



# 8月论文分享



 分享人：王雅恬



中山大學

SUN YAT-SEN UNIVERSITY

# BLNet: A Fast Deep Learning Framework for Low-Light Image Enhancement with Noise Removal and Color Restoration

IEEE



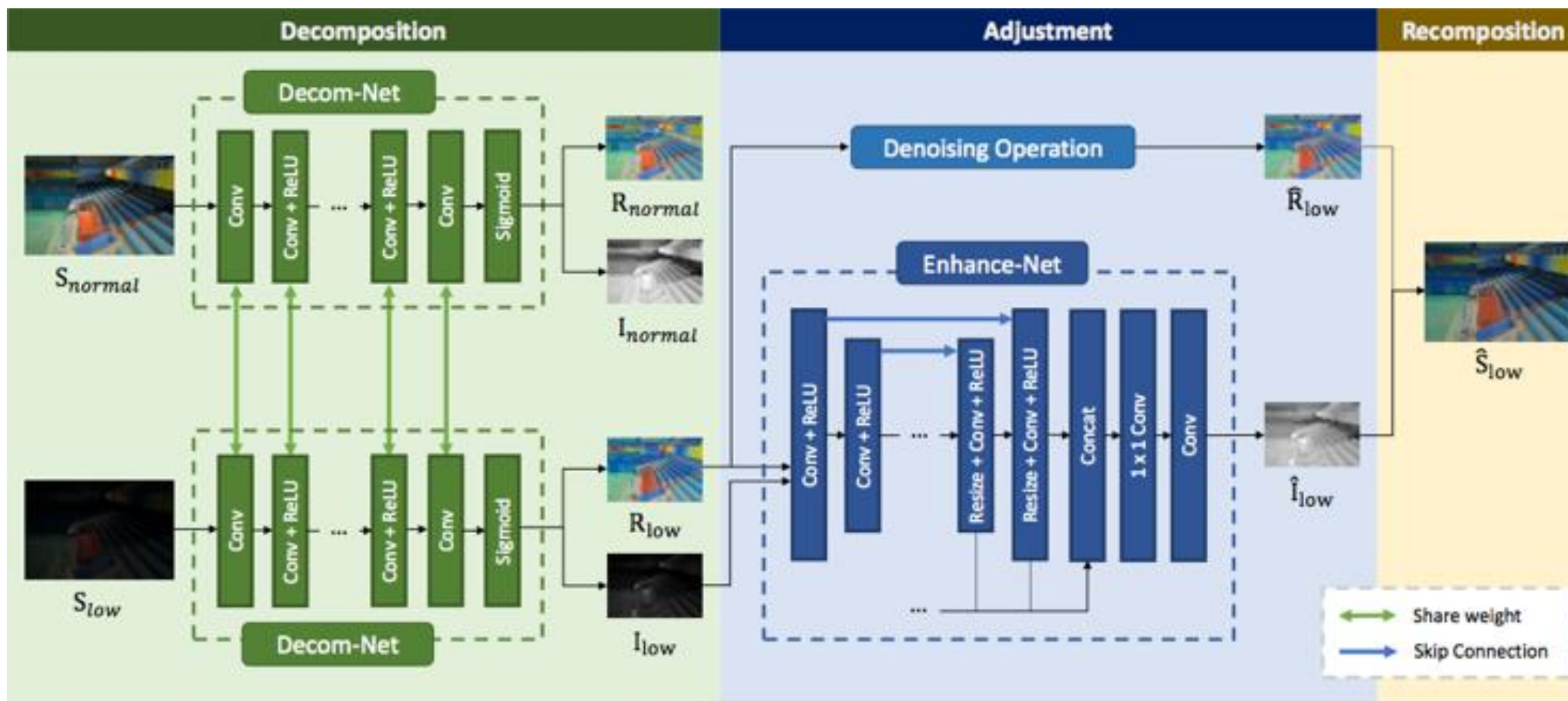
## Retinex理论

- 一张图像 $S$ ，可以被分解为照明层 $I$ 和反射层 $R$ 两个部分，其中照明层为单通道图像，反射层为3通道图像。

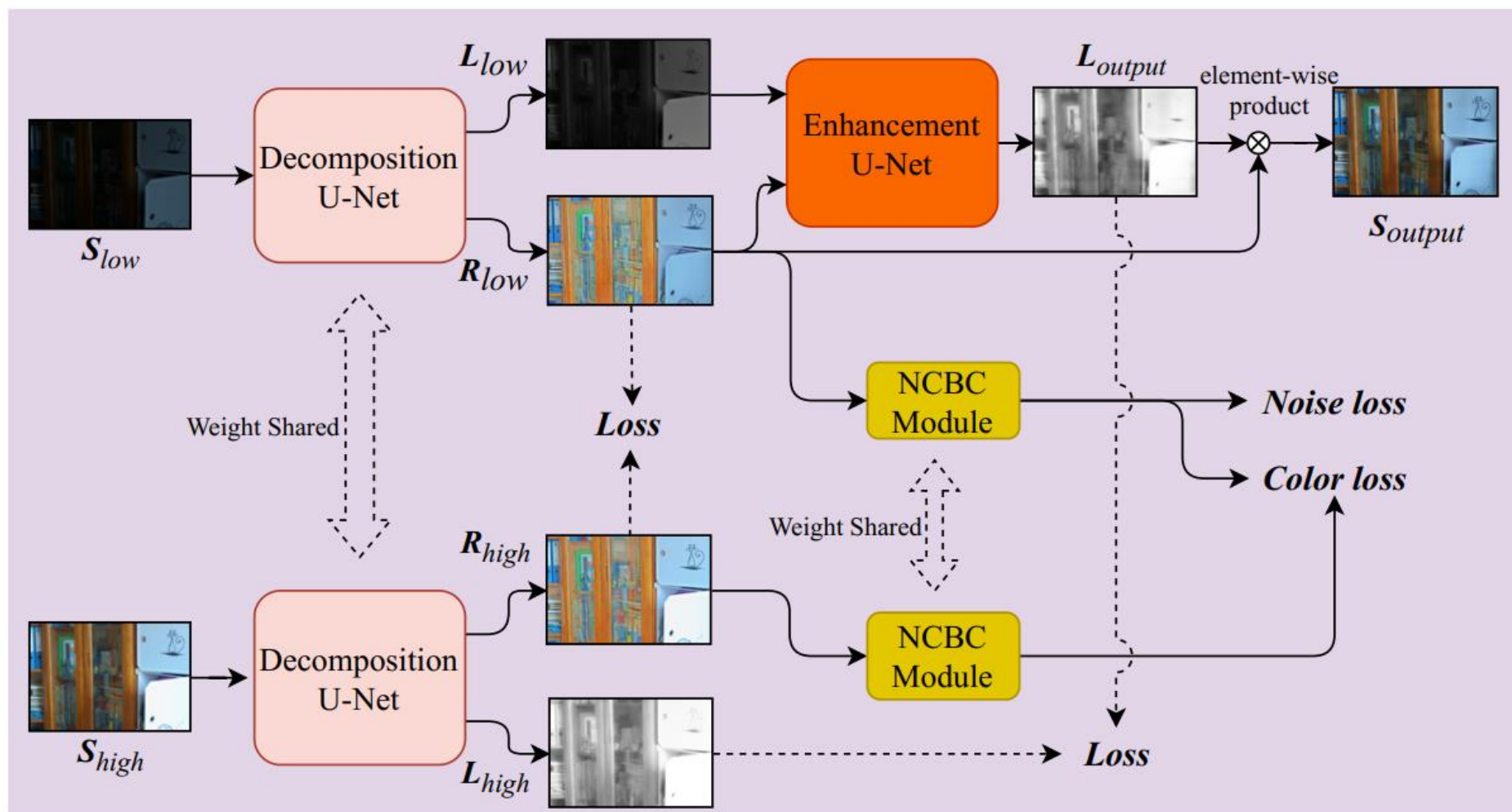
$$S = I \circ R$$

- 照明层 $I$ ：低频信息，亮度的强度和分布。
- 反射层 $R$ ：高频信息，语义信息如场景、纹理、细节等。

## RetinexNet (来自BMVC 2018)

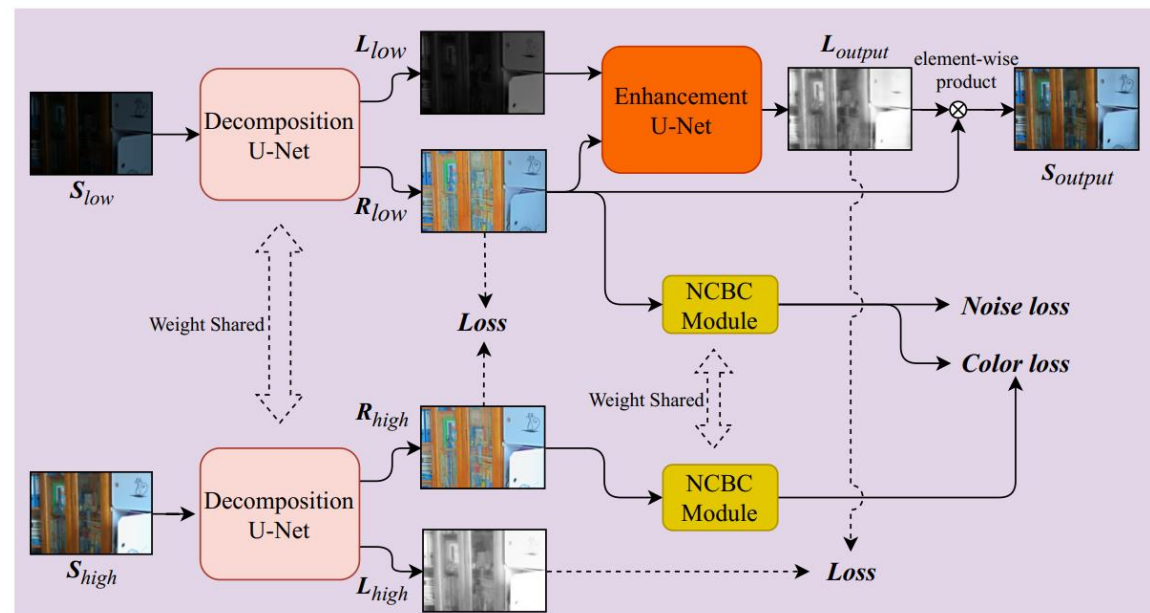


## 网络总览



## 网络总览

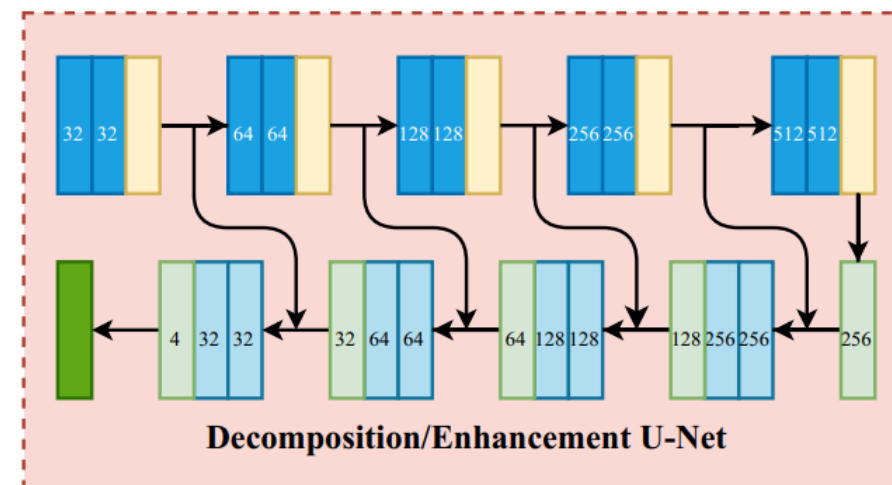
