CTF-Net: A Coarse-To-Fine Network For Low Light Image Enhancement

Fengji Ma[a,1], Junyi Chai[b,2], Haitao Li[a,3] and Jinping Sun [a,*,4]

[a] School of Electronic and Information Engineering, Beihang University, Beijing 100081, China
[b] School of Electrical and Computer Engineering, Purdue University, West Lafayette 47906, the United States

ARTICLE INFO

Keywords:
Low light image enhancement
Coarse to fine
Two-step image enhancement
Retinex theory

ABSTRACT

Conventional low-light enhancement algorithms often fail to adequately enhance detail information or avoid color distortion. In this paper, we propose a novel coarse-to-fine low-light image enhancement algorithm, which formulates the process as a two-step decomposition and refinement problem. Our algorithm contains two steps: coarse enhancement and refinement. The first physical prior based coarse enhancement module (CEM), consisting of illumination estimation network (IEN) and noise estimation network (NEN), is constructed following the robust retinex decomposition to eliminate noise hidden in dark regions and remove the dark illumination map to obtain the first-step coarse enhancement result. The second feature-based detail refinement module (DRM), which consists of several feature fusion detail enhancement units (FFDE). The module fine-tunes the detail information and brightness. By utilizing multi-scale spatial information fusion, the output of each deconvolution layer is connected in the expansion path to enhance the image details, optimize the brightness (contrast) of the image, and restore the high-quality clear image. Extensive experiments on various benchmarks demonstrate the advantages of our method over state-of-the-art methods both qualitatively and quantitatively. We will release our code upon acceptance.

1. Introduction

Images taken in low-light environment suffer from severe quality degradation, which significantly affects downstream computer vison tasks, including semantic instance segmentation, face detection, object tracking, etc. Recent advances in image enhancement have greatly promoted the fields of high-level computer vision tasks that require images of high visibility.

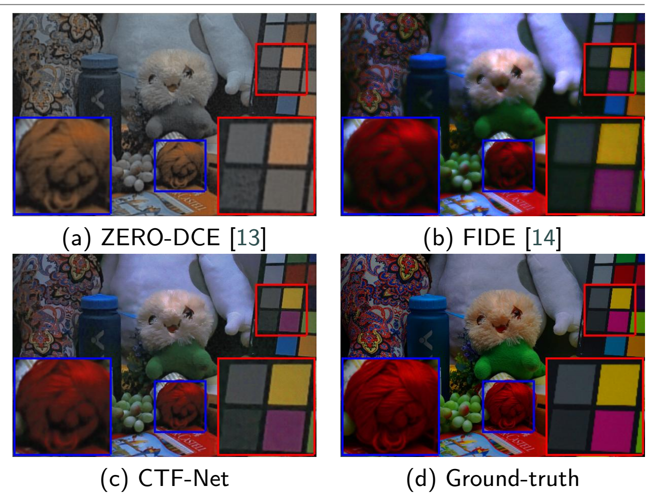
Conventionally, image enhancement has been achieved by histogram adjustment [1, 2] and Retinex theory [3] based illumination estimation [4–11]. These methods, however, do not take into consideration intensive noise hidden in the dark regions, and the noise could be amplified in the output image. Correspondingly, [12] proposes the robust retinex theory with an extra noise term:

$$\mathbf{I} = \mathbf{R} \circ \mathbf{T} + \mathbf{N} \tag{1}$$

where $\circ$ represents element-wise product. $\mathbf{I}$, $\mathbf{R}$, $\mathbf{T}$ and $\mathbf{N}$ represent input, reflectance map, illumination map and noise.

With the development of deep learning in recent years, learning-based low-light image enhancement algorithms emerge in endlessly. Deep-learning based low-light enhancement is generally divided into two parts: the one-step, or end-to-end, and the two-step. In one-step methods, the low-light images pass through the network, and the network outputs the final normal-light images. Two-step methods utilize two different networks for different levels of adjustment.

