# 计算机视觉

# 第 16 章 智能图像去噪与暗光处理



主讲教师: 全红艳

计算机科学与技术学院

# 本次课程内容

- 1. 自动编码去噪技术
- 2. 自动编码器原理
- 3. 常见的自动编码器
- 4. 降噪编码器工作

- 5. 降噪编码实例剖析
- 6. 栈式暗光自动编码
- 7. 基于CNN的暗光增强



# 1. 自动编码去噪技术(AUTOENCODER)介绍

- ◆ 自动编码器提出的背景
  - 编码、解码:对于神经网络,假设输出与输入相同,训练调整各层参数,得到每一层中的权重(权重就是编码解码的参数)
  - ▶ 特征: 输入图像I有多种特征表达, 每一层代表一种特征

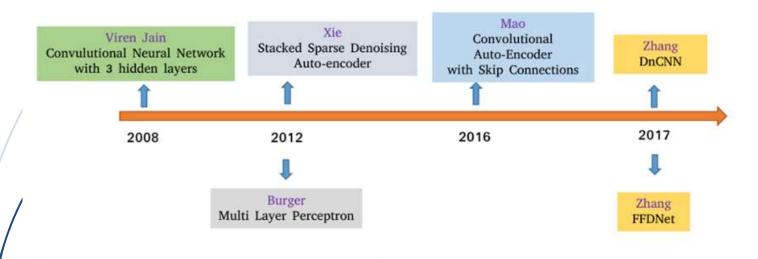
自动编码器的原则

尽可能复现输入信号的神经网络

需要捕捉有代表性的特征, 找到输入信息的主要成分

# 自动编码去噪技术发展

# > Deep Learning image denoising



<sup>[1]</sup> Jain V, Seung H S. Natural image denoising with convolutional networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2008;769-776.
[2] Xie J, Xu L, Chen E. Image denoising and inpainting with deep neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012;341-349.

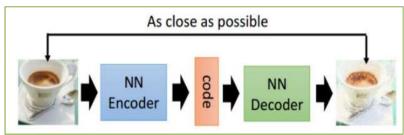
<sup>[3]</sup> Harmeling S. Image denoising: Can plain neural networks compete with BM3D?[C]// Computer Vision and Pattern Recognition. IEEE, 2012:2392-2399.
[4] Mao X J. Shen C, Yang Y B. Image Restoration Using Convolutional Auto-encoders with Symmetric Skip Connections[j]. 2016.

<sup>[5]</sup> Zhang K, Zuo W, Chen Y, et al. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising()]. IEEE Transactions on Image Processing, 2017, 26(7):3142.

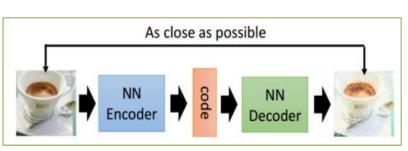
<sup>[6]</sup> Zhang K, Zuo W, Zhang L, FFDNet: Toward a Fast and Flexible Solution for CNN based Image Denoising[]]. 2017.

# 2. 自动编码器原理

- 采用重构的策略 (编码之后再解码重构)
- 目标: 重构误差最小
- 自动编码过程:调整编码-解码encoder和decoder的参数的过程,使得重构误 差最小
- /损失计算方法: 重构结果与原输入之间的差异







reconstruction

# 自动编码器(AUTOENCODER-AE)

- ◆ 自动编码器始于数据的压缩
  - > 对数据相关特征的压缩
  - > 使用神经网络提取与原始训练集高度相关的特征
  - 压缩后数据有损,在降维的过程中会丢失信息
  - ▶ 月标:编码解码后尽量无损

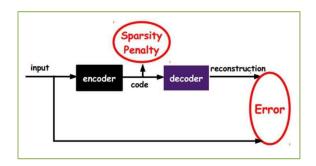
# 自动编码器深层网络训练与应用

- ◆ 深层网络训练
  - > 2012年, CNN中对自动编码器逐层预训练, 可训练更深层网络
  - > 2014年出现Batch Normalization技术,改进的初始化策略,可以训练更深 网络
  - > 2015年底,出现残差网络ResNet,可以训练任意深度的神经网络
- ◆/自劾编码器主要应用:图像(数据)去噪、可视化降维、生成数据

