

采用卷积自编码器网络的图像增强算法

王万良, 杨小涵, 赵燕伟, 高楠, 吕闯, 张兆娟

(浙江工业大学 计算机科学与技术学院, 机械工程学院, 浙江 杭州 310023)

摘 要: 将图像增强方法低光网络 (LLNet) 应用于实际场景下的彩色图像时会产生大量冗余参数, 为此基于 LLNet 提出卷积自编码器网络 (CAENet) 的图像增强方法. 将 LLNet 方法中的低光处理模块与网络训练衔接在一起; 采用卷积网络代替传统自编码器的编码和解码方式. 实验结果表明: CAENet 能够有效节约时间成本, 减少网络参数, 使网络训练更加高效, 得到更好的图像低维表示. 在 Corel5k 数据集上的实验结果表明, CAENet 在减少网络参数的同时, 能有效提高图像光感和色感; 在高分辨率数据集上的实验结果表明, 针对图像细节方面, CAENet 能够保留细节不失真; 针对含噪低光图像, CAENet 能在增强图像的同时达到去噪的效果, 证明 CAENet 具有较强的鲁棒性.

关键词: 图像处理; 图像增强; 深度学习; 卷积神经网络; 降噪自编码器

中图分类号: TP 311.1 **文献标志码:** A **文章编号:** 1008-973X(2019)09-1728-13

Image enhancement algorithm with convolutional auto-encoder network

WANG Wan-liang, YANG Xiao-han, ZHAO Yan-wei, GAO Nan,
LV Chuang, ZHANG Zhao-juan

(School of Computer Science and Technology, College of Mechanical Engineering,
Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: When the image enhancement method LLNet (the low-light net) was applied to three-channel images, there're a lot of redundant parameters. To solve this problem, a framework called CAENet (convolutional auto-encoder network) was proposed. Firstly, CAENet combined a low light processing module with a network training module. Secondly, in the encoding and decoding stages, CAENet used a convolutional network to replace the traditional fully connected network. The experimental results show that connecting low-light processing modules with network training can effectively save time costs. At the same time, the use of convolutional networks can reduce network parameters, making network training more efficient, and obtain better low-dimensional representation of images. The experimental results on the Corel5k dataset show that CAENet can effectively improve the image light perception and color perception while reducing network parameters. The experimental results on high-resolution datasets show that CAENet can preserve details for image details without distortion. In addition, for the noisy low-light image, CAENet can enhance the image while achieving the denoising effect, which proves that CAENet has strong robustness.

Key words: image processing; image enhancement; deep learning; convolutional neural network; denoising auto-encoder

收稿日期: 2019-01-24. 网址: www.zjujournals.com/eng/article/2019/1008-973X/201909012.shtml

基金项目: 国家自然科学基金资助项目 (61873240).

作者简介: 王万良 (1957—), 男, 教授, 从事人工智能、机器自动化、网络控制研究. orcid.org/0000-0002-1552-5075.

E-mail: zjutwll@zjut.edu.cn

在现代生活中,摄影设备技术已经取得很大进步,拍摄出来的照片能够很大程度地还原真实场景,但是在环境较为恶劣的情况下,比如低光环境,就会拍摄出低质量的低光图像^[1],这给图像识别、分类等后续的图像处理工作增加了难度.对于工业生产、医学研究等对图像质量要求较高的领域^[2],低光图像可能导致决策者作出错误的决定.图像增强算法作为一种能够提高图像整体光感质量的算法,对于图像分析具有至关重要的作用.

图像增强算法能够提高图像对比度,使增强后的图片更加清晰,而对比度增强算法在处理图像时可能会出现图像过度增强或不自然的效果^[3].目前常用的图像增强算法有直方图均衡(Histogram equalization, HE)算法、基于Retinex的图像增强算法等.HE算法通常用来增加图像全局对比度,使图像亮度很好地分布在直方图上.HE算法的优点是原理简单,缺点是增强之后的图像会出现伪像和强饱和度^[4]等不好的效果.Reza^[5]于2003年提出了CLAHE算法,该算法很好地抑制了图像的噪声.Retinex理论由Land于1971年提出,该理论认为人眼对于色彩的感知取决于物体表面的反射特性,而与入射强度无关^[6].1997年基于Retinex的单尺度Retinex(single scale retinex, SSR)^[7]图像增强算法首次被提出,该算法类似于在自然视觉科学中广泛使用的高斯差分(difference of gaussian, DOG)函数,将反射率视为最终增强结果.带颜色恢复的多尺度Retinex(multi-scale retinex with color restoration, MSRRCR)^[8]图像增强算法在边缘增强和色彩均衡^[9]方面有很大进步.Parthasarathy等^[10]于2012年提出了一种具有颜色恢复功能的自动化多尺度Retinex算法(automated multi-scale retinex with color restore, AMSRCR),该算法从增强的图像中获取参数值,为MSRRCR提出了一种完全自动化的方法.Petro等^[11]于2014年提出了具有色度保存的多尺度Retinex算法(the multi-scale retinex with chromaticity preservation, MSRCP),该算法在图像颜色恢复和对比度增强方面具有较好的效果.Fu等^[12]于2016年提出了图像增强算法SRIE(simultaneous reflectance and illumination estimation),该算法能够更好地保留图像反射率,还能在一定程度上抑制噪声.Guo等^[1]于2017年提出了低光图像增强(low-light image enhancement, LIME)算法,该算法通过找到三通道中的最大值来估计每个像素的光照强度,可以相应地实现图像增强.

近年来,深度学习作为机器学习领域的一个重要研究方向,在图像处理、自然语言处理、语音识别等领域取得了突破性进展^[13].Shen等^[14]提出了一种基于卷积神经网络和Retinex理论的低光图像增强模型,证明了多尺度Retinex图像增强算法相当于前馈卷积神经网络,有效提高了图像亮度.Li等^[15]提出了基于卷积神经网络的图像增强框架,首先将低光图像作为卷积网络的输入,接着将得到的输出结果用于Retinex算法进行低光图像增强.

上述基于深度学习的图像增强算法具有较好的增强效果,但都是基于有监督的学习方式,这种学习方式增加了寻找有标签低光图像数据集的难度,因此将无监督^[16]学习方式应用在图像增强领域很有必要.Lore等^[17]提出了一种基于自动编码器的图像增强框架——低光网络(the low-light net, LLNet),该框架利用自编码器的去噪能力,将自编码器应用在图像增强领域,在训练时将低光图像输入到网络中进行特征学习,最后得到重构明亮图像,达到增强图像的效果.

虽然LLNet在无监督学习方式下具有不错的图像增强效果,但仍是在传统的单通道灰度图像上进行讨论.为了增加框架的普适性,本文选择三通道的彩色图像来进行框架的设计和训练.由于三通道彩色图像不同于单通道灰度图像,每个像素点会有三个数值来表示,在网络训练时若使用传统的编码和解码方式会产生大量冗余参数,同时LLNet网络的训练集预处理工作独立于网络存在,这大大增加了网络训练的前期准备工作.因此基于以上问题,在LLNet的基础上,提出一种新的基于卷积自编码器的图像增强框架——CAENet(convolutional auto-encoder network).

为了不失一般性,本文首先对LLNet图像增强方法进行相关介绍和分析;其次介绍CAENet,详细描述CAENet在图像低光处理阶段和特征提取阶段作出的改进,重点阐述在特征提取阶段利用卷积操作取代传统的编码方式,说明CAENet整体方法的具体实现过程;最后对所提算法进行实验验证.

1 低光网络

自编码器(Autoencoder)是一种从样本数据中自动学习的无监督算法,网络利用非线性数据映

射得到的低维特征来反映得到高维的非线性映射空间,因此自编码器网络能够提取非线性特征^[18].去噪自编码器^[19]的原理与自编码器相同,都是尽量使重构输出数据等于输入数据,两者的不同之处在于,去噪自编码器的输入是被人为破坏的噪声数据.原始干净数据为 \mathbf{x} ,进行加噪处理后,输入网络得到的特征图^[19]表示为

$$\mathbf{h}(\tilde{\mathbf{x}}) = f(\mathbf{W} \times \tilde{\mathbf{x}} + \mathbf{b}). \quad (1)$$

式中: $\tilde{\mathbf{x}}$ 为将输入 \mathbf{x} 进行加噪处理后的噪声数据, \mathbf{W} 为网络权重, \mathbf{b} 为偏置.该阶段为编码阶段.