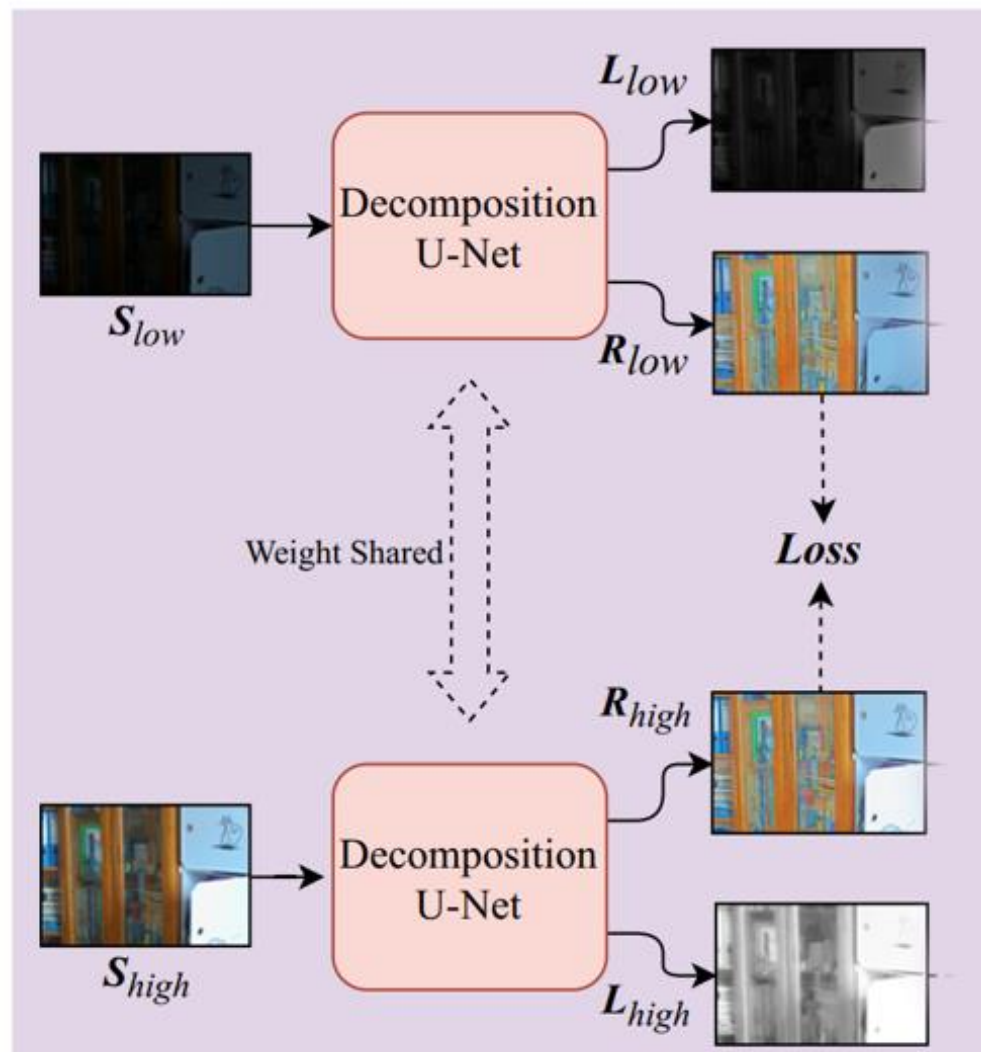
- 分解阶段：用U-Net将图像分解为照明层和反射层。此处，提出NCBC（Noise and Color Bias Control Module）模块，抑制反射层中的噪声，同时控制色调平衡。
- 增强阶段：用一个结构与分解阶段极其相似的U-Net。
- 通过精心设计的损失函数，用两个U-Net就能同时达到亮度增强、去除噪声、恢复色调、保留细节的目标。同时解决弱光增强领域中遇到的种种不足之处。





## 分解网络的设计

- RetinexNet: 用一个简单的卷积网络, 会放大噪声和造成色调失真。
- KinD: 用一个浅层的U-Net, 但反射层的噪声还是被放大了, 仍有严重色调失真。
- 论文方案: 用一个更深层的U-Net, 原因是, U-Net的这种先下采样再上采样的结构有去噪的功能, 尽量避免噪声被放大。
- 同时, 引入NCBC模块来解决反射层的噪声和色调失真。
- 受上面两个方法的影响, 用了RetinexNet提出的重建损失reconstruction loss, 来约束分解过程。