# 3. 常见的自动编码器

## (1) 稀疏自动编码器 SPARSE AUTOENCODER

- 在AutoEncoder的基础上加上L1的Regularity 限制(目的是产生稀疏权值 矩阵)
- ▶ 约束每一层中的节点,使其大部分为0,实现稀疏化
- > 限制每次得到的表达code尽量稀疏

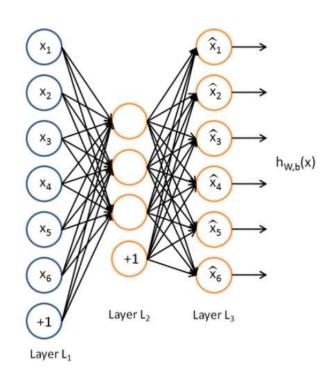


- input: X code:  $h = W^T X$ 

- loss:  $L(X; W) = ||W h - X||^2 + \lambda \sum_{j} |h_j|$ 

# (1) 稀疏自动编码器



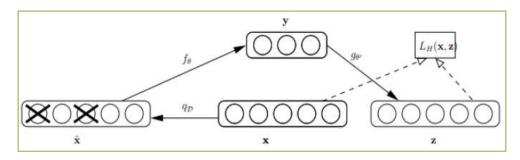


# (2) 收缩自动编码器 (CONTRACTIVE AUTOENCODER, CAE)

- ◆ CAE是自动编码器 (AE) 的一个变种
- ◆ 实质:不改变结构,在AE损失函数中加入了一个规则项,这样CAE可以被看作是 岭回归的神经形式

# (3) 降噪编码器及实例

- ♦ 2008年,Vincent提出了去噪自动编码概念(de-noising Autoencoder,dAE)
- ◆ dAE与人体的感官系统获取信息原理类似,例如人眼看物体时,如果物体某一小部分被 遮住了,人依然能够将其识别出来
- ◆ dAE的原理基于编码、解码原理:

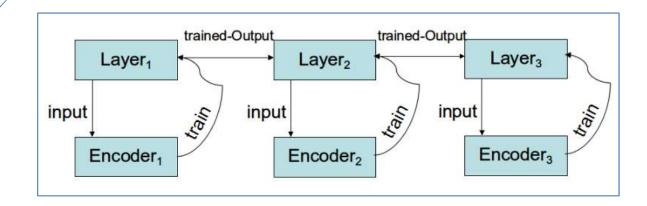


1 Vincent P, Larochelle H, Bengio Y, et al. Extracting and Composing Robust Features with Denoising Autoencoders[C]// International Conference on Machine Learning. 11ACM, 2008.

# (4) 栈式自动编码器及实例

## STACKED AUTOENCODER

- ◆ Bengio等人在2007年提出Stacked AutoEncoder,用于非监督学习
- ◆ 逐层初始化(Layer-wise Pre-training):通过逐层非监督学习的预训练,来初始化深度网络的参数,替代传统的随机方法



# 4. 降噪编码器工作 (De-noising Autoencoder, DA)

- 2008--Extracting and Composing Robust Features with Denoising Autoencoders
  - 该技术被广泛使用,无监督初始化步骤中,将输入映射到中间隐层特征 (编码)表示
  - > 学习噪声损坏图像的特征, 具有一定的鲁棒性





# 图像去噪自编码器

- 2010--Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion
  - 构建叠加去噪层的深层网络, 对普通自动编码器堆叠后的简单变种
  - 通过深度置信网络(DBN)改善性能,使得分类误差显著降低
  - 无监督学习

Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion.

We explore an original strategy for building deep networks, based on stacking layers of denoising autoencoders which are trained locally to denoise corr...