在得到噪声数据的特征图之后,网络已经学习到输入数据的隐含特征,接着网络会用已经学习到的特征来重构出原始干净的数据.最终网络得到的重构无噪数据如下式所示:

$$\begin{aligned} \mathbf{y}(\tilde{\mathbf{x}}) &= g[\mathbf{W}' \times \mathbf{h}(\tilde{\mathbf{x}}) + \mathbf{b}'] = \\ &g[\mathbf{W}' \times f(\mathbf{W} \times \tilde{\mathbf{x}} + \mathbf{b}) + \mathbf{b}']. \end{aligned} \quad (2)$$

式中: \mathbf{W}' 为权重, \mathbf{b}' 为偏置.该阶段为解码阶段.

基于降噪自编码器原理,LLNet 将自编码器应用在图像增强领域中. LLNet 的结构示意图如图 1 所示,输入低光图像,经过多层隐藏单元,最终得到重构的增强图像.

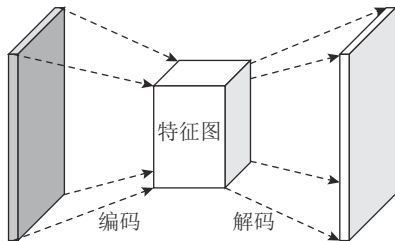


图 1 低光网络 (LLNet) 结构图

Fig.1 Structure diagram of low light net (LLNet)

得到重构增强图像之后,网络开始进行参数微调.首先进行原始干净图像与重构增强图像之间损失函数的计算.损失函数计算公式^[17]如下:

$$\begin{aligned} L_{\text{DA}}(D; \theta) &= \frac{1}{N} \sum_{i=1}^N \|y_i - \hat{y}(\mathbf{x}_i)\|_2^2 + \beta \sum_{j=1}^K \text{KL}(\hat{\rho}_j \| \rho) + \\ &\frac{\lambda}{2} [\|\mathbf{W}\|_F^2 + \|\mathbf{W}'\|_F^2]. \end{aligned} \quad (3)$$

式中: $\theta = \{\mathbf{W}, \mathbf{b}, \mathbf{W}', \mathbf{b}'\}$ 为模型参数, D 为低光图像, N 为样本数量, $\text{KL}(\hat{\rho}_j \| \rho)$ 为 KL 散度, γ, β 为通过交叉验证确定的标量超参数.

当损失函数计算完成后网络会采用向后传播算法,以最小化损失函数的原则进行网络微调. LLNet 的模型结构图如图 2 所示.

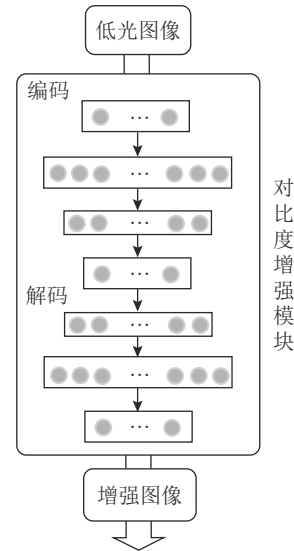


图 2 LLNet 模型结构示意图

Fig.2 Diagram of module structure of LLNet

尽管 LLNet 方法具有很好的图像增强效果,该网络会产生大量冗余参数,同时数据前期处理的时间成本较大.为了克服以上缺点,本文提出基于卷积自编码器的图像增强方法——CAENet.

2 采用卷积自编码器的图像增强网络

基于 LLNet 方法,提出基于卷积自编码器的图像增强框架——CAENet. 首先,为了减少前期数据处理时间,CAENet 将 LLNet 的人为低光处理模块添加到整体网络框架中,实现了框架统一,避免了多平台处理存在的诸多不便因素;其次,为了将该方法更好地应用到三通道彩色图像中,受卷积自编码器^[20]启发,将卷积操作当作自编码器的编码操作来得到低光图像的低维特征表示,结构图如图 3 所示.网络训练时输入处理后的低光图像,随后经过卷积网络进行编码,得到特征图,此时网络学习到低光图像的隐含特征,接着进行反卷积得到重构明亮图像,最后计算损失函数以便网络调参.

2.1 低光处理

训练网络之前,首先需要对输入图像进行破坏,即低光处理.与去噪自编码原理相同,该操作主要是为了避免隐藏层学习到没有意义的恒等函数,其次也是为了网络能够学习到更加具有鲁棒性的特征表达^[21].

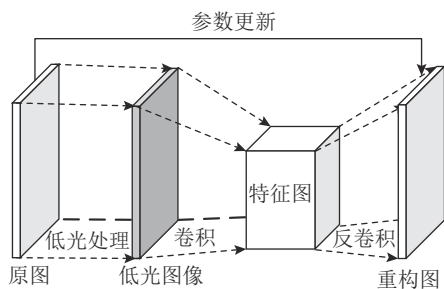


图3 采用卷积自编码器的图像增强网络 (CAENet) 结构示意图

Fig.3 Schematic diagram of convolutional auto-encoder network (CAENet)

如图4所示为在流形学习方法^[22]下的图像增强过程. 其中叉表示训练数据, 圆圈代表等概率损失过程, 箭头演示了将原始训练数据 x 转化为低光数据 x_d 的过程. 在进行图像增强时, 网络将编码得到的特征来进行反映射, 即圆圈圈内箭头所代表的过程: 将低光图像 x_d 映射成为重构图像 x .

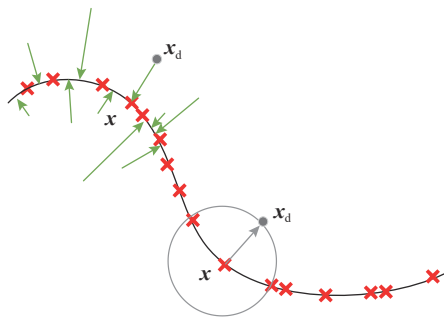


图4 流形学习方法下的图像增强原理图

Fig.4 Image enhancement schematic diagram under manifold learning method

如图5所示, 在进行图像低光化处理时, 由于LLNet方法需要人为操作, 并且处理时用到不同的平台, 这无形中增加了模型预处理时间^[21]. CAENet将该阶段与网络训练阶段衔接起来, 网络就可以将原始明亮的图像数据集用作输入数据来进行训练, 减少了图像的人为前期准备工作, 节省了时间成本, 同时统一了图像处理平台, 为网络训练不同数据集提供了灵活的转化方式.

2.2 卷积自编码器

在传统的自编码器网络层中, 神经元的连接方式通常是全连接方式, 如图6所示. 图中, X 为输入数据, Y 为重构数据, h 为网络层中的神经元. 从图中可以看出, 在处理高维数据时, 这种神经元连接方式会使得网络产生大量冗余参数, 尤其对于三通道彩色图像, 参数会变得更加多. 又由于

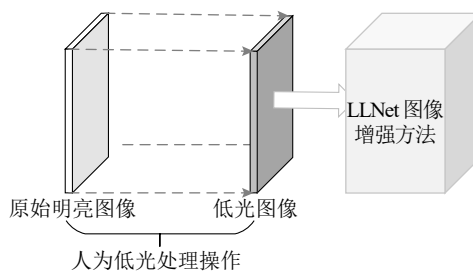


图5 LLNet 低光处理阶段示意图

Fig.5 Diagram of low light processing stage of LLNet

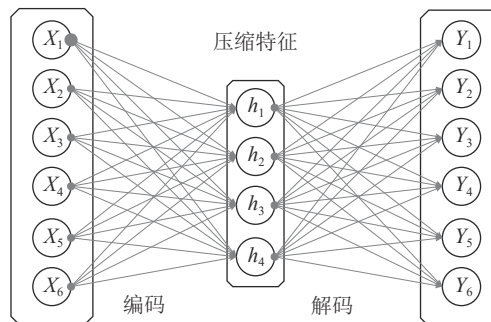


图6 传统自编码器神经元全连接方式示意图

Fig.6 Diagram of full connection of traditional autoencoder

在传统自编码器网络中, 每个网络层参数是全局的, 这样就无法保留空间局部性^[20], 使得网络学习速率变得很慢.