## 分解阶段的损失函数 $L_{decom} = 1.0L_{rc} + 0.1L_{smooth} + 0.01L_{equal}$

➤ 重建损失  $L_{rc} = \sum_i \sum_j \lambda_{ij} \|R_i * L_j - S_j\|_1$  (弱光和正常图像间权重取1, 否则取0.001)

➤ 约束弱光和正常光分解出的反射层  $L_{equal} = \|R_{low} - R_{high}\|_1$

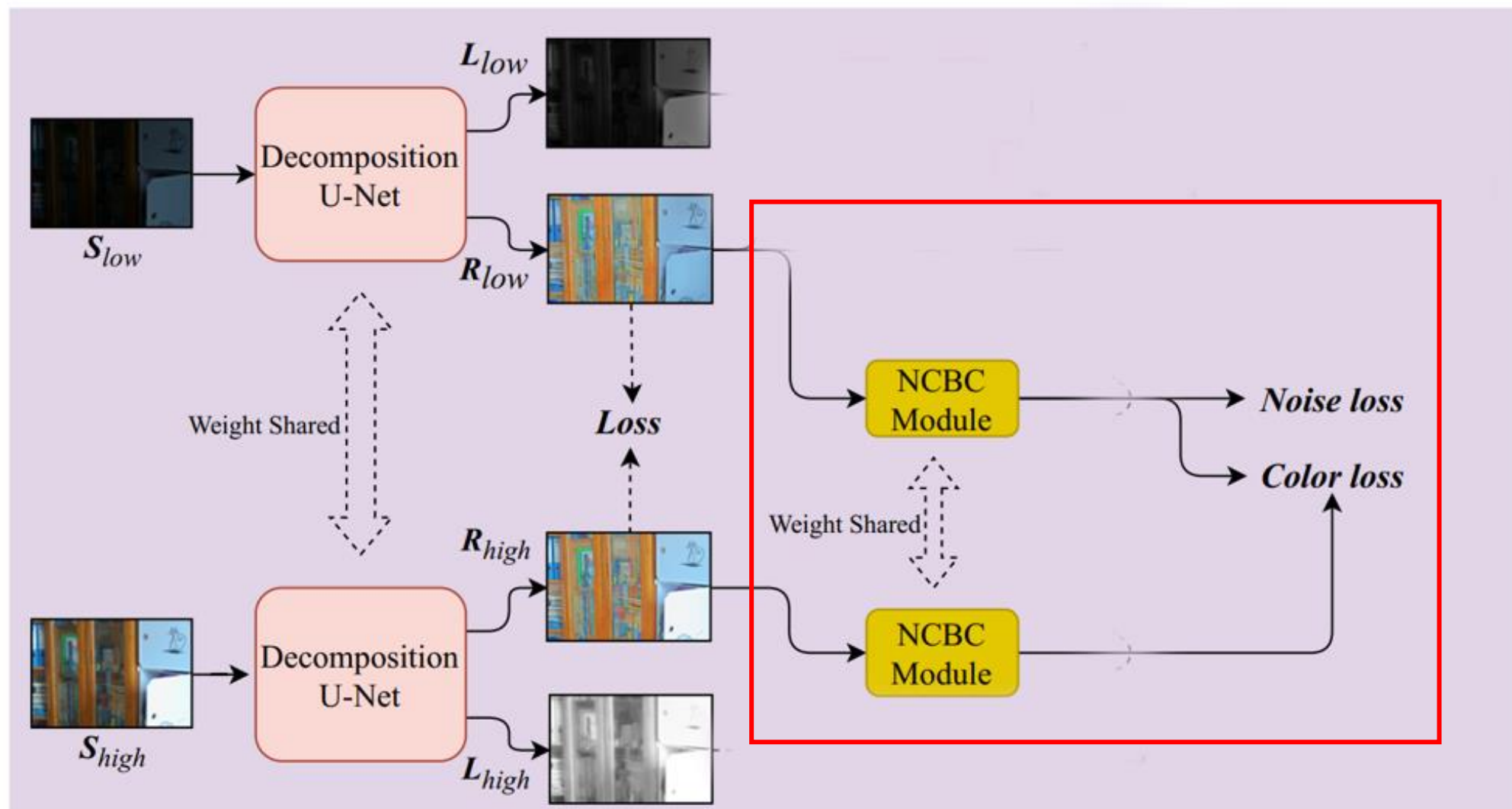
➤ 以往的工作中, 用平滑损失来约束照明层, 使其只含有低频信息例如光照的分布和强度, 所有的高频信息如细节和纹理, 都保留在反射层中。本方法也用了RetinexNet提出的平滑损失, 来使分解出的照明层光滑。

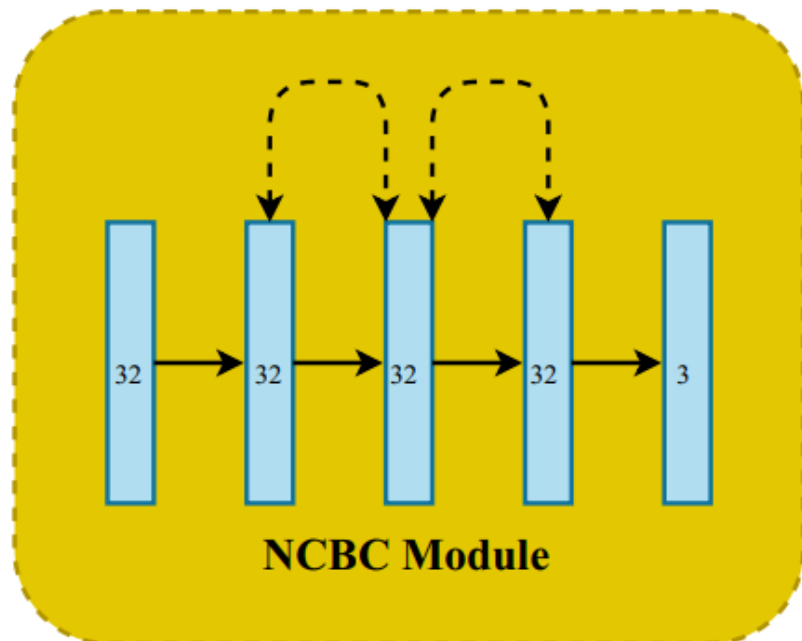
$$L_{smooth} = \sum_i \left\| \nabla_h L_i * e^{-\lambda * \nabla_h R_i} \right\| + \sum_i \left\| \nabla_v L_i * e^{-\lambda * \nabla_v R_i} \right\| \quad \text{(弱光和正常图像分别计算)}$$

- 按照以往的工作：此时，照明层足够光滑、只含低频信息，分解后产生的噪声都只出现在反射层。
- 但是，本论文认为，可以使照明层含有一些高频信息，在经过增强后被恢复出来！

## 改进之处：NCBC模块

- 以往的工作用TV损失（Total variance loss）来平滑反射层。 $L_{TV}^{low} = \|\nabla_h \phi(R_{low})\|_2^2 + \|\nabla_v \phi(R_{low})\|_2^2$
- 通过实验，论文发现：直接用TV损失来平滑反射层并不利于去噪，还可能引起图像失真。为了不引用额外的去噪网络（导致整个网络处理速度变慢，且在反射层引起色调失真），论文提出NCBC模块，来抑制噪声、控制反射层色调。
- 同时，通过此模块的损失设计（noise loss），被TV loss平滑掉的反射层中的细节和噪声，都被转移到了照明层上。因此，不像其他方法，我们的照明层也包含了一些高频信息，比如细节和纹理。这样一来，我们可以把细节保留在照明层中，这意味着，在经过增强网络后，细节能够被其修复和重建。





