Towards Unsupervised Deep Image Enhancement with Generative Adversarial Network

Zhangkai Ni, Graduate Student Member, IEEE, Wenhan Yang, Member, IEEE,
Shiqi Wang, Member, IEEE, Lin Ma, Member, IEEE, and Sam Kwong, Fellow, IEEE

摘要

摘要.Improving the aesthetic quality of images is challenging and eager for the public. To address this problem, most existing algorithms are based on supervised learning methods to learn an automatic photo enhancer for paired data, which consists of low-quality photos and corresponding expert-retouched versions. However, the style and characteristics of photos retouched by experts may not meet the needs or preferences of general users. In this paper, we present an unsupervised image enhancement generative adversarial network (UEGAN), which learns the corresponding image-to-image mapping from a set of images with desir characteristics in an unsupervised manner, rather than learning on a large number of paired images. The proposed model is based on single deep GAN which embeds the modulation and attention mechanisms to capture richer global and local features.Based on the proposed model, we introduce two losses to deal with the unsupervised image enhancement: (1) fidelity loss, which is defined as a ‘2regularization in the feature domain of a pre-trained VGG network to ensure the content between the enhanced image and the input image is the same, and (2) qualityloss that is formulated as relativistic hinge adversarial loss to endow the input image the desired characteristics. Both quantitative and qualitative results show that the proposed model effectively improves the aesthetic quality of images

关键词： Terms—Unsupervised learning, image enhancement,global attention, generative adversarial network.

1 引言

随着图像增强在数字时代的背景和重要性的提高，其中摄影已成为日常生活中无处不在的一部分。文章强调了专业级图像编辑与业余摄影师能力之间的差距，后者通常难以在其照片中实现所需的美学品质。作者讨论了手动图像编辑的局限性，这是一项耗时且需要技能的任务，以及在自动化此过程以满足多样的审美偏好方面面临的挑战。

当前自动图像增强方法存在着明显的缺陷。它强调了这些方法对监督学习的依赖，这需要大量的带有前后图像的数据集。作者认为，这种对配对训练数据的依赖是一个重要的限制，因为它无法适应无配对数据，也无法捕捉用户之间广泛的审美偏好。他们提出了需要一种无监督学习方法来进行图像增强，能够适应各种风格和偏好，而无需配对的训练数据。这为介绍他们的新方法创造了条件，旨在弥合这些差距。

2 相关工作

在相关工作部分，作者提供了对图像增强领域的全面概述。他们将现有方法分为两大类：传统方法和基于学习的方法。传统方法进一步分为直方图调整、锐化掩蔽和Retinex-based技术

，每种方法都具有一系列优势和缺点。然后，作者讨论了基于学习的方法，强调了该领域近年来朝着这些方法的转变。

基于学习的方法分为监督学习和无监督学习。作者对各种监督学习方法进行了详细回顾，注意到它们对带有配对图像的大型数据集的依赖。他们指出这些方法在满足多样的审美偏好和处理无配对数据方面的固有局限性。这个讨论导致了对无监督学习方法的探讨，作者认为这是未来研究的一个有前景的方向。他们对现有的无监督模型进行了批判性分析，指出了这些模型存在的差距和局限性，而他们提出的模型旨在解决这些问题。

3 本文方法

3.1 本文方法概述

作者提出了一种无监督图像增强生成对抗网络（UEGAN），它在无监督方式下学习图像到图像的映射，而不是基于大量成对图像的学习。该模型基于单一的深度生成对抗网络（GAN），嵌入了调制和注意机制，以捕捉更丰富的全局和局部特征。作者引入了两种损失函数来处理无监督图像增强：(1) 保真度损失，定义为预训练的 VGG网络特征域中的L2正则化，以确保增强后的图像和输入图像的内容相同；(2) 质量损失 ，以相对真实的铰链对抗损失来赋予输入图像所需的特征。