P Vincent, H Larochelle, I Lajoie, ...- 《Journal of Machi...-被引量: 1762 - 2010年

来源: AMS / ACM / ResearchGate / worldcat.org / ResearchGate ~



# 去噪自编码器的典型工作

- ◆ 2016--Medical image denoising using convolutional denoising autoencoders
  - 利用小样本量的卷积层构造去噪自编码器,有效地对医学图像进行去噪,提高去噪性能
- > 2017--Denoising auto-encoder based image enhancement for high resolution sonar image
  - > 利用卷积神经网络的自编码算法对声纳图像降噪,在自动编码器中学习大量声纳图像
- ◆ 2017--Medical Image Denoising Using Convolutional Denoising

  Autoencoders
  - 利用小样本的卷积层构造去噪自编码器,有效地对医学图像进行去噪,提高去噪性能



# 去噪自编码器的典型工作

- **♦ 2018-- Denoising Adversarial Autoencoders** 
  - 利用自动编码器从噪声样本中恢复目标,提出对抗性自动编码器方法去噪,利用 对抗性训练
- **♦ 2018--Blind Denoising Autoencoder** 
  - "盲去噪":从噪声样本本身获得模型结构,基于自动编码器的解决方案的盲去噪方法
- > 2018--Mitigation of through-wall interference in radar images using denoising autoencoders
  - 利用去噪自编码器将破坏后的穿墙雷达图像去噪为清晰结果
- **♦ 2019--Semantic denoising autoencoders for retinal optical coherence tomography**
  - ▶ 自动编码器与预先训练好的Resnet分类器相结合进行正则化,去噪后保留疾病特征.提高精细细节



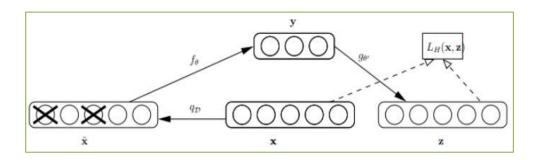
# 5. 降噪编码实例剖析

- A. 栈式暗光自然图像自动编码器的实例剖析(LLNet)
- B. 基于卷积神经网络的暗光图像增强(Low-light Image Enhancement)实例 剖析
- C. 基于先验知识的图像去噪方法



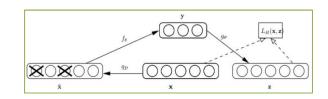
# (1) 降噪自动编码器的实例剖析(DAE)

- ◆ 通过对原始数据X进行人为损坏加随机噪声,得到X'
- ◆ X'作为网络的输入,重构出原始数据X
- ◆ 通过对损坏数据到修复结果的学习,利用编码-解码(Encoding-Decoding)的策略得到 图像修复的编码规律

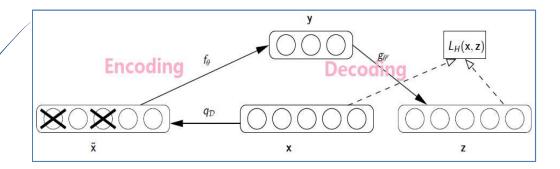




# 降噪编码器dAE实例



整体框架



14 incent P, Larochelle H , Bengio Y , et al. Extracting and Composing Robust Features with Denoising Autoencoders[C]// International Conference on Machine Learning. 19ACM, 2008.

# 降噪编码器网络性能

◆ 重构目标是最小化损失:

$$\theta^{\star}, \theta'^{\star} = \underset{\theta, \theta'}{\operatorname{arg \, min}} \frac{1}{n} \sum_{i=1}^{n} L\left(\mathbf{x}^{(i)}, \mathbf{z}^{(i)}\right)$$
$$= \underset{\theta, \theta'}{\operatorname{arg \, min}} \frac{1}{n} \sum_{i=1}^{n} L\left(\mathbf{x}^{(i)}, g_{\theta'}(f_{\theta}(\mathbf{x}^{(i)}))\right)$$



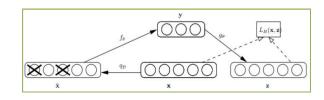
利用重构交叉熵计算损失,即利用负对数似然函数进行计算:

$$L_{\mathbf{H}}(\mathbf{x}, \mathbf{z}) = \mathbf{H}(\mathcal{B}_{\mathbf{x}} || \mathcal{B}_{\mathbf{z}})$$

$$= -\sum_{k=1}^{d} [\mathbf{x}_{k} \log \mathbf{z}_{k} + (1 - \mathbf{x}_{k}) \log(1 - \mathbf{z}_{k})]$$

Yincent P, Larothelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders[C]// International Conference on Machine Learning. 2008

# 降噪编码器实例细节



#### 整体框架

## 加噪过程

- > Choose a fixed proportion of components of x at random.
- Reset their values to 0.