包含：一个简单CNN，和两个损失函数

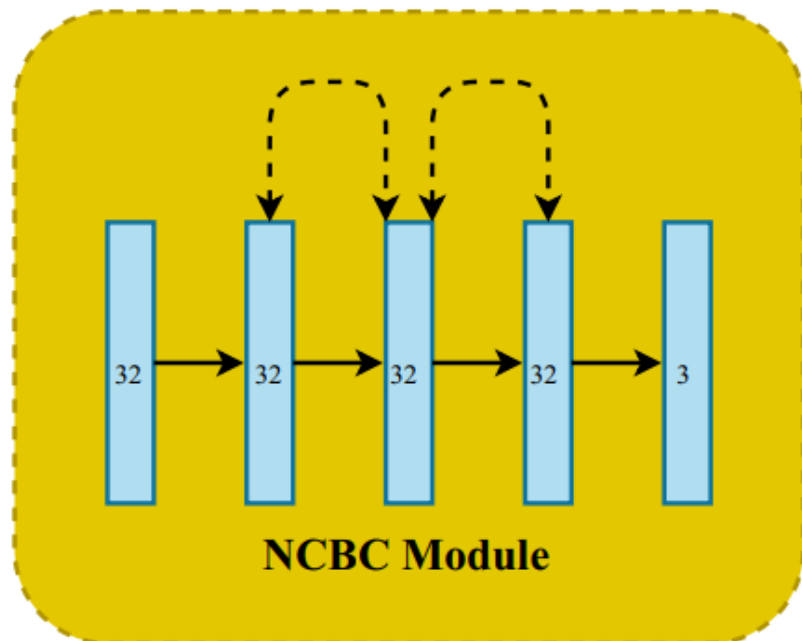
输入：  $R_{low}$  和  $R_{high}$

输出：参与损失项的计算，用  $\phi$  表示

$$L_{TV}^{low} = \|\nabla_h \phi(R_{low})\|_2^2 + \|\nabla_v \phi(R_{low})\|_2^2 \quad (\text{平滑反射层})$$

$$L_{MSE} = \|R_{low} - R_{high}\|_2^2 \quad (\text{使弱光域的反射层尽量接近正常光})$$

$$L_{noise} = 0.05L_{TV}^{low} + 1.0L_{MSE} \quad (\text{抑制反射层噪声})$$



包含：一个简单CNN，和两个损失函数

输入：  $R_{low}$  和  $R_{high}$

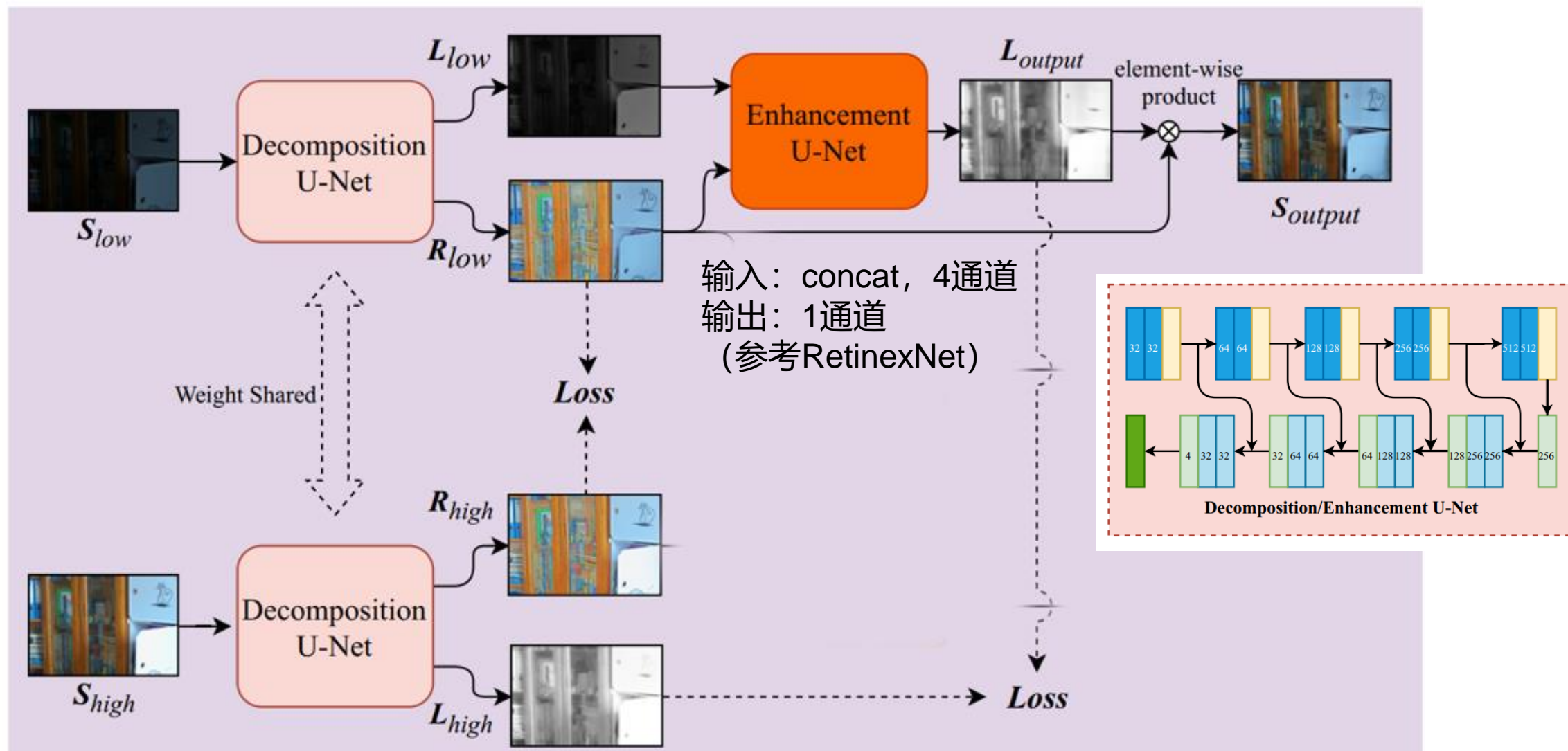
输出：参与损失项的计算，用  $\phi$  表示

$$L_{color} = \|\phi(R_{low}) - \phi(R_{high})\|_1$$

(使低光域反射层的全局信息尽量接近正常光，以保证色调不失真)

模块总损失函数：  $L_{NCBC} = 0.2L_{noise} + 0.1L_{color}$





## 增强阶段的损失函数

➤结构：与分解网络极其相似的U-Net。

➤效果主要依赖于损失函数。  $L_{eh} = 1.0L_{rc} + 1.0L_{bri} + 1.0L_{per} + 1.0L_{grad}$

$$L_{rc} = \|R_{low} * L_{output} - S_{high}\|_1 \quad (\text{重建损失})$$

$$L_{bri} = \|L_{output} - L_{high}\|_1$$

$$L_{per} = \frac{1}{CHW} \|\phi(R_{low} * L_{output}) - \phi(S_{high})\|_2^2 \quad (\text{用感知损失, 保持纹理信息的不变, } \phi \text{ 代表VGG16提取})$$

$$L_{grad} = \|\nabla_h(R_{low} * L_{output}) - \nabla_h(S_{high})\|_1 + \|\nabla_v(R_{low} * L_{output}) - \nabla_v(S_{high})\|_1 \quad (\text{保持锐度和平滑度之间的平衡, 不要太过平滑})$$

## 实验细节

- 数据集：LOL训练集和验证集
- 还在几个被广泛使用的数据集上进行了测试：DICM、LIME、MEF
- 量化评估指标：PSNR、SSIM、LPIPS、FSIM、UQI
- 用 Angular Error 和 DeltaE 作为色调失真的计算指标

$$AngularError = \arccos\left(\frac{\langle S_{output}, S_{high} \rangle}{\|S_{output}\| \cdot \|S_{high}\|}\right)$$

## 消融实验

➤ NCBC模块可以  
增强亮度、  
抑制噪声、纠  
正色偏、除去  
光晕伪影

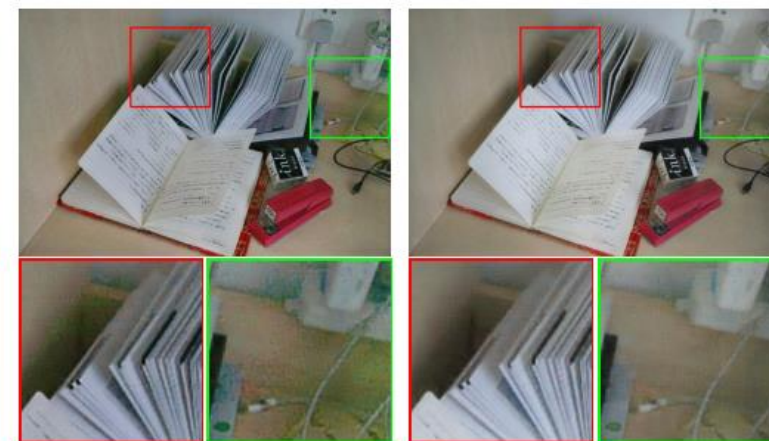


(a) RetinexNet [13](17.78/0.49/0.83) (b) Ours w/o NCBC(17.70/0.83/0.94)



(c) Ours with NCBC(22.06/0.87/0.62) (d) GroundTruth

Fig. 3. Ablation study on the LOL dataset (PSNR/SSIM/Color Bias Average).