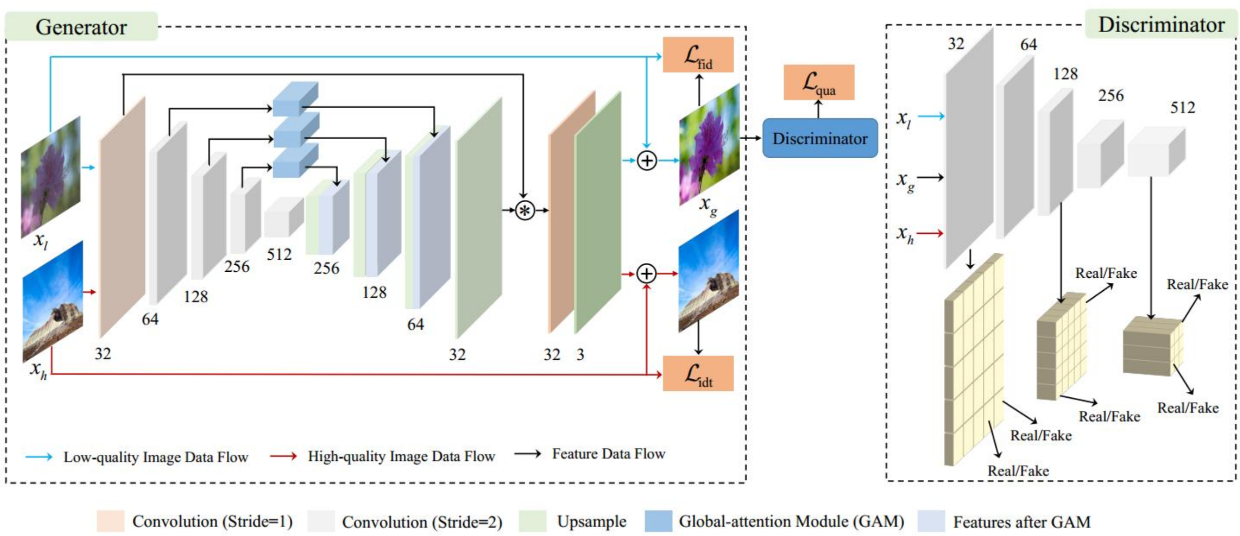

图 1. 方法示意图

4 复现细节

4.1 与已有开源代码对比

4.1.1.该换L1loss为**charbonnier loss**

原论文LOSS:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{qua}}\mathcal{L}_{\text{qua}}^{G} + \lambda_{\text{fid}}\mathcal{L}_{\text{fid}} + \lambda_{\text{idt}}\mathcal{L}_{\text{idt}},$$

其中：

$$\mathcal{L}_{\text{idt}} = \mathbb{E}_{x_h \sim P_h}\left[\left\|x_h - G(x_h)\right\|_1\right].$$

更改其中的L1loss为charbonnier loss

Charbonnier Loss，也称为L1 Charbonnier Loss或Smooth L1 Loss，是一种平滑的L1 Loss（绝对值误差损失）。它在计算目标与预测之间的差异时，相比标准的L1 Loss更加平滑，对于异常值（outliers）更具鲁棒性：

$$L(x,y) = \sqrt{(x-y)^2 + \varepsilon^2}$$

代码改动如下：

```
class CharbonnierLoss(nn.Module):
    def __init__(self,epsilon=1e-3):
        super(CharbonnierLoss, self).__init__()
        self.epsilon2=epsilon*epsilon

    def forward(self,x):
        value=torch.sqrt(torch.pow(x,2)+self.epsilon2)
        return torch.mean(value)
```

4.1.2.在编码的最后阶段加入**通道注意力**机制,代码改动如下：

（1）：

```
class SE(nn.Module):

    def __init__(self, in_channel, ratio=4):
        super().__init__()


        self.avg_pool = nn.AdaptiveAvgPool2d(output_size=1)

        self.fc1 = nn.Linear(in_features=in_channel, out_features=in_channel//ratio, bias=False)

        self.relu = nn.ReLU()

        self.fc2 = nn.Linear(in_features=in_channel//ratio, out_features=in_channel, bias=False)

        self.sigmoid = nn.Sigmoid()


    def forward(self, inputs):  # inputs 代表输入特征图

        b, c, h, w = inputs.shape

        x = self.avg_pool(inputs)

        x = x.view([b,c])

        x = self.fc1(x)

        x = self.relu(x)

        x = self.fc2(x)
```

```
        x = self.sigmoid(x)

        x = x.view([b,c,1,1])

        outputs = x * inputs
        return outputs
```