2\fincent P, Larochelle H, Bengio Y, et al. Extracting and Composing Robust Features with Denoising Autoencoders[C]// International Conference on Machine Learning. 21ACM, 2008.

# 数据集问题

- ♦ basic: subset of original MNIST digits:
  - ◆ 10 000 training samples, 2 000 validation samples, 50 000 test samples



rot: applied random rotation (angle between 0 and  $2\pi$  radians)



bg-img: background is random patch from one of 20 images



bg-rand: background made of random pixels (value in 0...255)



rot-bg-img: combination of rotation and background image

Vincent P , Larochelle H , Bengio Y , et al. Extracting and Composing Robust Features with Denoising Autoencoders[C]// International Conference on Machine Learning. ACM, 2008.

# 6. 栈式暗光自动编码算法实例

- ◆ Lore K G, Akintayo A, Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement[J]. Pattern Recognition, 2015, 61:650-662.
- ◆ 提出<mark>栈式自动稀疏去噪编码器 (stacked sparse denoising autoencoder -SSDA)</mark> 识别低光图 像中的信号特征
- ◆ 对高动态范围图像(High-Dynamic Range,简称HDR)中图像像素不饱和的部分进行自适应增强



# LLNet 特点

- ◆ 利用非线性方法模拟暗光的输入条件,对自然图像进行变暗处理,构建训练数据集。经过训练,该 网络可以增强暗光图像
- ◆ 利用深度神经网络构建SSDA,不使用卷积层
- ◆ 实现的SSDA可以自适应学习,用于自然低光环境去操

# LLNet 研究路线

- ◆ 训练数据生成,通过对互联网数据库中可用图像修改,模拟低光环境
- ◆ 两种学习模型结构:
  - > 对比度增强和去噪同时学习
  - > 对比度增强和去噪顺序学习
- ◆ 对训练后网络性能进行了评估:
  - > 与其他方法进行了比较,包括合成噪声和人工变暗。
  - > 证明所训练模型可以用于手机相机获得的真实低光图像的增强

# LLNet 系统结构

## 两种去噪模块

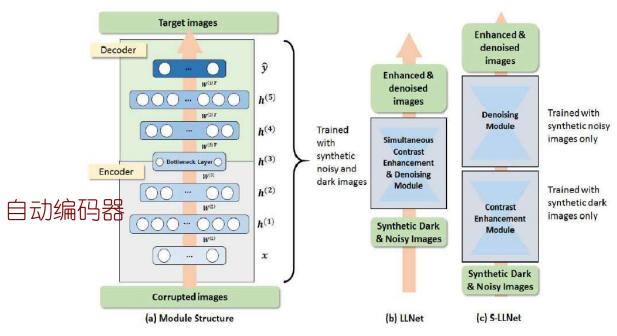
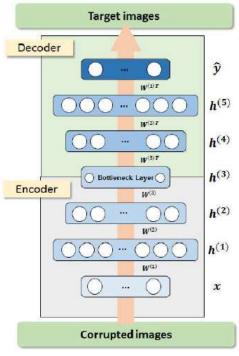


Figure 1. Architecture of the proposed framework: (a) An autoencoder module is comprised of multiple layers of hidden units, where the encoder is trained by unsupervised learning, the decoder weights are transposed from the encoder and subsequently fine-tuned by error backpropagation; (b) LLNet with a simultaneous contrast-enhancement and denoising module; (c) S-LLNet with sequential contrast-enhancement and denoising modules. The purpose of denoising is to remove noise artifacts often accompanying contrast enhancement.

