$PSNR = 10 \times \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (7)$$

其中,MSE 为原始图像与重建图像之间的均方误差。

4.3 实验结果与分析

将所提框架与深度学习降噪领域的先进方法进行定量及定性评估。对模型在各种噪声影响下的性能进行广泛评估,包括高斯噪声、伯努利噪声及泊松噪声,3 种噪声影响下的噪声图如图 6 所示。表 1—表 3 展示了各个模型在 3 种不同类型噪声影响下的性能。由于本文使用了自编码器作为基础架构,因此选择模型架构相似的 DL-AE 模型以及较为复杂但去噪效果较优的 FFDNet 网络作为对比模型进行了大量实验。

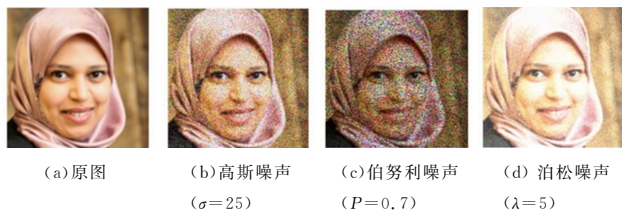


图 6 各种噪声影响下的图片

Fig. 6 Pictures under the influence of various noises

表 1 列出了高斯噪声影响下实验模型在 thumbnails 128×128 数据集上的 PSNR 均值;图 7 显示了模型在高斯噪声(σ=15)影响下训练阶段的 Loss 值以及测试集的 PSNR 均值;图 8 显示了各个模型在高斯噪声影响下,在 thumbnails 128×128 数据集上的去噪效果;图 9 显示了实验模型在高斯噪声影响下,在 CBSD68 数据集上的去噪效果。

表 1 高斯噪声影响下实验模型在 thumbnails 128×128 数据集上的 PSNR 均值

Table 1 Mean PSNR of experimental model under the influence of Gaussian noise on thumbnails 128×128 dataset

噪声水平	σ=15	σ=25	σ=50
DnCNN ^[10]	31.50	28.34	27.18
FFDNet ^[11]	31.77	29.43	26.81
DudeNet ^[12]	31.81	28.26	27.14
DL-AE ^[2]	30.12	28.38	26.72
AE-MCR ² (Ours)	32.86	30.98	27.50

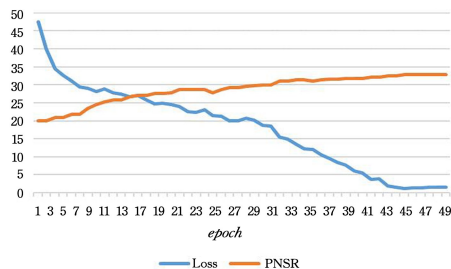


图 7 高斯噪声(σ=15)时模型的 Loss 值和测试集 PNSR 均值

Fig. 7 Loss values and test set PSNR means of the model with Gaussian noise(σ=15)

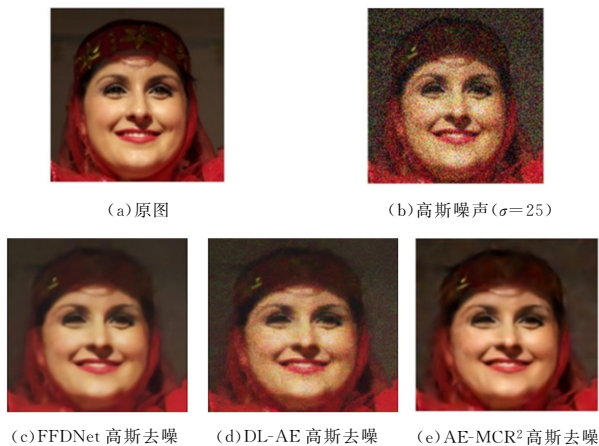


图 8 高斯噪声影响下模型在 thumbnails 128×128 数据集上的去噪效果

Fig. 8 Denoising effect of the model under the influence of Gaussian noise on thumbnails 128×128 dataset