卷积网络具有局部连接和权值共享的特点, 这2个特点能够很好地解决以上问题.

1) 局部连接, 即局部感知. 在卷积网络中, 神经元只对图像的局部信息进行感知, 进行感知的图像局部信息即为感受野. 当采用局部连接时, 卷积层的每个神经元仅与图像局部相连, 这大大减少了网络参数值. 如果要得到图像更多的特征信息, 则需要更多的卷积核来进行卷积, 最后只需在最高层将得到的特征信息综合起来即可. 图7显示了三通道彩色图像的卷积过程. 图像具有三通道表示图像的像素点由红、绿、蓝三色表示, 对应图7的左边部分, 而图7右边部分表示共有 n 个卷积核. 在卷积过程中, 图像感受野的各个通道先与1个卷积核进行卷积, 得到3个卷积结果后再相加, 得到该区域的第1个特征图. 接着该图像与第2个卷积核进行卷积相加得到第2个特征图, 直到与第 n 个卷积核卷积得到 n 个特征图.

在网络训练过程中, 网络内部是一个不透明的黑匣子, 而且由于不同的训练对象具有差异性, 在实际情况中并没有具体的理论支撑来说明卷积核的大小和个数应该如何设置. 在卷积网络中, 小尺寸的卷积核会比大尺寸的卷积核生成更

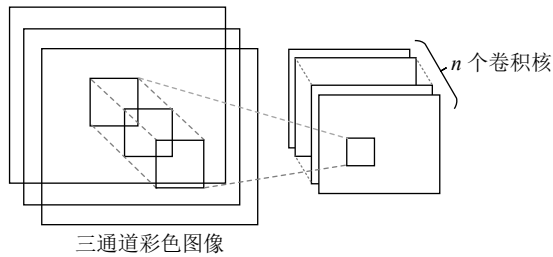


图 7 三通道彩色图像的卷积过程示意图

Fig. 7 Diagram of convolutional process of three-channel image

少的参数个数, 这点在处理三通道图像时显得尤为重要. 这是由于三通道图像的一个像素点由 3 个数值来表示, 当使用较大卷积核进行卷积操作时会产生大量的参数, 使得网络训练变得耗时. 并且, 使用多个小的卷积核的感受野和使用少量大的卷积核的感受野相同. 基于此, 本文在设定卷积核大小时根据卷积核越小越好的原则.

a. 卷积核大小为 1×1 . 大小为 1×1 的卷积核一般用来数据降维或者升维, 但是在 CAENet 中卷积核要大于 1 才能对感受野进行卷积编码得到特征, 因此 CAENet 不采用大小为 1×1 的卷积核;

b. 卷积核大小为偶数中的最小数 2×2 . 由于偶数大小的卷积核无法保证输入的特征图和输出特征图大小尺寸不变, 因此 CAENet 不采用大小为 2×2 的卷积核;

c. 卷积核大小为奇数中除了 1 以外的最小数 3×3 . 奇数大小的卷积核能够保证输入特征图和输出特征图大小尺寸不变, 同时奇数卷积核有中心像素, 能够确定卷积核的位置. 因此 CAENet 采用大小为 3×3 的卷积核来进行图像特征提取.

2) 权值共享. 在一幅图像中, 某一区域内的特征可能与其他区域内的特征相似, 比如相邻区域内的特征有极大可能是相同的, 因此在进行网络训练时不需要图像每一处的区域都拥有各自的权值, 可以通过共享权值来进行训练. 该特点能够很好地减少网络参数数量.

综上, 在处理三通道彩色图像这类高维数据时, CAENet 将卷积层作为网络特征提取层.

2.3 采用卷积自编码器的图像增强网络

1) 低光处理阶段.

在将图像处理成为低光图像时, 由于 Gamma 矫正的处理为像素点级别, 不会给图像带来不连续的斑块, 也不用考虑图像光源的位置来调整图像^[23], 因此在图像低光化处理时可以使用 Gamma 矫正. Gamma 矫正的本质是调整图像的 Gamma 曲

线, 是图像预处理阶段经常使用的非线性算子. 在矫正过程中, 首先对图像进行检测, 得到图像中的深色部分和浅色部分, 使两者比例增大或减小, 从而调整图像全局对比度^[24], 得到高光或者低光图像. 图像调整公式^[17]如下:

$$\mathbf{x}_d = A \times \mathbf{x}_1^\gamma. \quad (4)$$

式中: A 为常量, \mathbf{x}_1 为需要调整的原始高光图像. 当 $\gamma < 1$ 时, 可以提高图像对比度, 达到图像增亮的效果; 当 $\gamma > 1$ 时, 可以降低图像对比度, 达到图像低光处理效果; 而当 $\gamma = 1$ 时, 对图像不具有任何效果.

由于自然场景下的图像会有一定的噪声出现, 为了增加框架的鲁棒性, 可以在图像预处理成为低光图像时加入噪声来进行网络的增亮和去噪的训练. 本框架除了加入常见的高斯噪声外, 还加入椒盐噪声、泊松噪声和斑点噪声来增加网络的鲁棒性.

得到低光图像 \mathbf{x}_d 之后, 就可以将 \mathbf{x}_d 输入到网络隐层中进行图像增亮和去噪处理. 在进行网络训练时将卷积、池化等操作加入训练过程, 即编码阶段, 这样可以实现特征不变性提取的无监督学习, 能够很好地提取图片特征, 得到稳定的训练参数. 同时为了得到重构图像, 将反卷积操作加入解码阶段. 这一过程如图 8 所示.

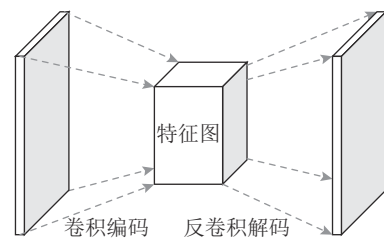


图 8 卷积编码以及反卷积解码示意图

Fig. 8 Diagram of convolutional coding and deconvolutional decoding

2) 卷积编码阶段.

隐层处理的第一阶段为图像编码, 包括卷积操作和池化操作. 首先 \mathbf{x}_d 经过卷积网络^[25] 进行编码, 得到比原图小很多且特征更加明显的特征图^[20]:

$$\mathbf{h}(\mathbf{x}_d) = \sigma(\mathbf{x}_d \mathbf{W}^k + \mathbf{b}^k). \quad (5)$$

式中: \mathbf{W} 为已经初始化成功的卷积核, 共有 k 个, 在 CAENet 中 $k=3$, 每个卷积核搭配一个偏置 \mathbf{b} , σ 为激活函数. 卷积核与输入 \mathbf{x}_d 卷积之后形成 k 个特征图 $\mathbf{h}(\mathbf{x}_d)$.

经过卷积操作之后,接着对图像进行池化操作,最终得到特征图 h' . 池化操作具有 2 个优点:

a. 降低特征图维度. 图像通过卷积操作得到多张特征图之后,假如直接进行后续计算,那么计算量将会变的很大. 这时需要池化操作来减少参数数量以降低维度,进而降低计算量;

b. 空间不变性^[26]. 空间不变性是指池化过程会保留图片位置关系的矩阵和大部分重要信息,这对之后的解码过程有至关重要的作用.

3) 反卷积解码阶段.

池化操作是编码阶段的最后一步,完成后网络开始进行解码过程. 在解码阶段,可以进行反卷积^[27]操作来重构图像. 由于不合理的卷积核大小和步长可能会使反卷积操作产生棋盘效应,即可能会出现矩阵块重叠的问题,在本框架中,使用上采样和卷积的方式来进行图像信息复原.