# 自动编码器

◆ 噪声自动编码器(denoising autoencoders , DA)模块由多层隐藏单元组成,其中编码器通过无监督学习进行训练,解码器权重从编码器转移,然后通过错误反向传播进行微调。



模块单独进行训练,使用仅暗光图像训练集和仅含噪音的训练集

KG, Akintayo A, Sarkar S. LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement[J]. Pattern Recognition, 2015, 61:650-662.

# LLNet自动编码器

- ◆ y ∈ R<sup>N</sup> 表示未损坏源图像
- ◆ x ∈ R<sup>N</sup> 表示用损失加噪图像
- ♦ x = My, M  $\in$  R<sup>N×N</sup> M为是非分析高维矩阵,假设M将干净图像进行了破坏
- ◆ DA: 为前馈学习函数的定义,通过以下表达式来描述m的每个元素
- ullet recoding functions  $h(\mathbf{x}) = \sigma(\mathbf{W}\mathbf{x} + \mathbf{b})$ 
  - decoding function  $\hat{\mathbf{y}}(\mathbf{x}) = \sigma'(\mathbf{W'h} + \mathbf{b'})$

# LLNet自动编码器

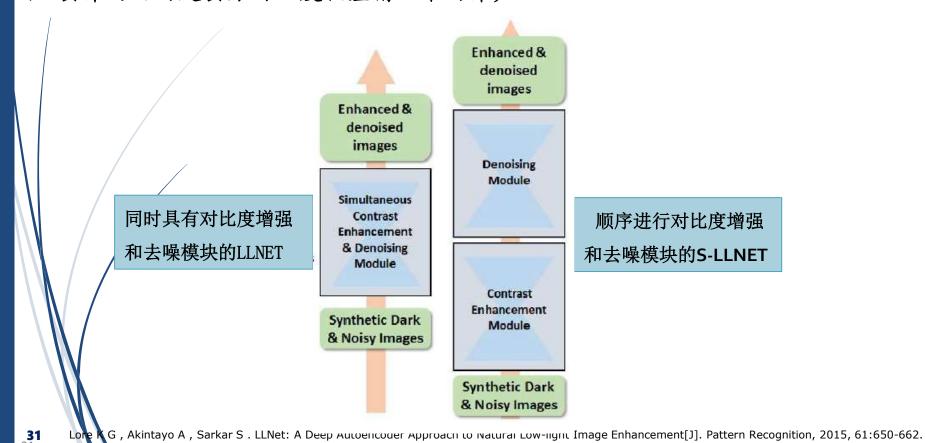
- $\mathbf{W} \in \mathcal{R}^{K \times N}$  and  $\mathbf{b} \in \mathcal{R}^{K}$  are the weights and biases of each layers of encoder  $\mathbf{W}' \in \mathcal{R}^{N \times K}$  and  $\mathbf{b}' \in \mathcal{R}^{K}$  are the weights and biases for each layer of the decoder.
- $h(\mathbf{x}) \in \mathcal{R}^K$  is the activation of the hidden layer  $\hat{\mathbf{y}}(\mathbf{x}) \in \mathcal{R}^N$  is the reconstruction of the input (i.e. the output of the DA)

# LLNet 网络的参数

- ◆ 网络参数:LLNET由3个DA层组成,第一个DA层取尺寸为17×17像素(即289个输入单元)的输入图像。第一层有867个隐藏单元,第二层有578个隐藏单元,第三层有289个隐藏单元,成为瓶颈层。
- ◆ 在第三DA层之外形成前三层的解码对应,因此第四层和第五层分别有578和867个隐藏单元。 输出单元与输入单元具有相同的尺寸,即289。
- ◆ 对网络进行了30个阶段的预训练,前两个DA层的预训练学习率为0.1,最后一个DA层的预训练学习率为0.01,而前200个微调阶段的学习率为0.1,之后为0.01,只有在验证误差改善小于 0.5%时才会停止微调。对于s-LLnet,每个模块的参数是相同的。