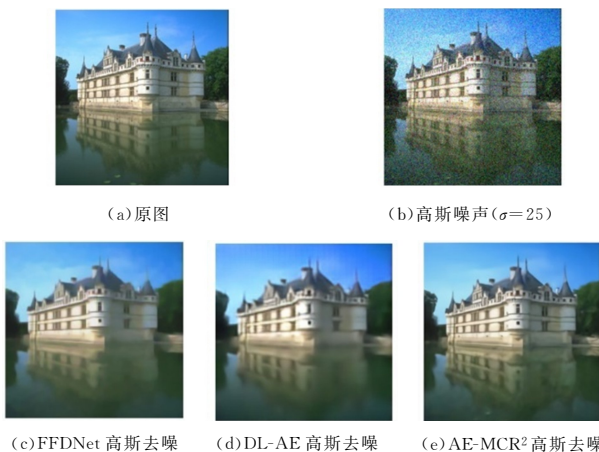


图 9 高斯噪声影响下模型在 CBSD68 数据集上的去噪效果

Fig. 9 Denoising effect of the model under the influence of Gaussian noise on CBSD68 dataset

实验表明,本文模型在各个程度的高斯噪声影响下都取得了较为优异的表现。具体而言,当 $\sigma=15$ 时,本文模型相较于表现最好的 FFDNet 模型,PSNR 均值提高了约 1 dB,相较于 DL-AE 模型,PSNR 均值提高了约 2.7 dB;当 $\sigma=25$ 时,FFDNet 模型表现依然有竞争力,本文模型相较于 FFDNet 模型,PSNR 均值提高了约 1.5 dB,相较于 DL-AE 模型,PSNR 均值提高了约 2.6 dB;当 $\sigma=50$ 时,各个模型的 PSNR 均值均下降明显,本文模型相较于最好的 DnCNN 模型,PSNR 均值提高了约 0.3 dB,相较于 DL-AE 模型,PSNR 均值提高了约 0.8 dB。如图 8 所示,本文模型可以有效地去除图片中的高斯噪声点,并保留较为丰富的图像细节。DL-AE 模型所生成的去噪图片还含有部分的高斯噪声点,FFDNet 模型则在去除噪声得同时损失了部分图像细节。模型在 CBSD68 数据集上的泛化效果如图 9 所示,AE-MCR²模型的去噪效果更为明显且保留了图像的细节,FFDNet 模型的去噪图片与 DL-AE 的去噪图片则有略微的图像细节损失。

表 2 显示了伯努利噪声影响下实验模型在 thumbnails 128×128 数据集上的 PSNR 均值;图 10 显示了本文模型在伯努利噪声($P=0.7$)影响下训练阶段的 Loss 值以及测试集的 PSNR 均值;图 11 显示了各个模型在伯努利噪声影响下,

在 thumbnails 128×128 数据集上的去噪效果;图 12 显示了实验模型在伯努利噪声影响下,在 CBSD68 数据集上的去噪效果。

表 2 伯努利噪声影响下实验模型在 thumbnails 128×128 数据集上的 PSNR 均值

Table 2 Mean PSNR values of the experimental model under the influence of Bernoulli noise on thumbnails 128×128 dataset (dB)

噪声水平 (值越高,噪声越弱)	$P=0.7$	$P=0.8$	$P=0.9$
DnCNN ^[10]	32.49	37.52	37.29
FFDNet ^[11]	33.57	38.09	38.16
DudeNet ^[12]	34.15	38.59	37.65
DL-AE ^[2]	32.90	36.86	37.23
AE-MCR ² (Ours)	36.94	39.68	39.74

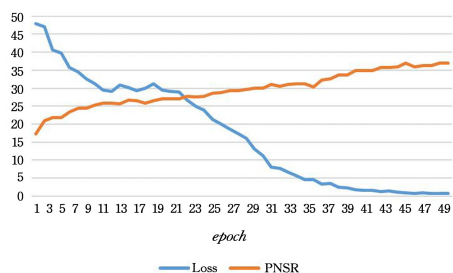


图 10 伯努利噪声($P=0.7$)时模型的训练 Loss 值和测试集 PSNR 均值

Fig. 10 Training Loss values and test set PSNR means of the model with Bernoulli noise($P=0.7$)