(a) Our model without NCBC (1.16/1.14/1.15) (b) Our model with NCBC (1.03/0.99/1.01)



(c) Our model without NCBC (1.73/1.87/1.80) (d) Our model with NCBC (0.86/1.24/1.05)

Fig. 4. Ablation study on LOL dataset (color bias mean/color bias median/color bias average).



## 消融实验

➤调整NCBC模块中的noise loss的权重

$$L_{noise} = 0.05L_{TV}^{low} + 1.0L_{MSE}$$

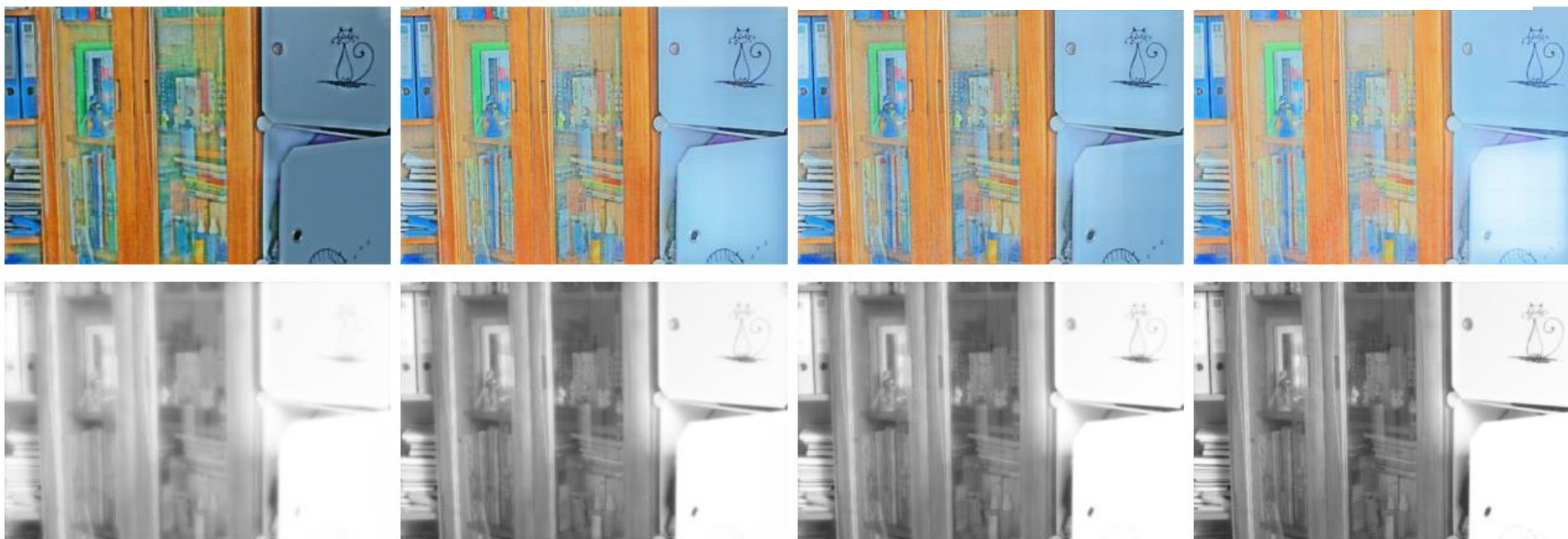


Fig. 5. The images in the first and second row from left to right represent the low-light image's reflectance and GroundTruth's illumination (low-light image's illumination is too dark to compare) decomposed by our decomposition net without the NCBC Module, using noise loss with weight coefficients of 0.2, 0.4 and 0.7.

从左到右，随着noise loss权重增大，反射层趋于光滑，照明层趋于尖锐

# 和SOTA的量化评估

有监督指标:

Traditional Methods	PSNR↑	SSIM↑	LPIPS↓	FSIM↑	UQI↑
Input	7.7733	0.1914	0.4173	0.7190	0.0622
MSRCR [11]	13.1728	0.4615	0.4404	0.8450	0.7884
BIMEF [35]	13.8752	0.5949	0.3673	0.9263	0.7088
LIME [12]	16.7586	0.4449	0.4183	0.8549	0.8805
Dong [36]	16.7165	0.4783	0.4226	0.8886	0.8078
SRIE [34]	11.8552	0.4954	0.3657	0.9085	0.5033
MF [33]	16.9662	0.5075	0.4092	0.9236	0.8572
NPE [32]	16.9697	0.4839	0.4156	0.8964	0.8943
RRM [19]	13.8765	0.6636	0.3476	0.8821	0.7275
LECARM [38]	14.4099	0.5448	0.3687	0.9288	0.6406
JED [20]	13.6857	0.6509	0.3549	0.8812	0.7143
PLM [61]	16.2620	0.4617	0.4284	0.8265	0.8892
DIE [39]	14.0181	0.5188	0.3910	0.9172	0.7027
DL Methods	PSNR↑	SSIM↑	LPIPS↓	FSIM↑	UQI↑
MBLLEN [18]	17.8583	0.7247	0.3672	0.9262	0.8261
RetinexNet [13]	16.7740	0.4249	0.4670	0.8642	0.9110
GLAD [14]	19.7182	0.6820	0.3994	0.9329	0.9204
RDGAN [16]	15.9363	0.6357	0.3985	0.9276	0.8296
Zero-DCE [17]	14.8671	0.5623	0.3852	0.9276	0.7205
Zhang [48]	19.4968	0.7003	0.3911	0.8514	0.8521
EnlightenGan [15]	17.4828	0.6515	0.3903	0.9226	0.8499
Our model w/o NCBC	18.4446	0.7605	0.3514	0.9261	0.9259
Our model	20.1447	0.7918	0.3126	0.9454	0.9371

Angular Error, 评估色调:

Traditional Methods	Mean↓	Median↓	Average↓	DeltaE↓
MSRCR [11]	3.7421	4.5877	4.1649	27.4496
BIMEF [35]	3.4004	3.5187	3.4595	33.8820
LIME [12]	3.2096	4.0825	3.6460	21.1816
Dong [36]	3.3499	4.1481	3.7490	25.3349
SRIE [34]	3.4488	4.0751	3.7620	44.3194
MF [33]	3.3277	3.9810	3.6544	24.5488
NPE [32]	3.5588	4.2505	3.9046	22.6374
RRM [19]	3.3745	3.5821	3.4784	32.9843
LECARM [38]	3.4091	3.9668	3.6979	34.5143
JED [20]	3.4064	3.8651	3.6357	33.8342
PLM [61]	3.4274	3.7085	3.4829	22.0553
DIE [39]	3.4597	4.1164	3.7880	34.1583
Deep Learning Methods	Mean↓	Median↓	Average↓	DeltaE↓
MBLLEN [18]	3.2716	4.4620	3.8669	21.5774
RetinexNet [13]	3.7501	4.4975	4.3589	21.3550
GLAD [14]	3.3110	3.8021	3.5565	16.0393
RDGAN [16]	4.3899	5.3027	4.8463	26.3796
Zero-DCE [17]	4.1051	4.6860	4.3955	31.4451
Zhang [48]	3.3744	3.9104	3.6424	17.6652
EnlightenGan [15]	4.5296	5.2536	4.8916	21.9113
Our model w/o NCBC	2.8645	3.9083	3.3864	16.0604
Our model	2.1785	2.3870	2.3024	13.5264