# 自动编码器

◆ 去噪的目的是去除对比度较差情况下的噪声



# 模型训练

- ◆ 训练模型: 使用422500片进行训练, 从169个标准测试图像中提取1。
- ◆ 将图像像素标准化为0到1之间。从同一图像的随机位置(以及随机变暗和噪声参数) 生成了2500个patch小块。
- ◆ 使用matlab命令imaadjust对17×17像素的patch进行非线性变暗,以随机应用 gamma调整,产生幂运算,进行增强

# 实验及对比

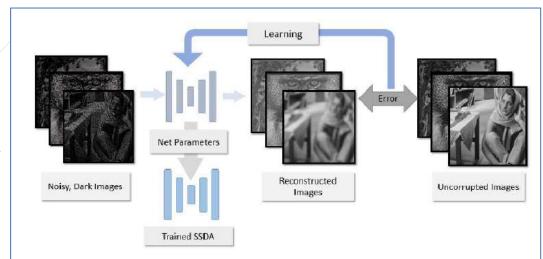


Figure 2. Training the LLNet: Training images are synthetically darkened and added with noise. These images are fed through LLNet where the reconstructed images are compared with the uncorrupted images to compute the error, which is then backpropagated to finetune and optimize the model weights and biases.

# 实验及对比

- ◆ LLNET和SLLNET输出重建结果,利用17×17块、跨距大小为3×3的块进行重建的
- ◆ 训练: Nvidia的TitanxGPU使用Theano的深度学习框架,耗时约30小时
- ◆ 预测: 增强图像 512×512像素的尺寸, 在GPU上为0.42秒。

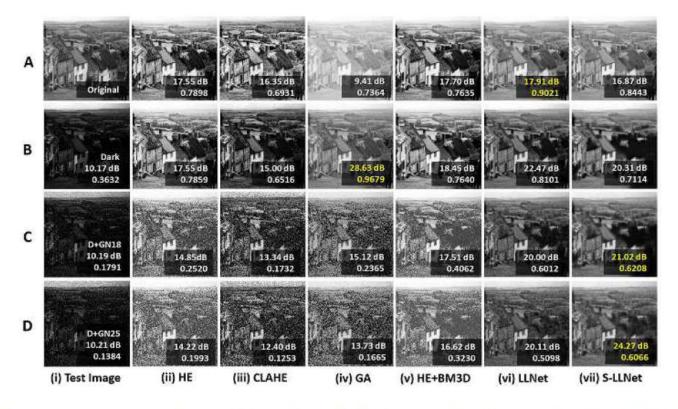


Figure 4. Comparison of methods of enhancing 'Town' when applied to (A) original already-bright, (B) darkened, (C) darkened and noisy ( $\sigma = 18$ ), and (D) darkened and noisy ( $\sigma = 25$ ) images. Darkening is done with  $\gamma = 3$ . The numbers with units dB are PSNR, the numbers without are SSIM. Best viewed on screen.

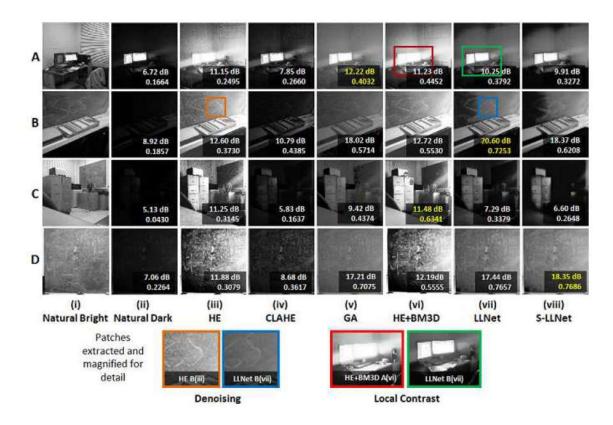
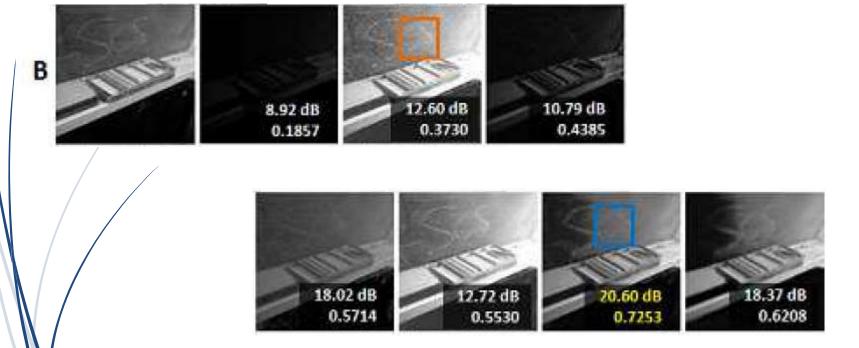
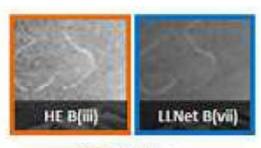