通过定量考量,本文模型在伯努利噪声的影响下也取得了较优的表现。当 $P=0.7$ 时,图片受噪声干扰较大,相较于其他模型的 PSNR 均值,本文模型平均有 3~4 dB 的提升;当 $P=0.8$ 时,各个模型均有不错的表现,但本文模型在 PSNR 均值上仍有 1~2 dB 的提升;当 $P=0.9$ 时,图片受到的噪声干扰最弱,本文模型相较于其他模型有 2 dB 左右的 PSNR 均值提高。同时,如图 11 所示,当 $P=0.7$ 时,本文模型几乎可以不损失任何细节以去除伯努利噪声。



图 11 伯努利噪声影响下模型在 thumbnails 128×128 数据集上的去噪效果

Fig. 11 Denoising effect of the model under the influence of Bernoulli noise on thumbnails 128×128 dataset

相较于本文模型,DL-AE 模型也可较好地保留图像细节,但仍存在部分噪声点未去除,FFDNet 模型在去除噪声的

同时有部分图像细节损失。伯努利去噪模型在 CBSD68 数据集上的泛化效果如图 12 所示。

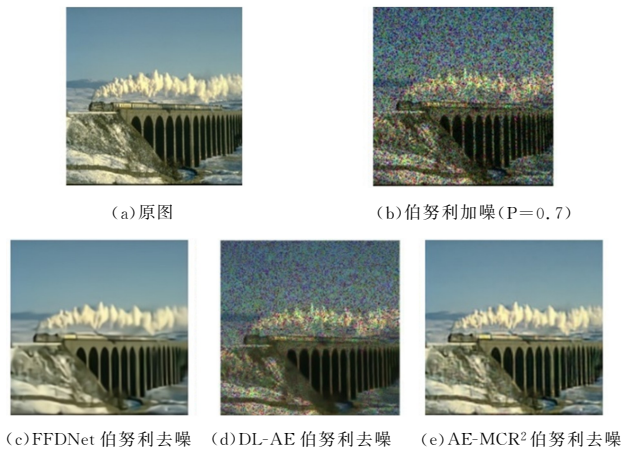


图 12 伯努利噪声影响下模型在 CBSD68 数据集上的去噪效果
Fig. 12 Denoising effect of the model under the influence of Bernoulli noise on CBSD68 dataset

DL-AE 模型的泛化效果较差,有大量的噪声点未去除,FFDNet 模型泛化效果较好,但有图像细节损失,AE-MCR²模型则可去除大部分噪声点,且图像细节损失较小。

表 3 列出了泊松噪声影响下实验模型在 thumbnails 128×128 数据集上的 PSNR 均值;图 13 显示了模型在泊松噪声($\lambda=5$)影响下训练阶段的 Loss 值以及测试集的 PSNR 均值;图 14 显示了各个模型在泊松噪声影响下,在 thumbnails 128×128 数据集上的去噪效果;图 15 显示了实验模型在泊松噪声影响下,在 CBSD68 数据集上的去噪效果。

表 3 泊松噪声下实验模型在 thumbnails 128×128 数据集上的 PSNR 均值

Table 3 Mean PSNR values of the experimental model with Poisson noise on thumbnails 128×128 dataset (dB)

噪声水平 (值越高,噪声越强)	$\lambda=5$	$\lambda=6$	$\lambda=7$
DnCNN ^[10]	31.34	28.29	24.10
FFDNet ^[11]	31.61	28.17	25.80
DudeNet ^[12]	32.06	29.06	26.63
DL-AE ^[2]	29.61	28.38	24.19
AE-MCR ² (Ours)	34.97	31.44	27.85