在上采样操作中采用内插值法来进行图片大小复原,即在原有图像的基础上,在像素点之间采用合适的插值算法插入新的元素. CAENet 方法使用最邻近内插值法,该算法在待插入像素点的相邻 4 个像素点中选取距离最近的像素点,将其像素赋值给待插入像素,得到与原图像大小相等的特征图 y .

上采样操作已将特征图大小复原,因此解码阶段的卷积操作不需要考虑图像大小问题,只须进行简单卷积操作即可. 具体卷积操作公式^[6]如下:

$$Y = \sigma(yW'^k + b'^k). \quad (6)$$

式中: Y 为重构图像.

至此,网络训练过程(见 图 9)完成.

图像维度变换过程如图 10 所示. 输入图像经过共享卷积层可得特定维度的卷积特征图^[28]. 首先将 3 维图像加深为 128 维,直到编码结束,数据维度变为 32 维. 在此过程中,网络利用前层学习

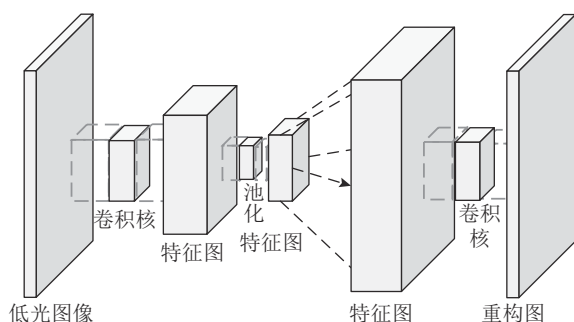


图 9 CAENet 训练过程

Fig.9 Process of CAENet training

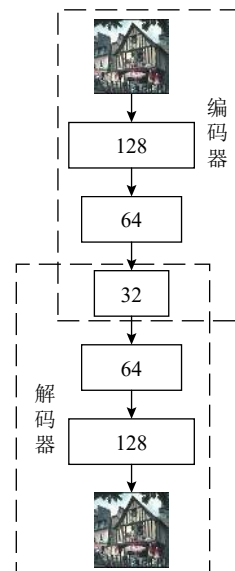


图 10 图像维度变换过程

Fig.10 Process of change in image dimensions

得到的特征图与卷积核进行卷积操作提取局部特征,使网络学习更加细节化. 解码操作最后将图片深度还原为 128 维. 但是由于图片正常维度是 3 维,还需要一个 3 维的卷积核来对图片进行降维处理.

CAENet 整体框架的伪代码如下所示.

算法: CAENet

Input: Original_Image x

Begin:

Dark_image $x_d = A^*x^v$

For ($e \leq \text{epoch}$) **do:**

Feature map $h = \text{convolution } x_d \text{ by Eq.(5);}$

Pooling Feature map h' ;

Upsample Feature map y ;

Reconstructed image $Y = \text{convolution } y \text{ by Eq.(6);}$

Parameter Update;

End;

第一步进行图像预处理,第二步进行迭代计算. 在第二步中,当迭代数小于预设的迭代数时,网络进行编码和解码来训练数据集,并进行参数更新,直到迭代数不满足迭代条件,网络完成训练. 代码执行具体流程图如图 11 所示.

3 实验分析

3.1 实验环境及对比算法

为了验证本文框架的图像增强能力,对比 CAENet 与目前常用的图像增强算法的效果,对比

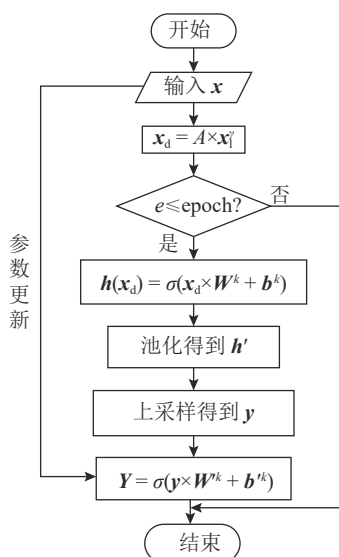


图 11 CAENet 伪代码流程图

Fig. 11 Flow chart of CAENet pseudo code

算法包括标准直方图均衡算法 HE、对比度有限的自适应直方图均衡算法 CLAHE、带颜色恢复的多尺度 Retinex 算法 MSRCR、具有颜色恢复功能的自动化多尺度 Retinex 算法 AMSRCR、具有色度保存的多尺度 Retinex 算法 MSRCP、SRIE 算法和 LIME 算法。

本次实验使用公共数据集 Corel5k 进行网络训练。该数据集由科雷尔 (Corel) 公司整理, 用于图像实验的 5 000 张真实场景图像包括自然风光、动物、人像等类别, 其中训练集共有 3 330 张图像, 测试集共有 320 张, 图像大小为 128×128。为了进一步证明 CAENet 能够更好的用于高清图像, 本次实验还使用 100 张高清图像来进行实验验证。本次实验环境使用曙光 XMachine W580-G20 服务器, 该服务器配置了双路 CPU, 适用于深度学习计算, 具有处理速度快、扩展性强、易于管理等优点。为了更加直观地对比不同算法效果, 本次实验采用结构相似度 (structural similarity index, SSIM)^[29]、峰值信噪比 (peak signal-to-noise ratio, PSNR)^[30] 和自然图像质量评价指标 (natural image quality evaluator, NIQE)^[31] 作为评估指标。

3.2 实验参数设置

在本实验中, CAENet 采用 3 层卷积编码层和 2 层反卷积解码层, 3 层卷积编码的卷积核个数分别为 128、64 和 32 个, 池化步长为 2; 2 层反卷积的卷积核个数分别为 64 和 128 个, 卷积核大小为 3×3。为了重构图像回归到三通道, 网络还需要 3 个 3×3 的卷积核来进行特征图降维。训练集的迭

代次数为 40, 使用 Tensorflow 框架进行网络搭建。

至于其他对比算法, HE 算法使用 Python 下的 skimage 库, 该库中的 histogram 函数可以对图像直方图进行处理; CLAHE 算法使用 MATLAB 中的集成函数 adapthisteq, 执行默认参数 (8×8 图像图块, 对比度增强限制为 0.01, 全范围输出, 均匀直方图分布并且分布参数为 0.4); MSRCR、AMSRRCR 和 MSRCP 算法均使用 json 集成文件来进行算法实现, SRIE 和 LIME 算法均使用 Matlab 来进行测试。

3.3 网络训练过程

CAENet 算法在训练过程中首先提取低光图像的特征, 为了使处理后的图像尽量接近原图, 调整网络参数以达到最好的增强效果。从图 12 可以看出, 随着网络训练迭代次数的增加, 图像逐渐变亮, 图像细节信息 (如图中房子倒影等) 逐渐显现。

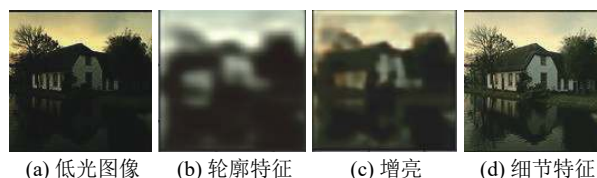


图 12 CAENet 训练中的图片特征逐步提取

Fig. 12 Image feature extraction step by step using CAENet algorithm

3.4 在 Core5k 数据集上的实验对比

如图 13 所示为各个算法在 Corel5k 数据集上的处理效果, 其中图 13(b)~(h) 分别为 HE、CLAHE、MSRCR、AMSRRCR、MSRCP、SRIE 和 LIME 算法的增强效果, 图 13(i) 显示了 CAENet 算法的增强效果。从图中可以看出, 对比算法在增强图片时会存在色彩过于明亮、图像饱和度过高等问题, 从主观层面来说, 得到的图像与真实场景不符。而 CAENet 算法能够很好地增强图像, 整体看来明亮程度和色彩饱和度都最接近真实场景, 比如天空的颜色是令人舒适的蓝色, 金字塔的亮度也没有过高。