Gray Entropy (GE) 和 Color Entropy (CE), 评估图像信息量:

Deep Learning Methods	GE↑	CE↑	Time cost↓
Low-Light Input	4.7432	14.3837	-
GroundTruth	7.0400	21.3163	-
MBLLEN [18]	7.1421	21.3664	80 ms
RetinexNet [13]	6.8346	21.1266	20 ms
GLAD [14]	7.1141	21.5237	25 ms
RDGAN [16]	6.6327	20.1053	30 ms
Zero-DCE [17]	6.5964	18.9072	2 ms
Zhang [48]	7.0673	21.4085	20 ms
EnlightenGan [15]	7.0664	21.3476	20 ms
Our model w/o NCBC	7.0336	21.3651	7 ms
Our model	7.1450	21.5882	7 ms



## 和SOTA的量化评估

本方法在耗时上存在优势，原因在于：

1. 网络结构简单，分解和增强都是U-Net结构，计算速度快
2. 去噪和色调调整的步骤由损失函数进行，没有使用额外的网络，在测试时不耗费额外时间

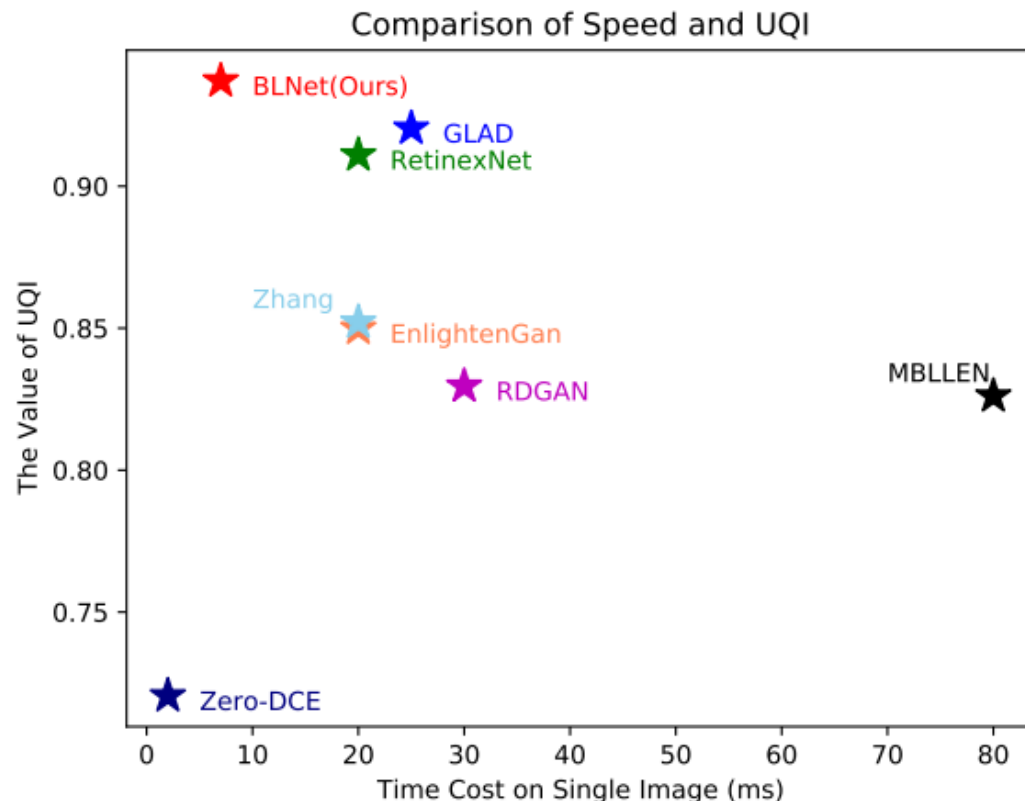
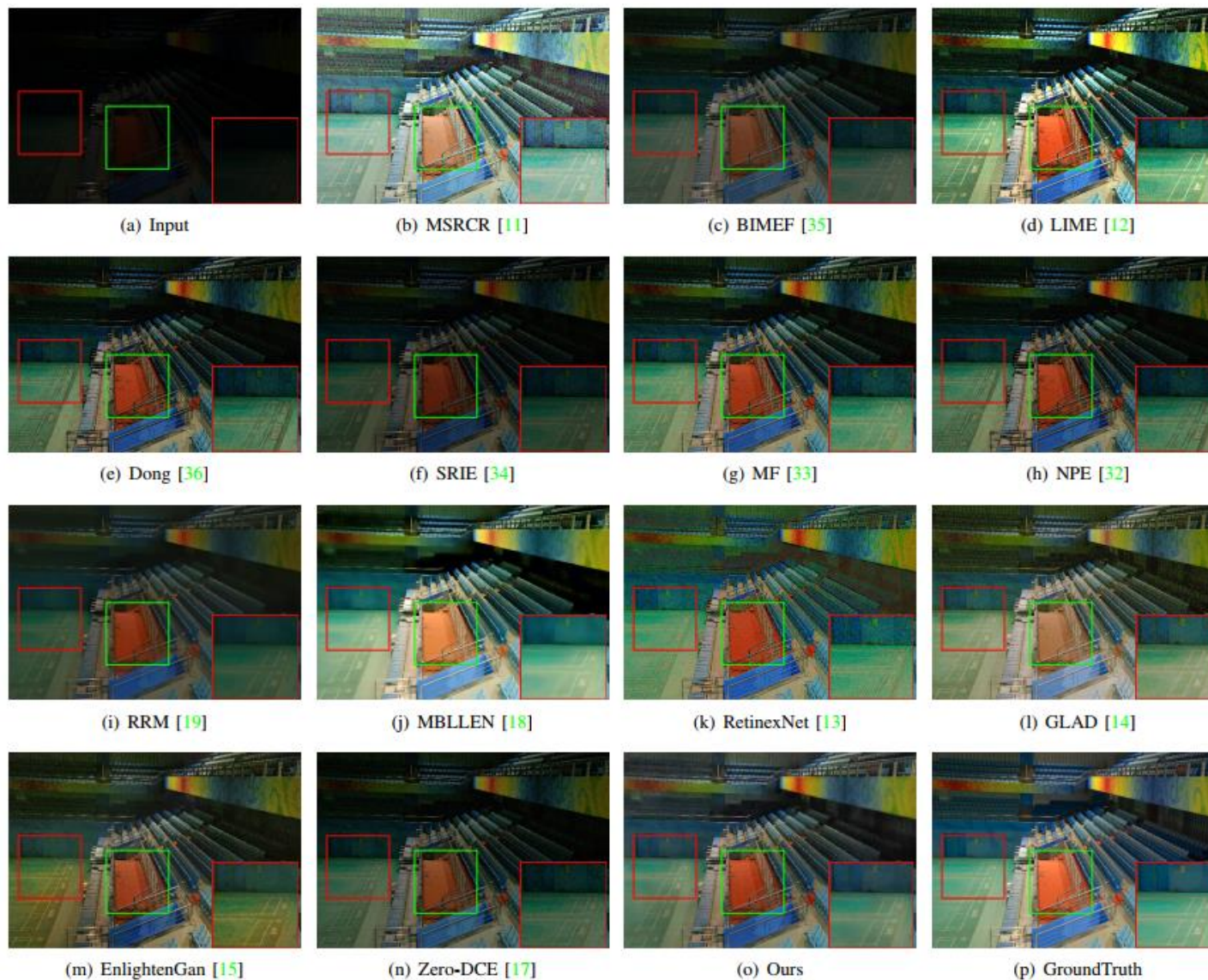


