

Low_Light

▸ 2017 3

▸ 2020 21

▸ 2021 30

▼ 语义先验

- ▼ (MM 2020) Integrating Semantic Segmentation and Retinex Model for Low Light Image Enhancement

🔗 [Integrating Semantic Segmentation...](#)

▪

Minhao Fan, Wenjing Wang
Peking University

Wenhan Yang
City University of Hong Kong

Jiaying Liu*
Peking University

▼ Motivation

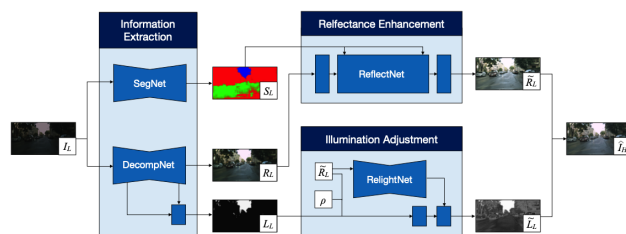
- 语义信息可以为低光照图像增强提供丰富的信息

▼ Contribution

- 提出一种结合Retinex模型和深度学习的方法，并且加入了语义先验。
- 语义先验通过空间变换处理特征来指导光照和反射率的联合增强，从而提高区域恢复的恢复质量。
- 建立了一种新的低光照图像合成模型，考虑曝光、颜色、噪声等因素。并且收集了100张图像。

▼ Method

▪



- 包含三个模块，四个子网络
- DecompNet：利用Retinex理论，将input分解为反射图和照度图。用成对的低光和正常光图像训练。
- SegNet：一个轻量的U-Net，只用来区分sky、ground、foreground objects。
- Reflectance Enhancement：基于RIR组成，结合语义特征，设计了一种SRIR，输出一个预测的反射图。

- Illumination Adjustment: 调节亮度

▼ Experiment

- Dataset: 基于Cityscapes和Camvid生成的2458对图像，2118训练，340验证。100张真实低光照图像用作测试。
- Train: 每个子网络单独训练，分为四个阶段。
- results:
- code

[GitHub - XFW-go/ISSR](https://github.com/XFW-go/ISSR)

▼ 思考

- 没有用到perceptual loss，主要用到MSE、SSIM、Grad。
- 文中只区分了天空陆地和前景，类别太少，对于真实场景增强，不够用

▼ (IJCAI 2018)When Image Denoising Meets High-Level Vision Tasks A Deep Learning Approach

<https://arxiv.org/pdf/1706.04284.pdf>

▪

Ding Liu¹, Bihan Wen¹, Xianming Liu², Zhangyang Wang², Thomas S. Huang¹ *

¹ University of Illinois at Urbana-Champaign, USA

- 第一个将去噪和语义分割及其他高级视觉任务结合的工作。

▼ Motivation

- 由于许多算法用MSE做去噪，因此重要的细节很容易丢失，导致图像退化。
- noise很容易影响网络的分类结果。

▼ Contribution

- 提出一种连接图像去噪和高级视觉任务的方法，同时最小化图像重建损失和高级任务损失。证明了语义信息对去噪的重要性。
- 证明了用共同损失训练网络，不仅可以增强去噪网络的效果，并且可以提高高级视觉任务的精确度。而且这样训练出来的去噪网络在各种高级视觉任务上都能够泛化。

▼ Method

▪

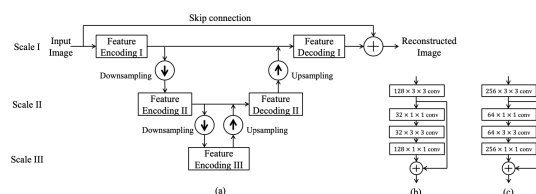


Figure 2: (a) Overview of our proposed denoising network. (b) Architecture of the feature encoding module. (c) Architecture of the feature decoding module.

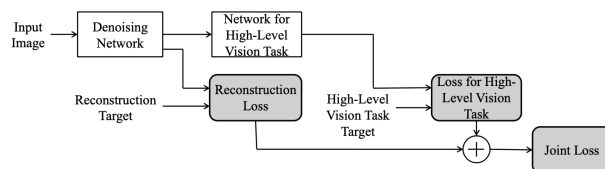


Figure 3: Overview of our proposed cascaded network.

- **Loss:**
重建损失使用MSE损失。
高级视觉任务使用CE损失。

▼ Experiment

- **Dataset:**
- **Train:** 高级视觉任务的网络使用一个在无噪声情况下训练好的网络，然后固定参数，用共同损失更新去噪网络的参数。

▼ (CVPR 2018) Recovering Realistic Texture in Image Super-resolution by Deep Spatial Feature Transform

[\[1804.02815\] Recovering Realistic T...](#)

Xintao Wang¹ Ke Yu¹ Chao Dong² Chen Change Loy¹
¹CUHK - SenseTime Joint Lab, The Chinese University of Hong Kong, ²SenseTime Research

- **SFTGAN:** 一种超分领域GAN-based-method

▼ Motivation

- 图像的纹理信息恢复的不好，通过改进损失函数有所改善，但还是存在问题。
- perceptual loss和adversarial loss带来的细节和原图的语义关联度不够，不能反映真实的类别。实验说明，先验可以很大程度影响模型的output，因此语义先验是有作用的。
- 为每个类别训练一个超分模型显然不可能。如果简单地把语义图和中间的feature map concat起来又没法充分发挥语义图的作用

▼ Contribution

- 提出了一种Spatial feature transform模块(SFT)
节约参数。
即插即用，很容易与现有模型结合。
可扩展，prior不仅仅可以是语义图，也可以是深度图等等。

▼ Method

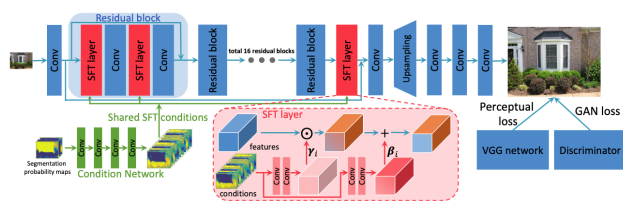


Figure 3. The proposed SFT layers can be conveniently applied to existing SR networks. All SFT layers share a condition network. The role of the condition network is to generate intermediate conditions from the prior, and broadcast the conditions to all SFT layers for further generation of modulation parameters.