尽管 CAENet 算法具有较好的图像颜色增强能力, 由于自编码器在进行编码和解码时会损失一部分数据信息, 而 Corel5k 数据集的图像像素并不高, 损失的这部分数据会对图像质量产生较大的影响, 使 CAENet 的处理效果在主观层面来说稍显模糊。为了证明损失的数据信息在高清数据集上不会产生较大影响, 接下来使用高分辨率数

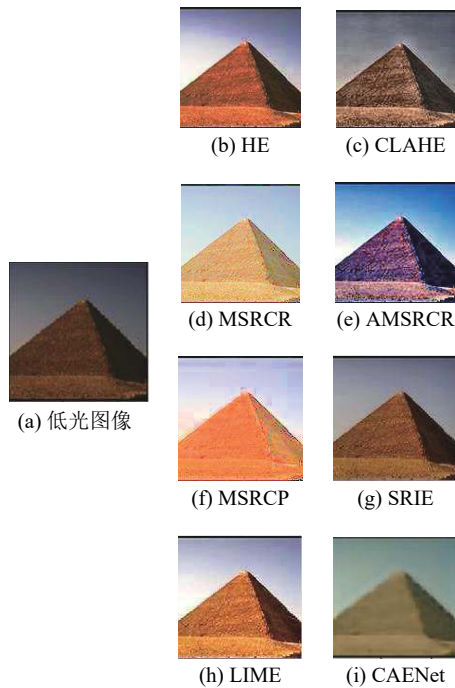


图13 不同算法在Corel5k数据集上的处理效果对比

Fig.13 Effect comparison of different algorithms on Corel5k dataset

据集来进行算法验证。

3.5 高清数据集实验对比

为了验证CAENet适用于高清图像,对比不同图像增强算法在高清数据集上的处理效果。

3.5.1 细节增强对比 如图14所示为在不同黑暗场景下的图像,其中图14(a)、(c)、(e)是原图,图14(b)、(d)、(f)是低光处理后的图像。本次实验主要对比的是细节部分,因此选取的3张图像都具有一些细节特征,如图14中的岩石分层、湖面的倒影以及远处建筑的门窗。

图15~17显示了不同算法对3种不同细节地方的处理效果。从图中可以明显看到,CAENet算法不仅可以调整图像整体亮度,还能增强图像细



图14 细节增强测试实验中的原始图像与低光图像

Fig.14 Original image and low light image in detail enhancement test

节特征,保证图像不失真。对于岩石星空图,从增强过后的图像中能够清晰的看到岩石的断层和天空的颜色变化;对于湖面细节图像,增强之后的树林倒影能被较为清晰地辨别;对于建筑物图像,远处的门窗在增强之后也清晰可见。这些实验结果证明,尽管自编码在进行数据训练时会损失一部分信息,但是对于高分辨率数据集图像来说,损失的数据没有对图像质量产生较大影响,图像增强之后的细节部分能够从主观层面分辨。因此,CAENet算法适用于高清数据集。

为了更加直观地比较CAENet和其他图像增强算法的处理效果,表1将结构相似度(SSIM)作为评价指标,对比了各个算法对3张图像的处理效果。对于岩石图像,由于图像整体颜色较黑,对于图像细节处的增强要求较高,从表中可以看到,除了CAENet,其他算法的SSIM数值相对较低。这说明在图像整体为弱光感时,对比算法对图像的增强效果不是很理想。对于湖面图像,树的颜色以及水中树的倒影是增强的关键,从表中可以看到,各个算法的SSIM数值都较为接近,其中HE算法和本文算法的值都较高,但是HE算法的增强效果从主观层面看失真较严重。对于建筑图像,增强的难点在于远处的建筑呈现低光状态,而近处的房屋内又是较为明亮的颜色,从表中看出,CLAHE算法和CAENet都有较高的SSIM数值,但是前者的视觉呈现效果较差。从表1中还可看出,其他对比算法在处理湖面图像时的SSIM数值较高,但是在处理岩石图像和建筑图像时其SSIM数值都明显下降,而CAENet算法在处理这3种不同场景下图像时都能够保持较为平均的SSIM数值,这表明其他图像增强算法在应用场景上具有一定的局限性,不能完全应用于多个场景,而CAENet算法在一定程度上弥补了这个不足。

表2显示了以上算法处理结果的自然图像质量评价指标(NIQE)的数值对比。该对比指标不同于SSIM,当数值越大时,表示图像质量越低,从表中可以看出,对于3种不同细节增强的图像,处理效果最好的依次是MSRCP、SRIE和CAENet算法。尽管对于岩石和湖面细节图像,CAENet算法没有达到最佳效果,但是与MSRCP、SRIE算法得到的数值十分接近,这体现出CAENet算法在处理不同图像时的效果稳定。

3.5.2 加噪实验对比 为了证明CAENet算法具有鲁棒性,在低光图像中加入噪声进行对比实验,加入的噪声分别为高斯噪声、椒盐噪声、斑点噪

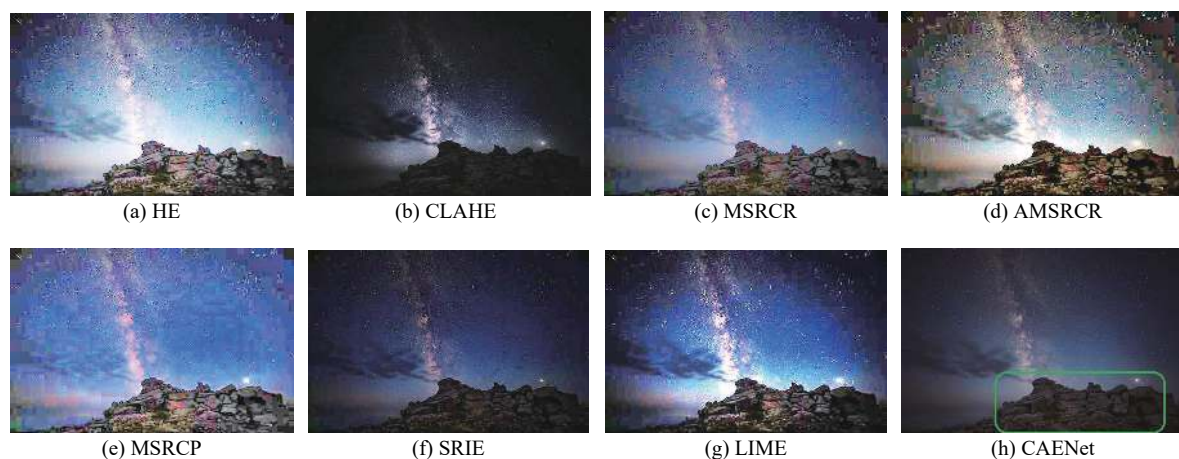


图 15 不同算法对岩石低光图像中岩石细节的增强效果对比

Fig.15 Comparison of enhancement effects of different algorithms on rock detail in rock low light images

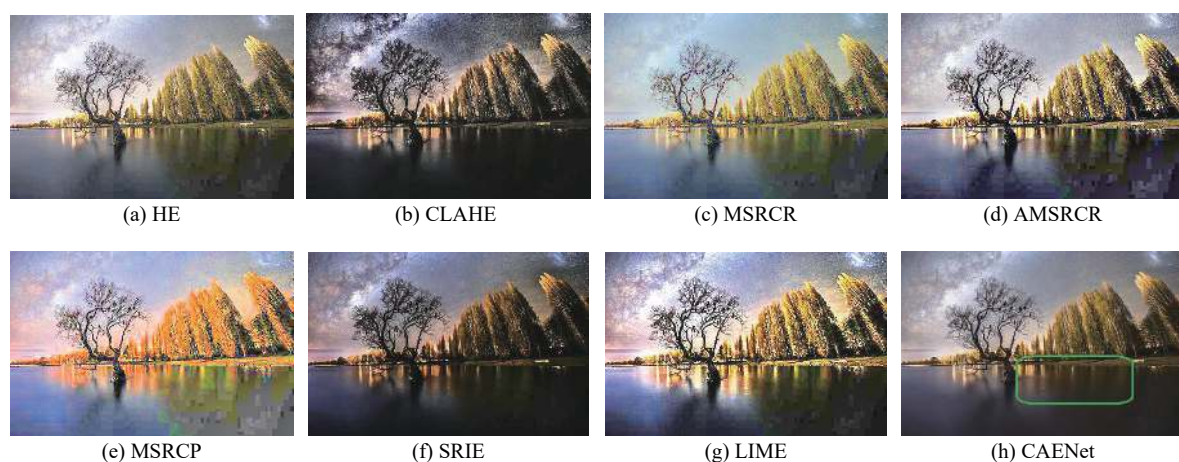


图 16 不同算法对湖面低光图像中水面倒影细节的增强效果对比

Fig.16 Comparison of enhancement effects of different algorithms on reflection of water surface in low-light images of lakes