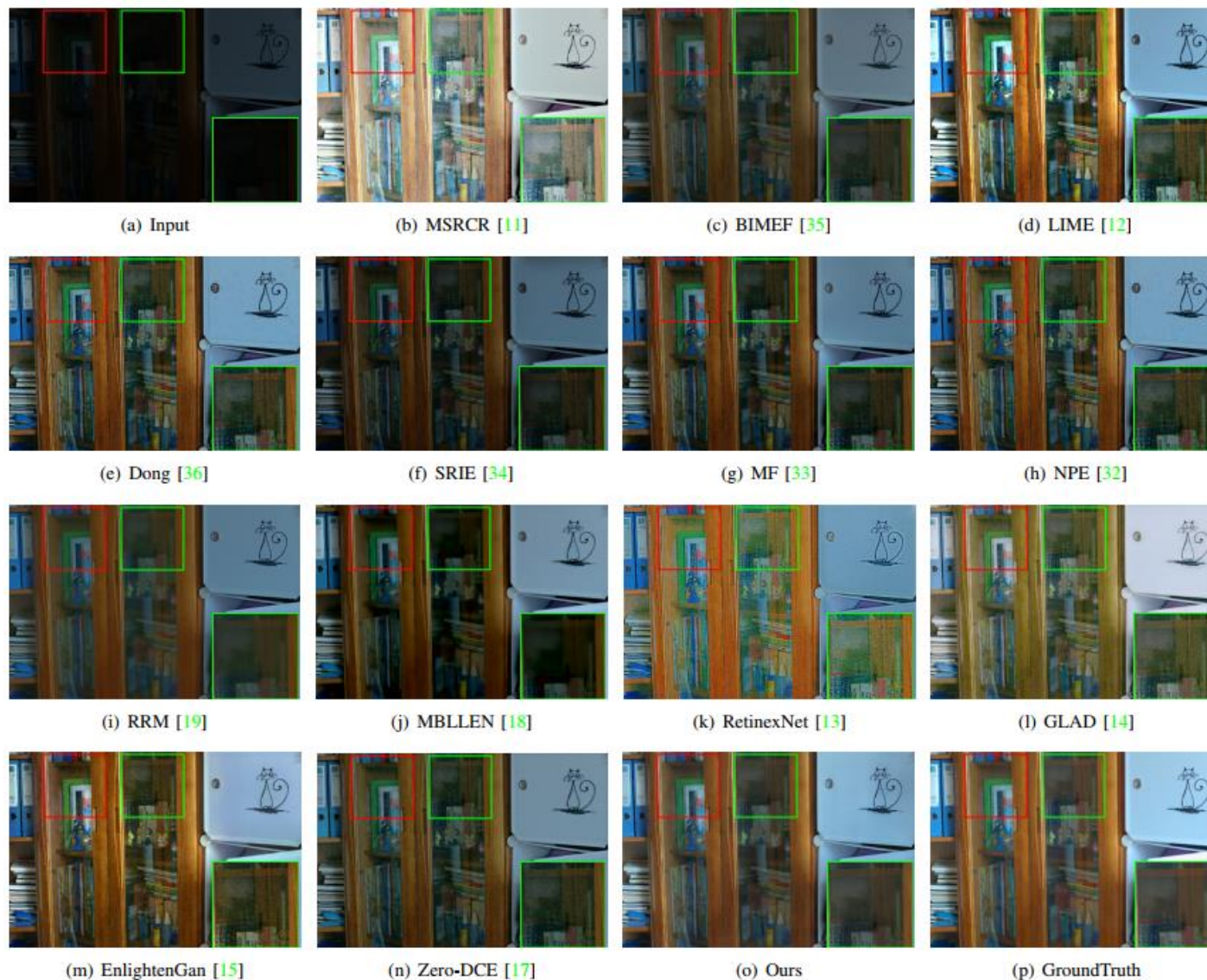
Fig. 11. Runtime cost and performance comparison of our method and other state-of-the-art deep learning methods on the LOL dataset.

## 和SOTA的视觉评估





## 和SOTA的视觉评估



## 不足之处

- 如果图像中存在过暗区域，仍会出现光晕伪影和局部阴影



Fig. 12. Some failure cases of our results.

## 工作亮点

- 在RetinexNet的工作上进行了改进，解决了前者的不足，效果有明显提升。
- 提出NCBC模块，用损失函数替代专门的网络，在保证结果质量的同时降低处理耗时。
- 实验结果展示比较全面，在量化指标中，用Angular Error来评估色调，用GE和CE来评估图像信息量，以展示算法在这两个层面上的效果。



中山大學

SUN YAT-SEN UNIVERSITY

# Learning to Enhance Low-Light Images with a Synthetic-Real Interaction Training Strategy

ICDH 2020





## 背景

- 现存的深度网络忽视了网络强大的描写能力。它们倾向于设计复杂的网络结构，包含大量参数，定义多种损失函数。这导致了参数的冗余和计算上的压力。
- 为了解决这个问题，我们研究了一个新颖的训练策略，去挖掘网络强大的描写能力。这篇工作是领域内首次尝试这个思路的工作。

## 网络结构

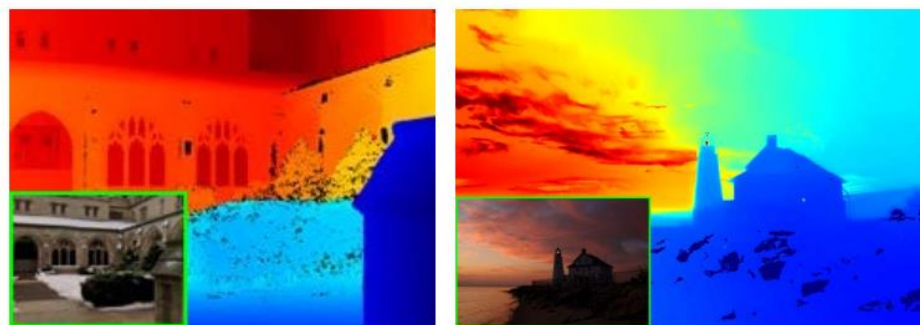
- 设计了一个具有强大描写能力的网络，同时具有更少的参数。
- 参考Retinex理论，用这个网络从图像中得到其照明层，再相除得到其反射层，将其反射层作为最终增强结果。
- 含26个卷积层的残差网络，文章详细描述了结构设计。

## Synthetic-Real 迭代训练策略

- 广受认可的一个观点是，训练数据的分布决定了深度学习的表现。
- 但是，如何对训练数据去建立训练的流程，也是同等重要的。
- 现存的弱光增强工作，都考虑通过一个单步的训练pipeline去训练参数。一个广泛使用的训练策略是，直接去使用真实的成对数据集，比如MIT-Adobe 5K dataset，学习一个从弱光到正常光间的映射是困难的，因此给网络结构和损失函数的设计带来许多阻碍，导致设计出来的网络含有很大的参数量。
- 但是，我们设计的网络是轻量的，因此不能通过单步训练pipeline获得理想结果。
- 因此，考虑用训练策略来达到满意的结果。

## Synthetic-Real 迭代训练策略

- 不仅用了真实图像对，也合成了成对数据。
- 考虑平滑度之间的联系，将深度作为照明层。此处用的是NYU Depth dataset。



The Depth Map

The Illumination Map

Figure 2: Visual comparison of the depth and illumination maps. The left bottom corner of each sub-figure demonstrates the input image.