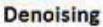
Figure 5. Comparison of methods of enhancing naturally dark images of (A) computer, (B) chalkboard, (C) cabinet, and (D) writings. Selected regions are enlarged to demonstrate the denoising and local contrast enhancement capabilities of LLNet. HE (including HE+BM3D) results in overamplification of the light from the computer display whereas LLNet was able to avoid this issue. In the natural case, the best performers in terms of PSNR and SSIM coincide. Best viewed on screen.



## 暗光增强的结果

Patches extracted and magnified for detail







#### 7. 基于CNN的暗光图像增强实例剖析

- ◆ LLCNN: A convolutional neural network for low-light image enhancement
- ◆ 基于CNN的暗光图像增强方法
- ◆ 利用多尺度特征映射 (feature maps), 避免了梯度消失问题
- ◆ 使用\$SIM(结构相似性)损失训练模型,尽可能地保留图像纹理细节
- ◆ 自适应地增强暗光图像的对比度



#### **LLCNN**

- ◆ 设计的CNN网络称为暗光卷积神经网络(LLCNN)
- ◆ 利用不同卷积核进行学习,结合多尺度特征映像产生增强的结果
- ◆ 目的是获取正常光照条件,并保留原始特征和纹理
- ◆ 集成了SSIM损失, 重建精确的图像纹理
- ◆ 希望达到深层神经网络学习的效果
- ◆ 处理消失梯度问题
- 网络设计思路受到Inception网络和残差网络启发



#### RESIDUAL LEARNING 残差学习(RESNET)—相关知识

- ◆ 残差网络 Kaiming He等提出:
- He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. 2015.
- ◆ 解决的问题: 当网络的层数急剧增加时, 导数的使梯度慢慢减小直至消失
- ◆ 假定输入是x,期望输出是H(x),需要学习的目标就是F(X)=H(x)-x, Resnet相当于将学习目标改变了,不再是学习H(x),而是H(x)-x,即残差
- ◆ 在ImageNet数据集,评估深度达152层时,

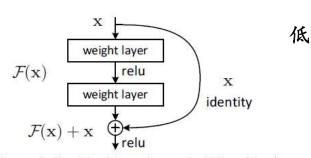


Figure 2. Residual learning: a building block.



#### INCEPTION结构—相关知识

- 为了获得高质量模型, 增加模型深度(层数)或宽度(层核或神经元数)出现问题:
  - 1. 参数太多, 若训练数据集有限, 容易过拟合
  - 2.网络越大计算复杂度越大,难以应用
  - 3.网络越深,梯度越往后穿越容易消失,难以优化模型
- Inception架构的主要思想是找出如何用密集成分来近似最优的局部稀疏解
- /Inception V4的结构结合了残差神经网络ResNet
- - goos [v1] Going Deeper with Convolutions, 6.67% test error, http://arxiv.org/abs/1409.4842
    - [v2] Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, 4.8% test error, http://arxiv.org/abs/1502.03167
    - [v3] Rethinking the Inception Architecture for Computer Vision, 3.5% test error, http://arxiv.org/abs/1512.00567
    - [v4] Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, 3.08% test error, http://arxiv.org/abs/1602.07261