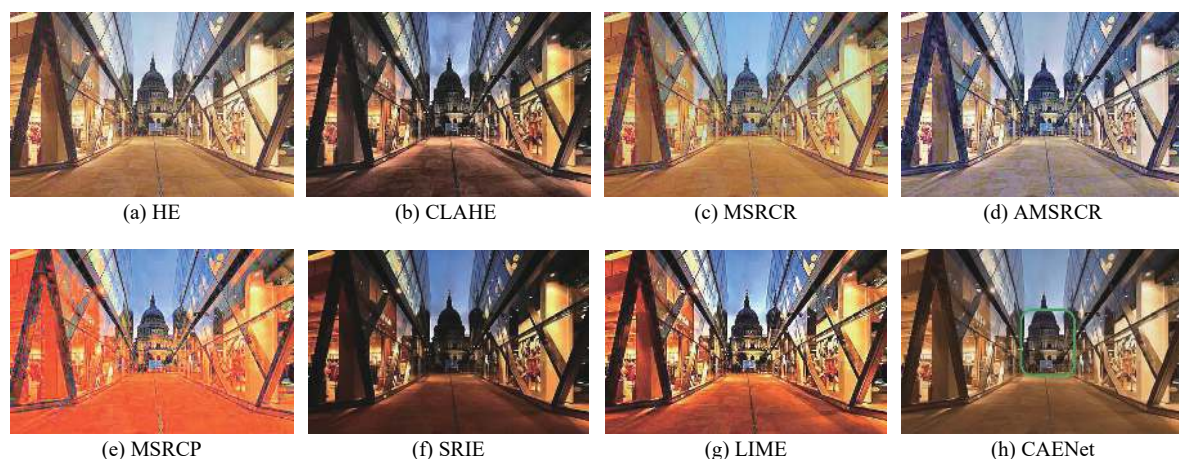


图 17 不同算法对建筑低光图像中门窗细节的增强效果对比

Fig.17 Comparison of enhancement effects of different algorithms on detail of doors and windows in low-light images of buildings

声和泊松噪声,如图18所示.在本次实验中,测试图像选取逆光图像,同时为了体现算法具有较强的鲁棒性和普适性,在进行测试时图像噪声数据大小随机给出.

如图19~22所示为不同算法对这4种噪声的处理效果.通过对比可以看出,本文算法在增强图像的同时能够在一定程度上减少噪声,而其他对比算法的降噪效果稍显不足.

表 1 不同图像增强算法与 CAENet 算法的结构相似度指标对比

Tab.1 Comparison of structural similarity index between different image enhancement algorithms and CAENet algorithm

SSIM	HE	CLAHE	MSRCR	AMSRCR	MSRCP	SRIE	LIME	CAENet
岩石细节图像	0.412 4	0.659 6	0.546 8	0.398 9	0.370 7	0.734 5	0.545 1	0.875 9
湖面细节图像	0.804 3	0.689 2	0.701 3	0.725 6	0.600 8	0.775 3	0.700 0	0.827 8
建筑细节图像	0.703 9	0.758 5	0.622 3	0.633 8	0.555 1	0.645 6	0.694 3	0.835 8

表 2 不同图像增强算法与 CAENet 算法的自然图像质量评价指标对比

Tab.2 Comparison of natural image quality evaluator between different image enhancement algorithms and CAENet algorithm

NIQE	HE	CLAHE	MSRCR	AMSRCR	MSRCP	SRIE	LIME	CAENet
岩石细节图像	7.900 6	8.459 4	8.163 6	7.825 8	7.179 6	8.970 7	8.259 4	7.428 4
湖面细节图像	4.687 2	5.135 0	5.317 0	4.610 2	4.665 8	4.151 3	5.714 1	4.526 7
建筑细节图像	5.086 8	5.061 9	5.357 1	5.603 4	5.186 1	4.827 3	5.681 6	3.864 8



图 18 低光含噪图像

Fig.18 Low light noisy images

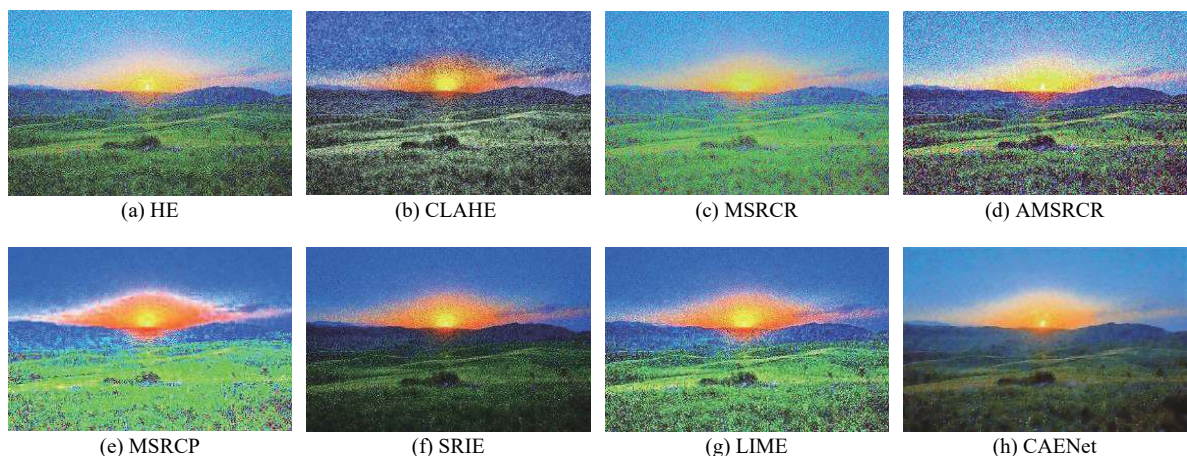


图 19 不同算法对高斯噪声的处理效果

Fig.19 Effect of different algorithms on Gaussian noise

图 23 显示了各个算法处理低光含噪图像的峰值信噪比 (PSNR)、相似度结构对比 (SSIM) 和自然图像质量评价指标 (NIQE). 在计算这 3 个值时, 由于图像中的噪声数据是随机给出的, 为了保持客观性, 对上述加噪图像一共进行 30 次图像测试, 最终计算平均值来进行比较. 从图 23(a) 可以看出, 在对 4 种噪声的处理中 CAENet 算法的

PSNR 值较大, 说明 CAENet 算法具有很好的鲁棒性, 能够在进行增强图像的同时较好地抑制噪声干扰. 从图 23(b) 可以看出, 各个算法对含有斑点噪声和泊松噪声的图像处理效果较为理想, 但是 CAENet 算法的 SSIM 值整体较高, 因此相较于其他对比算法具有一定的优势. 尽管 CAENet 算法具有较好效果, 对于不同噪声来说, CAENet 算法