## Synthetic-Real 迭代训练策略

- 用合成训练机制去引导生成平滑的照明层。合成训练阶段不是用来获取理想增强结果（即反射层）的，而是用来获取中间成分（即照明层）的。
- 之后，真实训练阶段是用来纠正照明层、获得理想增强结果的。

## 训练损失函数

➤简单的损失函数，只用到MSE。

$$\mathcal{L} = \begin{cases} \mathcal{L}_{MSE}(\mathbf{I}, \mathbf{I}_d), & \text{if synthetic,} \\ \mathcal{L}_{MSE}(\frac{\mathbf{L}}{\mathbf{I}}, \hat{\mathbf{R}}), & \text{if real,} \end{cases}$$

$\mathbf{I}$  : 弱光的照明层

$\mathbf{I}_d$ : 真实深度

$\mathbf{L}$ : 弱光输入图像

$\hat{\mathbf{R}}$ : 真实反射层 (指的可能是对应的正常光图像)

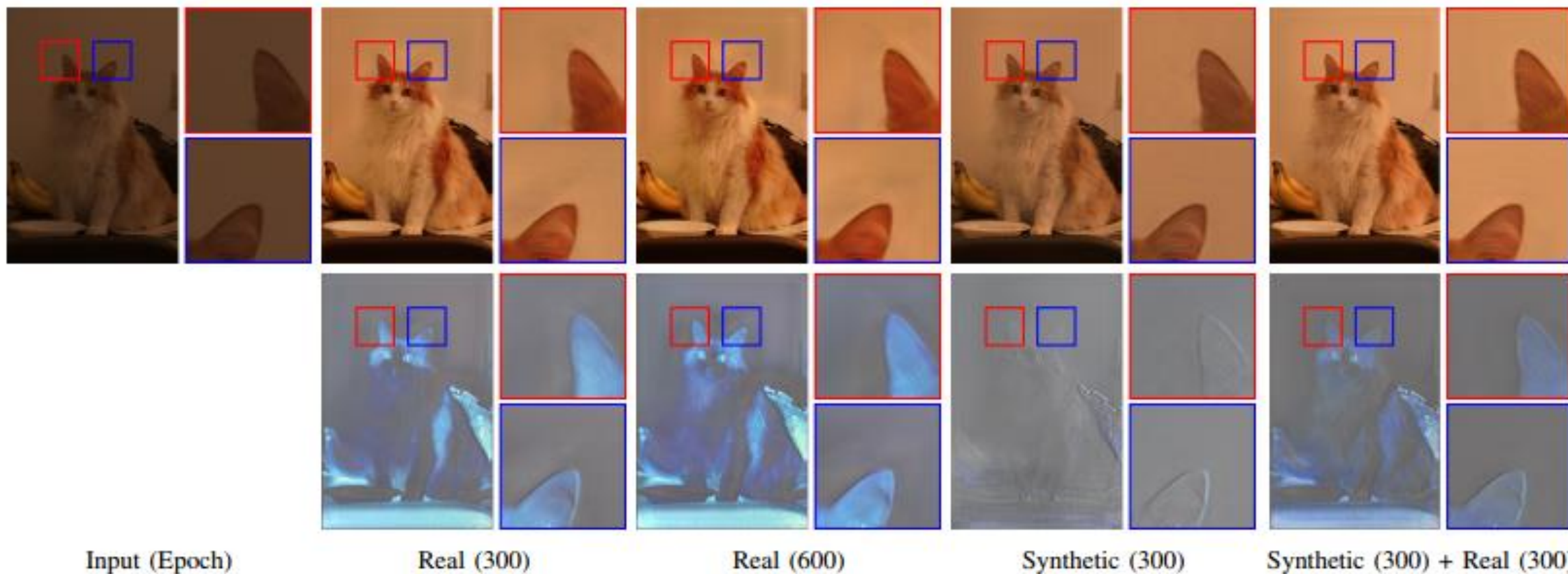
论文强调：使用的损失非常简单，无需任何正则化。  
与弱光增强领域的其他SOTA方法比，是突破性的。



- 论文的主要贡献在于这个 Synthetic-Real 迭代训练策略。
- 只用真实图像对训练，迭代了600个epoch，照明层和反射层仍存在伪影。因为网络和损失函数都太过简单，难以解决困难的任务。
- 只用合成图像对训练，可以避免伪影，但亮度和色调都不令人满意，因为合成图像和真实图像存在差异。
- 我们的 Synthetic-Real 训练策略，在相同的训练epoch下表现得很好，尤其是有效增强亮度、抑制伪影。

存在伪影

效果不真实



- 2次迭代的结果与1次迭代比，视觉效果和量化结果都提高了。
- (每个iteration含600个epoch)



PSNR/SSIM

Input

20.9986/0.8738

The 1st Interaction

22.7372/0.8990

The 2nd Interaction

## 效果展示

Metric	RetinexNet [7] (BMVC '18)	EnlightenGAN [10] (Arxiv '19)	DPE [13] (CVPR '18)	WhiteBox [14] (TOG '18)	DeepUPE [9] (CVPR '19)	Ours
PSNR/SSIM	11.9803/0.6836	15.1854/0.7460	15.9902/0.7754	20.0479/0.8346	20.8898/0.8391	<b>23.2041/0.8721</b>
Parameters # (Reduction)	2.1103M (57.75%)	8.6455M (89.69%)	3.3292M (73.22%)	25.6901M (96.53%)	2.9963M (70.25%)	<b>0.8915M</b>



Input

RetinexNet

EnlightenGAN

DPE

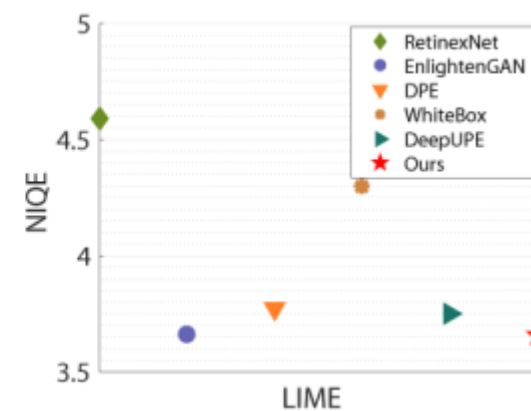
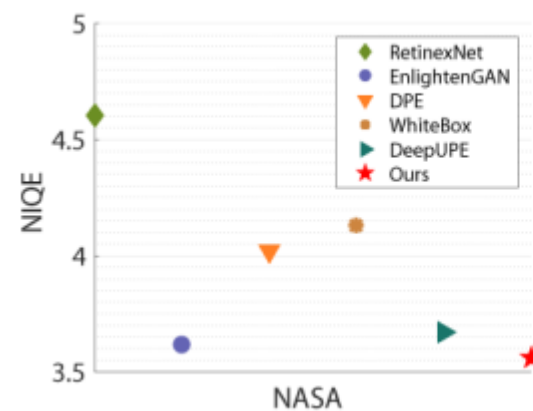
WhiteBox

DeepUPE

Ours



## 效果展示



- 亮点在于，提出用训练策略来弥补简单的网络结构，也能达到同等效果。
- 不足之处在于，没有对不同的网络结构进行实验对比，深入探究这个训练策略。
- 实验做得太过简单，采用的指标和测试数据集太少，说服力不够。

## 颜色损失

- 金老师说，可以思考为什么这里的色调调整方法取得了成功，它对比其他color loss有什么优势？
- 补充SOTA论文中用到的颜色损失（鸿俊在组会中提供）

Zero-DCE中的颜色损失：

$$L_{col} = \sum_{\forall (p,q) \in \varepsilon} (J_p^p - J_q^q)^2, \varepsilon = \{(R, G), (R, B), (G, B)\}$$

Deep-UPe中的颜色损失：

$$\mathcal{L}_c^i = \sum_p \angle((\mathcal{F}(I_i))_p, (\tilde{I}_i)_p),$$



➤用深度图去替代照明层的理论是错误的。金老师说，它们除了看上去都是光滑的，实际上没有任何相关之处。

Synthetic-Real 迭代训练策略

➤不仅用了真实图像对，也合成了成对数据。

➤考虑平滑度之间的联系，将深度作为照明层。此处用的是NYU Depth dataset。

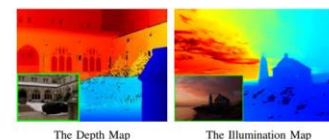


Figure 2: Visual comparison of the depth and illumination maps. The left bottom corner of each sub-figure demonstrates the input image.

➤但是，这种先达成一个目标，再达成另一个目标的做法，是可取的。

$$\mathcal{L} = \begin{cases} \mathcal{L}_{MSE}(\mathbf{I}, \mathbf{I}_d), & \text{if synthetic,} \\ \mathcal{L}_{MSE}(\frac{\mathbf{L}}{\mathbf{I}}, \hat{\mathbf{R}}), & \text{if real,} \end{cases}$$