#### GOOGLE INCEPTION V1 - V4 —相关知识

- ◆ Inception v1的网络,将1x1,3x3,5x5的conv和3x3的pooling结合,增加了网络的width,增加了网络对尺度的适应性
- ◆ v2: 在v1的基础上加入了BN层,每一层的输出都规范化到一个N(0, 1)的高斯,学习VGG用2个3x3的conv替代inception模块中的5x5,既降低了参数数量,也加速计算
- ◆ **v3**: 将7x7分解成两个一维的卷积(1x7,7x1),3x3也是一样(1x3,3x1), 既可以加速计算,又可以使网络深度进一步增加
- ◆ V4: Inception模块结合Residual Connection, ResNet结构可极大地加速 训练,性能有所提升,Inception-ResNet v2网络中设计了更深更优化的 Inception v4模型



## LLCNN网络结构—由特殊模块构成

- ◆ 特殊模块: The special-designed convolutional module
- ◆ 特殊模块设计:为了达到深度神经网络的效果, 受到残差网络和 inception的结构的启发
- ◆ 特殊模块设计采用两阶段学习:
  - ◆ /在第一阶段, 经过两路: 一个1×1的卷积层, 另一路经过2个3×3的卷积层, 两路结果相加作为第二阶段的输入, 特征数均为64
  - ◆ 第二阶段,经过两路:一个经3×3卷积卷积后输入残差学习模块;另一路直接 直接接入残差学习模块



### LLCNN网络结构

- ◆ 网络结构:
  - > 一个卷积层用于预处理以产生均匀的输入,另一个卷积层用于生成增强图像
  - 在这两个层之间使用几个特殊模块
  - 滤波器 (卷积核) 个数 (通道)
  - ◆ 最后一层卷积核个数:取决于颜色通道的数量
  - ◆ 其它层卷积核

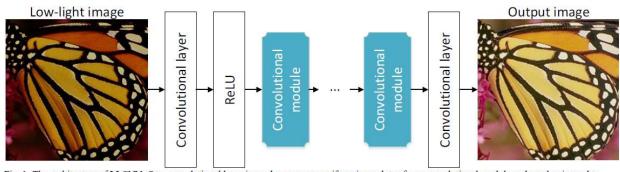


Fig. 1. The architecture of LLCNN. One convolutional layer is used to generate uniform input data of our convolutional module and another is used to generate output image. The number of convolutional module is flexible.



## 损失函数设计

**SSIM** 值范围为 (0.11. 如果为 1 ,表示两幅图像完全相同,因此,每个像素的损 1 - SSIM(p)

$$\ell_{\textit{ssim}} = \frac{1}{N} \sum_{p \in P} 1 - \textit{SSIM}(p)$$
 整体图像损失定  $\mathcal{S}_{\perp}$  . . .



## 无监督学习

- ◆ 在训练时,因为没有ground-truth图像
- ◆ 使用普通的自然图像作为目标图像, 其损失(破坏、加噪)之后的图像作为暗光图
- ◆ 利用网络生成的恢复结果图像与自然采集的目标图像进行比对,计算损失(使得其收敛为最小),调整网络参数,以求得最佳的(恢复图像)网络编码像。

## 实验结果



#### 通过对比组实验,从网络深度,Loss形式,CNN模型角度进行对比:

	<b>PSNR</b>	SSIM	LOE	SNM
Dark	12.70	0.476	50.82	0.118
SSR	22.33	0.908	25.93	0.571
MSRCR	25.97	0.894	65.93	0.612
SRIE	14.83	0.605	56.82	0.200
LIME	19.17	0.719	92.70	0.565
DHE	13.35	0.546	58.30	0.195
CLAHE	16.54	0.686	75.73	0.346
Li's	18.63	0.754	27.98	0.381
CNN baseline	26.08	0.890	31.68	0.545
LLCNN	35.20	0.957	10.27	0.598



Fig. 5. An example of the comparison with other contrast enhancement method for low-light images.





# End