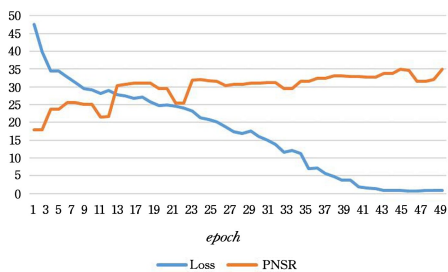


图 13 泊松噪声($\lambda=5$)时模型的训练 Loss 值和测试集 PSNR 均值
Fig. 13 Training Loss values and test set PSNR means of the model with Poisson noise($\lambda=5$)

除了高斯噪声以及伯努利噪声,本文还在泊松噪声的影响下进行了大量实验。当 $\lambda=5$ 时,相较于其他模型,本文模型的 PSNR 均值提升较多,约为 3~5 dB;当 $\lambda=6$ 时,本文模型比表现最好的 DudeNet 模型在 PSNR 均值上提升了

1.5 dB,比较差的 DL-AE 模型提升了约 3 dB;当 $\lambda=7$ 时,图片受到噪声干扰最大,本文模型比最优的 DudeNet 模型提升了约 1 dB 的 PSNR 均值,比 DL-AE 模型提升了约 3.5 dB。同时,如图 14 所示,当 $\lambda=5$ 时,本文模型可以去绝大部分噪声点,并且很好地复原了图像。相较于本文模型,DL-AE 模型对泊松噪声缺乏一定的鲁棒性,不能很好地去除图像中的噪声点,FFDNet 模型可以去除大部分噪声点。泊松去噪模型在 CBSD68 模型上的泛化效果如图 15 所示,AE-MCR²模型可以去除噪声点,但会损失一定的图像细节,DL-AE 模型则不能很好地对图像进行去噪处理,FFDNet 模型可以去除噪声点,但同样会损失一定的图像细节。

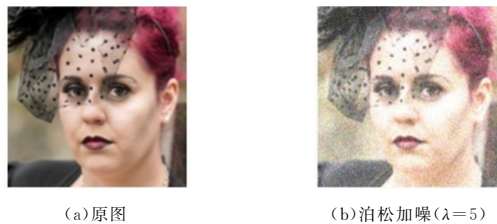


图 14 泊松噪声影响下模型在 thumbnails 128×128 数据集上的去噪效果
Fig. 14 Denoising effect of the model under the influence of Poisson noise on thumbnails 128×128 dataset

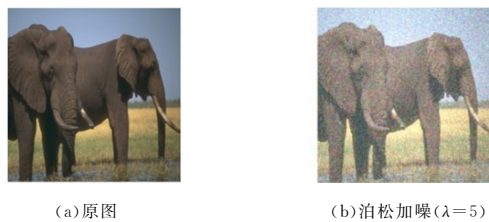


图 15 泊松噪声影响下模型在 CBSD68 数据集上的去噪效果
Fig. 15 Denoising effect of the model under the influence of Poisson noise on CBSD68 dataset

本文的消融实验如表 4 所列,其中 AE 模型以均方损失作为损失函数,其模型架构与 AE-MCR²模型完全相同,实验分别在高斯噪声($\sigma=25$)、伯努利噪声($P=0.7$)以及泊松噪声($\lambda=5$)影响下的 thumbnails 128×128 数据集上进行。

表 4 各消融实验的 PSNR 均值

Table 4 Mean PSNR value of each ablation experiment

噪声类型及水平	高斯 $\sigma=25$	伯努利 $P=0.7$	泊松 $\lambda=5$
AE-MSE	26.09	31.09	30.40
AE-MCR ² (Ours)	30.98	36.94	34.97

消融实验表明,AE-MCR²模型相比于 AE-MSE 模型,在不同类型的噪声下(高斯、伯努利和泊松噪声)都展现出更好的去噪性能,表现为更高的平均 PSNR 值。这些结果进一步验证了 AE-MCR²模型在去噪任务中的有效性和优越性。