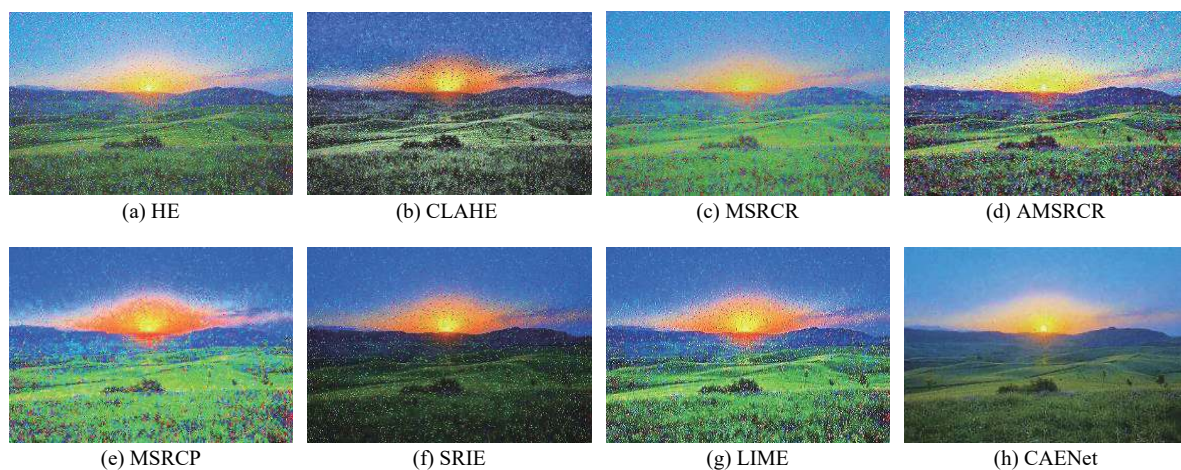


图 20 不同算法对椒盐噪声的处理效果

Fig.20 Effect of different algorithms on salt and pepper noise

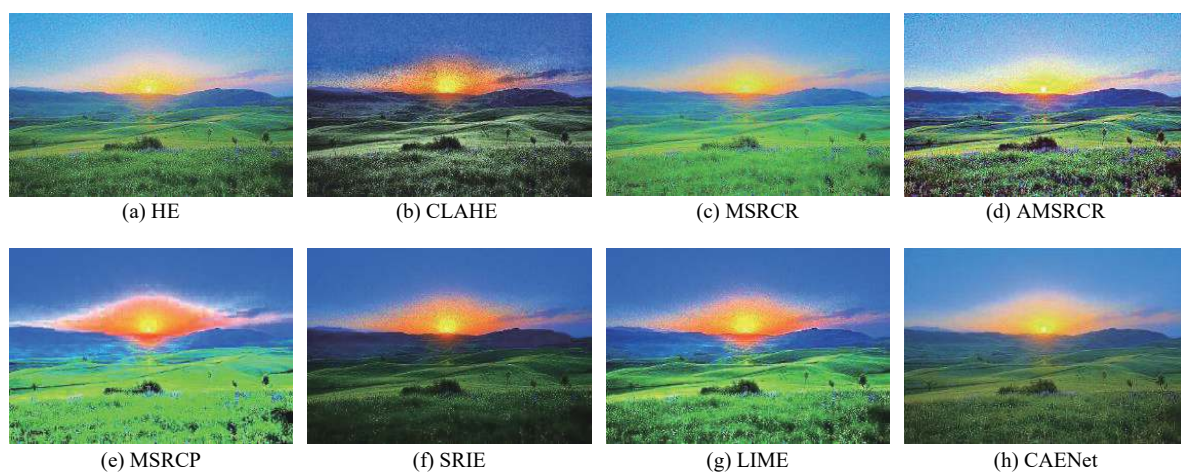


图 21 不同算法对斑点噪声的处理效果

Fig.21 Effect of different algorithms on speckle noise

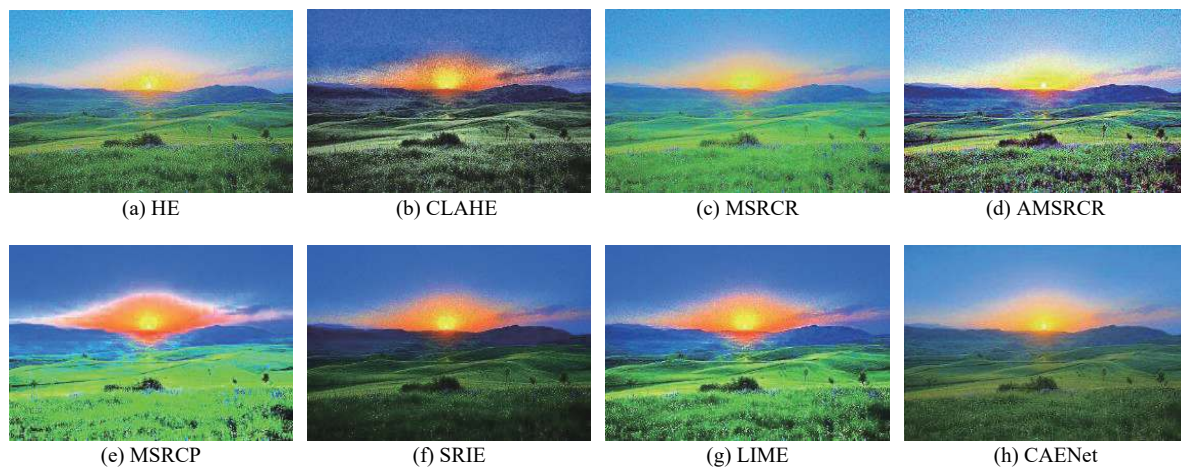


图 22 不同算法对泊松噪声的处理效果

Fig.22 Effect of different algorithms on poisson noise

的处理效果也不同, 其中对于高斯噪声的处理效果最为不理想. 从图 23(c) 可以看出, CAENet 算法对高斯噪声和椒盐噪声的处理效果优于其他算

法, 但是对于斑点噪声和泊松噪声, CAENet 算法相比于 MSRCP 算法来说稍显不足, 从视觉效果层面来说, CAENet 算法稍优于 MSRCP 算法.

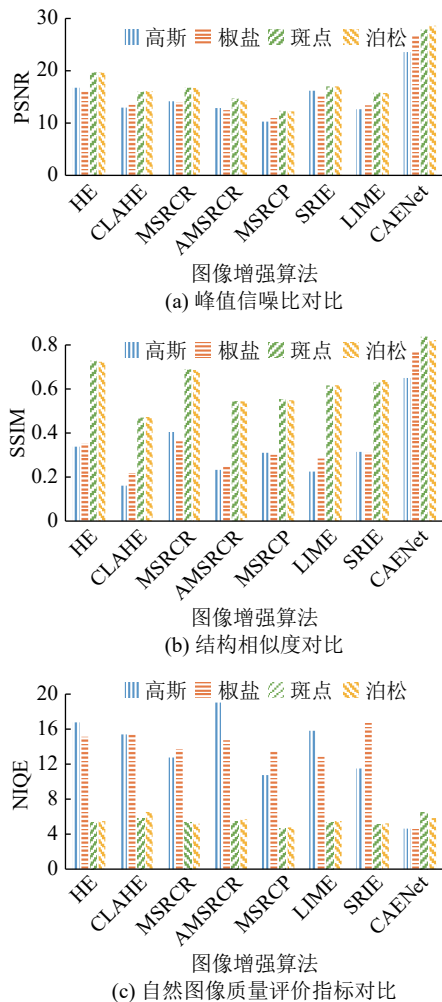


图23 不同图像增强算法评价指标的对比

Fig.23 Comparison of evaluation indexes by different image enhancement algorithms

4 结 论

本文针对图像增强算法 LLNet 在推广到三通彩色图像时会产生大量冗余参数的问题,利用卷积自编码器,提出了一种基于卷积自编码器的图像增强算法——CAENet.实验采用公共数据集 Corel5k 和高清数据集,将所提算法与 HE、CLAHE、MSRCR、AMSR、MSRCP、SRIE 和 LIME 等算法进行比较.实验结果表明,相比于其他图像增强算法,CAENet 能够在增强图像的同时保持图像细节不失真,同时对含噪低光图像也有较好的处理效果,为图像增强提供了一种新的方法.后续的工作将围绕探索训练数据规模和网络规模之间的关系展开,使该算法能够适用于大规模数据的处理.