（2）：

```
def forward(self, x):
        ### encoder
        x1 = self.enc1( x)
        x2 = self.enc2(x1)
        x3 = self.enc3(x2)
        x4 = self.enc4(x3)
        x5 = self.enc5(x4)
        x5 = self.ga5(x5)
        x5 = SE(x5)
```

4.2 实验环境搭建

- Python 3.6
- PyTorch 1.4.0
- tqdm 4.43.0
- munch 2.5.0
- torchvision 0.5.0

5 实验结果分析

本部分对实验所得结果进行分析， 详细对实验内容进行说明， 实验结果进行描述并分析。

5.1 数据集

采用了MIT-Adobe FiveK Dataset数据集

5.2 实验结果

采用了三个质量评估指标，分别为PSNR,SSIM和NIMA。

实验结果如下：

	PSNR	SSIM	NIMA
Ori	22.88	0.8882	4.76
Our_1	22.98	0.8888	4.96
Our_2	23.99	0.8991	5.10

其中，Our_1是将L1loss为charbonnier loss，Our_2是在Our_1的基础上加入了通道注意力机制

图 3. 实验结果示意

6 总结与展望

该方法是完全无监督的，不可避免地存在局限性。在结果图像上出现的一个典型的伪影是颜色偏差。尽管它们有时会巧合地产生更令人愉悦的结果，但这种调整改变了内容，使结果看起来不真实。我们的方法无法从生成的结果中去除噪声。然而，这种噪点在曝光不足的图像中很常见

参考文献

[1] framework and its application for image contrast enhancement," IEEE

[2] Transactions on Image Processing, vol. 18, no. 9, pp. 1921–1935, 2009.

[3] [2] G. Thomas, D. Flores-Tapia, and S. Pistorius, "Histogram specification:

[4] A fast and flexible method to process digital images," IEEE Transactions

[5] on Instrumentation And Measurement, vol. 60, no. 5, pp. 1565–1578,

[6] 2011.

[7] [3] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered

[8] difference representation of 2D histograms," IEEE Transactions on

[9] Image Processing, vol. 22, no. 12, pp. 5372–5384, 2013.

[10] [4] D. Coltuc, P. Bolon, and J.-M. Chassery, "Exact histogram specification,"

[11] IEEE Transactions on Image Processing, vol. 15, no. 5, pp. 1143–1152,

[12] 2006.

[13] [5] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae,

[14] "A dynamic histogram equalization for image contrast enhancement,"

[15] IEEE Transactions on Consumer Electronics, vol. 53, no. 2, pp. 593–

[16] 600, 2007.

[17] [6] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, "Deep

[18] bilateral learning for real-time image enhancement," ACM Transactions

[19] on Graphics, vol. 36, no. 4, p. 118, 2017.

[20] [7] A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool,

[21] "DSLR-quality photos on mobile devices with deep convolutional

[22] networks," in Proceedings of The IEEE International Conference on

[23] Computer Vision, 2017, pp. 3277–3285.

[24] [8] Z. Yan, H. Zhang, B. Wang, S. Paris, and Y. Yu, "Automatic photo adjustment using deep neural networks," ACM Transactions on Graphics,

[25] vol. 35, no. 2, p. 11, 2016.

[26] [9] W. Ren, S. Liu, L. Ma, Q. Xu, X. Xu, X. Cao, J. Du, and M.-H.

[27] Yang, "Low-light image enhancement via a deep hybrid network," IEEE

[28] Transactions on Image Processing, vol. 28, no. 9, pp. 4364–4375, 2019.