- condition network：把语义图转换为条件图，共享到所有的SFT块。
- SFT块：通过条件图映射为一对参数，对特征进行变换。
- Loss：
 - 用预训练的VGG输出特征图计算perceptual loss，用minimax 计算对抗损失。

▼ Experiment

- Dataset：预训练使用Imagenet，去掉30k以下的图片，取了450k训练图片。
fine-tune使用从搜索引擎收集的，训练10324，测试300。
- Train：
- results：数量分析，还有user study

code

[GitHub - xinntao/SFTGAN: CVPR18 -...](#)

▼ 思考

- 也是使用语义先验，类别有八个，但是还是不算多。没有提到语义图分割不对情况下的副作用。

▼ (CVPR 2021) GLEAN: Generative Latent Bank for Large-Factor Image Super-Resolution

■

Kelvin C.K. Chan¹ Xintao Wang² Xiangyu Xu¹ Jinwei Gu³ Chen Change Loy^{1*}
¹Nanyang Technological University, Singapore
²Applied Research Center, Tencent PGC ³SenseBrain

- Motivation
- Contribution
- Method
- Experiment
 - Dataset：

- Train:

▼ (ECCV 2020) Blind Face Restoration via Deep Multi-scale Component Dictionaries

🔗 [2008.00418] Blind Face Restoration...

▪

Xiaoming Li^{1,4,6} [0000-0003-3844-9308], Chaofeng Chen^{2,4} [0000-0001-6137-5162],
Shangchen Zhou³ [0000-0001-8201-8877], Xianhui Lin⁴ [0000-0002-8974-2064],
Wangmeng Zuo^{1,2,5} [0000-0002-3330-783X], and Lei Zhang^{4,6} [0000-0002-2078-4215]

▼ Motivation

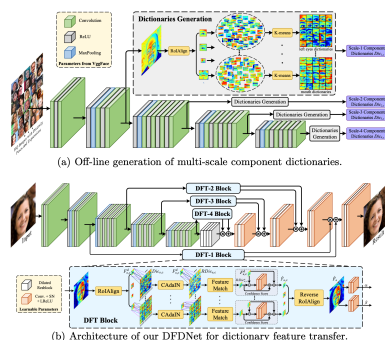
- blind image restoration对于恢复真实LQ图像有很强的实用价值
- 缺陷：其中两种方法，1. 需要一个来自同样人物的正面HQ图像。2. HQ和LQ中姿势和表情不同会影响重建的效果

▼ Contribution

- 使用一个五官部位字典作为参考备选，指导人脸重建。不需要对应的HQ图像
- 提出一个DFT block，使用CAdaIN消除输入和字典的风格差异，使用 confidence score控制融合程度。
- 渐进式方法训练

▼ Method

▪



- Off-line 多尺度部位字典生成：VggNet作为特征提取器，在多个尺度上，先使用RoIAlign裁剪出四种部位，之后使用K-means分为多个堆，作为字典。
- ▼ 采用同样的预训练VggNet作为encoder，保证同样的特征空间。
在DFT block中，首先通过RoIAlign，产生四个部位区域。假设input的部位和字典里的风格不一样，所以使用CAdaIN进行正则化，之后在Feature Match（用内积算相似度）在字典中选取相似的类。之后预测一个confidence score，在融合时更好的提供互补信息。最后Reverse RoIAlign把增强后的特征贴到相同位置。借用SFT块思想预测一对参数，输入到decoder。

- CAdaIN

$$RDic_{s,c}^k = \sigma(F_{s,c}^d) \left(Dic_{s,c}^k - \mu(Dic_{s,c}^k) \right) + \mu(F_{s,c}^d)$$

$$F_{s,c}^d = F_{s,c}^d + \frac{1}{\sigma(Dic_{s,c}^*)} \left(\sigma(Dic_{s,c}^*) - F_{s,c}^d \right)$$

- Confidence score

$$\hat{F}_{s,c} = F_{s,c}^d + RDic_{s,c}^* * \mathcal{F}_{Conf}(RDic_{s,c}^* - F_{s,c}^d; \theta_C),$$

- Loss:
reconstruction loss: 包含MSE和perceptual loss
adversarial loss

▼ Experiment

- Dataset:
FFHQ: 10000张图像用来建立字典
VggFace2: 训练
测试: 2000VggFace2, 2000CelebA
- Train:

- code

[GitHub - csxmli2016/DFDNet: Blind F...](#)

▼ 思考

- 不是完全的语义分割先验，只是将面部的一些部位进行裁剪并着重处理，整体结构比较复杂，需要先建立一个字典才能实现。
和有前脸参考的方法相比，本方法不需要成对的HQ图像，但是由于字典的限制，面部细节只能着重处理左右眼、鼻子、嘴巴这四个部位，还是有一定的提升空间

▪ 分支主题 5