结束语 本文提出了一种基于有损压缩编码的卷积自编码器 AE-MCR²,并根据最大编码率下降原则设计了新的损失函数,同时提出了降噪自编码器新的学习过程,即在传统自编码器的学习过程中添加了重构图像到重构隐变量的学习路径。由于传统的降噪自编码器方法中衡量原始数据与重构数据之间的距离较为困难,因此本文通过最大编码率下降原则计算隐变量与重构隐变量之间的距离以间接衡量原始数据与重构数据的差异,从而训练模型进一步去除图像中的噪声。本文在 thumbnails 128×128 数据集与 CBSD68 数据集上进行了大量实验,涉及了高斯噪声、伯努利噪声以及泊松噪声 3 种噪声类型。实验表明,AE-MCR²模型对多种类型的噪声均有优异的鲁棒性,并且在去除噪声的同时可以保留较好的图像细节,损失较少的图像信息。

参 考 文 献

- [1] DWIVEDI N, SINGH D K. Review of deep learning techniques for gender classification in images[C]// Harmony Search and Nature Inspired Optimization Algorithms: Theory and Applications (ICHSA 2018). Springer Singapore, 2019: 1089-1099.
- [2] BAJAJ K, SINGH D K, ANSARI M A. Autoencoders based deep learner for image denoising[J]. Procedia Computer Science, 2020, 171: 1535-1541.
- [3] LEHTINEN J, MUNKBERG J, HASSELGREN J, et al. Noise2Noise: Learning image restoration without clean data[J]. arXiv:1803. 04189, 2018.
- [4] ELAD M, KAWAR B, VAKSMAN G. Image denoising: The deep learning revolution and beyond—a survey paper[J]. SIAM Journal on Imaging Sciences, 2023, 16(3): 1594-1654.
- [5] XIE J, XU L, CHEN E. Image denoising and inpainting with deep neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25.
- [6] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [7] CANDÈS E J, WAKIN M B. An introduction to compressive sampling[J]. IEEE Signal Processing Magazine, 2008, 25(2): 21-30.
- [8] BARANIUK R G. Compressive sensing [lecture notes][J]. IEEE Signal Processing Magazine, 2007, 24(4): 118-121.
- [9] DUARTE M F, DAVENPORT M A, TAKHAR D, et al. Single-pixel imaging via compressive sampling[J]. IEEE Signal Processing Magazine, 2008, 25(2): 83-91.
- [10] ZHANG K, ZUO W, CHEN Y, 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.
- [11] ZHANG K, ZUO W, ZHANG L. FFDNet: Toward a fast and flexible solution for CNN-based image denoising[J]. IEEE Transactions on Image Processing, 2018, 27(9): 4608-4622.
- [12] TIAN C, XU Y, ZUO W, et al. Designing and training of a dual CNN for image denoising[J]. Knowledge-Based Systems, 2021, 226: 106949.
- [13] LIANG J, CAO J, SUN G, et al. Swinir: Image restoration using swin transformer[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 1833-1844.
- [14] MA Y, DERKSEN H, HONG W, et al. Segmentation of multivariate mixed data via lossy data coding and compression[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(9): 1546-1562.
- [15] YU Y D, KWAN H R C, YOU C, et al. Learning diverse and discriminative representations via the principle of maximal coding rate reduction[C]// Advances in Neural Information Processing Systems. 2020.
- [16] CHAN K H R, YU Y, YOU C, et al. ReduNet: A white-box deep network from the principle of maximizing rate reduction[J]. The Journal of Machine Learning Research, 2022, 23(1): 4907-5009.
- [17] HINTON G E, ZEMEL R. Autoencoders, minimum description length and Helmholtz free energy[J]. Advances in Neural Information Processing Systems, 1993, 6.
- [18] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C]// International Conference on Machine Learning. PMLR, 2017: 214-223.
- [19] ZHAO S, SONG J, ERMON S. Infovae: Information maximizing variational autoencoders[J]. arXiv:1706. 02262, 2017.
- [20] DAI X, TONG S, LI M, et al. Closed-loop data transcription to an ldr via minimaxing rate reduction[J]. arXiv: 2111. 06636, 2021.
- [21] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. arXiv:1412. 6980, 2014.



YUAN Zhen, born in 1999, master. His main research interests include image classification and computer vision.



LIU Jinfeng, born in 1971, Ph.D, professor, master supervisor. His main research interests include image processing and heterogeneous computing.