参考文献 (References):

[1] GUO X, LI Y, LING H. LIME: low-light image

enhancement via illumination map estimation [J]. **IEEE Transactions on Image Processing**, 2017, 26(2): 982-993.

[2] 吕永标, 赵建伟, 曹飞龙. 基于复合卷积神经网络的图像去噪算法 [J]. 模式识别与人工智能, 2017, 30(2): 97-105.

LV Yong-biao, ZHAO Jian-wei, CAO Fei-long. Image denoising algorithm based on composite convolution neural network [J]. **Pattern Recognition and Artificial Intelligence**, 2017, 30(2): 97-105.

[3] RIVERA A R, RYU B, CHAE O. Content-aware dark image enhancement through channel division [J]. **IEEE Transactions on Image Processing**, 2012, 21(9): 3967-3980.

[4] SINGH K, KAPOOR R. image enhancement using exposure based sub image histogram equalization [J]. **Pattern Recognition Letters**, 2014, 36: 10-14.

[5] REZA A M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement [J]. **Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology**, 2004, 38(1): 35-44.

[6] 方帅, 杨静荣, 曹洋, 等. 图像引导滤波的局部多尺度 Retinex 算法 [J]. 中国图象图形学报, 2012, 17(7): 748-755.

FANG Shuai, YANG Jing-rong, CAO Yang, et al. A local multiscale Retinex algorithm for image guided filtering [J]. **Journal of Image and Graphics**, 2012, 17(7): 748-755.

[7] JOBSON D J, RAHMAN Z, WOODDELL G A. Properties and performance of a center/surround retinex [J]. **IEEE Transactions on Image Processing**, 1997, 6(3): 451-462.

[8] 汪荣贵, 张新彤, 张璇, 等. 基于 Zernike 矩的新型 Retinex 图像增强方法研究 [J]. 中国图象图形学报, 2011, 16(3): 310-315.

WANG Rong-gui, ZHANG Xin-tong, ZHANG Xuan, et al. Research on the new Retinex image enhancement method based on zernike moment [J]. **Journal of Image and Graphics**, 2011, 16(3): 310-315.

[9] WANG J, LU K, XUE J, et al. Single image dehazing based on the physical model and MSRCR algorithm [J]. **IEEE Transactions on Circuits and Systems for Video Technology**, 2017, PP(99): 1-1.

[10] PARTHASARATHY S, SANKARAN P. An automated multi scale retinex with color restoration for image enhancement [C] // **IEEE 2012 National Conference on Communications (NCC)**. Kharagpur: IEEE, 2012: 1-5.

[11] PETRO A B, SBERT C, MOREL J M. Multiscale

- Retinex [J]. **Image Processing on Line**, 2014: 71–88.
- [12] FU X, ZENG D, HUANG Y, et al. A weighted variational model for simultaneous reflectance and illumination estimation [C] // **Computer Vision and Pattern Recognition (CVPR)**. Las Vegas, Nevada, USA: IEEE, 2016.
- [13] 黄立威, 江碧涛, 吕守业, 等. 基于深度学习的推荐系统研究综述 [J]. **计算机学报**, 2018: 1–29.
HUANG Li-wei, JIANG Bi-tao, LV Shou-ye, et al. Review of the recommendation system based on deep learning [J]. **Chinese Journal of Computers**, 2018: 1–29.
- [14] SHEN L, YUE Z, FENG F, et al. Msr-net: low-light image enhancement using deep convolutional network [J]. **ArXiv Preprint**. ArXiv: 1711. 02488, 2017.
- [15] LI C, GUO J, PORIKLI F, et al. LightenNet: a convolutional neural network for weakly illuminated image enhancement [J]. **Pattern Recognition Letters**, 2018, 104: 15–22.
- [16] 王万良. 人工智能及其应用: 第 3 版 [M]. 北京: 高等教育出版社, 2015: 319.
- [17] LORE K G, AKINTAYO A, SARKAR S. LLNet: A deep autoencoder approach to natural low-Light image enhancement [J]. **Pattern Recognition**, 2017, 61: 650–662.
- [18] 黄健航, 雷迎科. 基于边际 Fisher 深度自编码器的电台指纹特征提取 [J]. **模式识别与人工智能**, 2017, 30(11): 1030–1038.
HUANG Jian-hang, LEI Ying-ke. Radio fingerprint feature extraction based on marginal fisher depth autoencoder [J]. **Pattern Recognition and Artificial Intelligence**, 2017, 30(11): 1030–1038.
- [19] VINCENT P, LAROCHELLE H, BRNGIO Y, et al. Extracting and composing robust features with denoising autoencoders [C] // **Proceedings of the 25th International Conference on Machine Learning (ICML2008)**. Helsinki: ACM, 2008: 1096–1103.
- [20] MASCI J, MEIER U, CIRESAN D, et al. Stacked convolutional auto-encoders for hierarchical feature extraction [C] // **International Conference on Artificial Neural Networks (ICANN)**. Springer, Berlin, Heidelberg: 2011: 52–59.
- [21] 袁非牛, 章琳, 史劲亭, 等. 自编码神经网络理论及应用综述 [J]. **计算机学报**, 2019, 42(1): 205–232.
YUAN Fei-niu, ZHANG Lin, SHI Jin-ting, et al. Self-encoding neural network theory and application review [J]. **Chinese Journal of Computers**, 2019, 42(1): 205–232.
- [22] 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.
- [23] 丁毅, 李玉惠, 李勃. 基于图像不同亮度区域特征的 Gamma 矫正方法 [J]. **计算机技术与发展**, 2016, 26(6): 37–39+45.
DING Yi, LI Yu-hui, LI Bo. Gamma correction method based on features of different brightness regions of images [J]. **Computer Technology and Development**, 2016, 26(6): 37–39.
- [24] 张菲菲, 谢伟, 石强, 等. 人眼视觉感知驱动的梯度域低照度图像对比度增强 [J]. **计算机辅助设计与图形学学报**, 2014, 26(11): 1981–1988.
ZHANG Fei-fei, XIE Wei, SHI Qiang, et al. Contrast enhancement of gradient domain low illumination image driven by human visual perception [J]. **Journal of Computer-Aided Design and Computer Graphics**, 2014, 26(11): 1981–1988.
- [25] 张林, 程华, 房一泉. 基于卷积神经网络的链接表示及预测方法 [J]. **浙江大学学报: 工学版**, 2018, 52(3): 552–559.
ZHANG Lin, CHENG Hua, FANG Yi-quan. Link representation and prediction method based on convolutional neural network [J]. **Journal of Zhejiang University: Engineering Science**, 2018, 52(3): 552–559.
- [26] GU J, WANG Z, KUEN J, et al. Recent advances in convolutional neural networks [J]. **Pattern Recognition**, 2018, 77: 354–377.
- [27] ZEILER M D, KRISHNAN D, TATLOR G W, et al. Deconvolutional networks [C] // **Computer Vision and Pattern Recognition (CVPR)**. Las Vegas: IEEE, 2010: 2528–2535.
- [28] 袁公萍, 汤一平, 韩旺明, 等. 基于深度卷积神经网络的车牌识别方法 [J]. **浙江大学学报: 工学版**, 2018, 52(4): 694–702.
YUAN Gong-ping, TANG Yi-ping, HAN Wang-ming, et al. A vehicle identification method based on deep convolution neural network [J]. **Journal of Zhejiang University: Engineering Science**, 2018, 52(4): 694–702.
- [29] LOZA A, BULL D R, HILL P R, et al. Automatic contrast enhancement of low-light images based on local statistics of wavelet coefficients [J]. **Digital Signal Processing**, 2013, 23(6): 1856–1866.
- [30] YE X, WANG L, XING H, et al. Denoising hybrid noises in image with stacked autoencoder [C] // **Information and Automation (ICIA)**, Lijiang: IEEE, 2015: 2720–2724.
- [31] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a “completely blind” image quality analyzer [J]. **IEEE Signal Processing Letters**, 2013, 20(3): 209–212.