[29] [10] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia,
[30] "Underexposed photo enhancement using deep illumination estimation,"
[31] in Proceedings of The IEEE Conference on Computer Vision And Pattern
[32] Recognition, 2019, pp. 6849–6857.
[33] [11] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, "Learning photographic global tonal adjustment with a database of input/output image
[34] pairs," in Proceedings of The IEEE Conference on Computer Vision And
[35] Pattern Recognition, 2011, pp. 97–104.
[36] [12] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image
[37] translation using cycle-consistent adversarial networks," in Proceedings
[38] of The IEEE International Conference on Computer Vision, 2017, pp.
[39] 2223–2232.
[40] [13] Y.-S. Chen, Y.-C. Wang, M.-H. Kao, and Y.-Y. Chuang, "Deep photo
[41] enhancer: Unpaired learning for image enhancement from photographs
[42] with Gans," in Proceedings of The IEEE Conference on Computer Vision
[43] And Pattern Recognition, 2018, pp. 6306–6314.
[44] [14] C. H. Lin, C.-C. Chang, Y.-S. Chen, D.-C. Juan, W. Wei, and H.-T. Chen,
[45] "COCO-GAN: generation by parts via conditional coordinating," in
[46] Proceedings of the IEEE International Conference on Computer Vision,
[47] 2019, pp. 4512–4521.
[48] [15] Y. Hu, H. He, C. Xu, B. Wang, and S. Lin, "Exposure: A whitebox photo post-processing framework," ACM Transactions on Graphics,
[49] vol. 37, no. 2, p. 26, 2018.
[50] [16] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou,
[51] and Z. Wang, "EnlightenGAN: Deep light enhancement without paired
[52] supervision," arXiv Preprint arXiv:1906.06972, 2019.
[53] [17] H. Ibrahim and N. S. P. Kong, "Brightness preserving dynamic histogram
[54] equalization for image contrast enhancement," IEEE Transactions on
[55] Consumer Electronics, vol. 53, no. 4, pp. 1752–1758, 2007.
[56] [18] J. A. Stark, "Adaptive image contrast enhancement using generalizations
[57] of histogram equalization," IEEE Transactions on Image Processing,
[58] vol. 9, no. 5, pp. 889–896, 2000.
[59] [19] W. Ye and K.-K. Ma, "Blurriness-guided unsharp masking," IEEE
[60] Transactions on Image Processing, vol. 27, no. 9, pp. 4465–4477, 2018.
[61] [20] S. K. Mitra, H. Li, I.-S. Lin, and T.-H. Yu, "A new class of nonlinear
[62] filters for image enhancement," in Proceedings of The IEEE International Conference on Acoustics, Speech, And Signal Processing, 1991,
[63] pp. 2525–2528.
[64] [21] K. He, J. Sun, and X. Tang, "Guided image filtering," IEEE Transactions
[65] on Pattern Analysis And Machine Intelligence, vol. 35, no. 6, pp. 1397–
[66] 1409, 2012.

[67] [22] A. Polesel, G. Ramponi, and V. J. Mathews, “Image enhancement via
[68] adaptive unsharp masking,” IEEE Transactions on Image Processing,
[69] vol. 9, no. 3, pp. 505–510, 2000.
[70] [23] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, “Structure-revealing lowlight image enhancement via robust retinex model,” IEEE Transactions
[71] on Image Processing, vol. 27, no. 6, pp. 2828–2841, 2018.
[72] [24] S. Wang, J. Zheng, H.-M. Hu, and B. Li, “Naturalness preserved
[73] enhancement algorithm for non-uniform illumination images,” IEEE
[74] Transactions on Image Processing, vol. 22, no. 9, pp. 3538–3548, 2013.
[75] [25] X. Guo, Y. Li, and H. Ling, “LIME: Low-light image enhancement
[76] via illumination map estimation.” IEEE Transactions Image Processing,
[77] vol. 26, no. 2, pp. 982–993, 2017.
[78] [26] Z. Ying, G. Li, Y. Ren, R. Wang, and W. Wang, “A new low-light image
[79] enhancement algorithm using camera response model,” in Proceedings
[80] of The IEEE International Conference on Computer Vision, 2017, pp.
[81] 3015–3022.