An end-to-end multi-branch enhancement network is proposed by [15], which uses feature extraction module, enhancement module and fusion module to obtain final enhanced images. An encoder-decoder network is proposed by [16], which extracts both local and global features to approximate mapping from input image to illumination, and uses the reflectance map as enhancement result. EnlightenGAN [17] utilizes the generator with attention-guided U-Net [18], and uses global-local discriminator to enhance the low-light image. ZERO-DCE [13] utilizes deep curve estimation network for non-linear estimation of low-light image to obtain high-order parameter curve, and the enhancement



(a) ZERO-DCE [13] (b) FIDE [14]
(c) CTF-Net (d) Ground-truth

Fig. 1: A typical example of three types of low light image enhancement method. (a) ZERO-DCE (One-step enhancement method). (b) FIDE (Two-step enhancement method). (c) CTF-Net (Coarse-to-fine enhancement method). (d) Ground-truth. The proposed CTF-Net achieves visually pleasing result in terms of brightness (contrast), color, and image detail.

*Corresponding author
mafengji@buaa.edu.cn (F. Ma); chai28@purdue.edu (J. Chai); 15864587215@163.com (H. Li); sunjinping@buaa.edu.cn (J.S.)

is performed based on the parameter curve. A reflectance estimation network [19] is proposed to directly learn to map input image to its corresponding reflectance map.

The two-step Retinex-based enhancement network is first proposed by [20]. The first part is decomposition network for Retinex decomposition, and the second step is the enhancement network which adjusts the illumination map for low-light enhancement. Similarly, works including [21, 22] also follow the decomposition-enhancement procedure. EEMEFN [23] proposes to decompose the enhancement task into multi-exposure fusion and edge enhancement. A frequency-based two-step network [14] is proposed where the first part learns to recover low-frequency targets, and the second part adjusts high-frequency detail information according to the low-frequency targets.

In conclusion, one-step methods perform end-to-end enhancement with carefully designed network, while two-step methods decompose the enhancement tasks into several subtasks. Inspired by previous work, we propose a coarse-to-fine two-step low-light image enhancement algorithm. Our work is motivated by:

- Pre-Experiment results, which validates the hypothesis that the two-step coarse-to-fine network works better than the one-step networks and most two-step methods of adjusting the illumination map, and that under the same trainable network parameters, the results from the coarse-to-fine network are significantly better than those of one-step networks and two-step networks that adjusts the illumination map.

- An intuitive fact that conventional Retinex-based methods often result in distorted output image. As shown in Figure 1, one-step enhancement method fails to recover color information, while two-step method based on illumination adjustment results in over-exposure and loss of detail information, and all of these shortcomings significantly affects the quality of output image. Thus, a straightforward idea is to add an extra refinement network after obtaining coarse enhancement results to fine-tune local brightness and detail information. As will be stated as follows, the refinement network allows us to reduce the size of coarse enhancement network, while still achieving good performance.

We design a novel CTF-Net, which decomposes the enhancement task into two sub-tasks. In the first step, we design a lightweight network to pre-enhance the input image based on physical prior. Compared with feature-prone refinement network, the coarse network effectively improves image contrast, and reduces the model size. In the second step, we perform feature-based detail refinement module (DRM), which is consists of several feature fusion detail enhancement units (FFDE). Every unit utilizes spatial information and pixel-level relationship to fine-tune the coarse result, and the final result is obtained through fine adjustment of image brightness and detail enhancement.

In summary, we make the following contributions:

Table 1
Pre-Experiment results on LOL dataset [20]. The best result in terms of PSNR is highlighted in bold.

Metric	One-step	Two-step	Coarse-to-fine
PSNR	16.2532	16.4856	**16.9548**

- For the first time, we propose to decompose the enhancement task into Retinex-based coarse enhancement and feature-based refinement, and design a coarse-to-fine network (CTF-Net) with two subnetworks for the two subtasks respectively.

- Our CTF-Net contains both unsupervised coarse network for Retinex decomposition and supervised multi-level feature-based refinement network. Our coarse network is simple but effective, and for refinement network we effectively combine spatial and pixel-level information to perform brightness adjustment and detail enhancement.

- We show from experiments that our method performs significantly better than other SOTA methods on both full-reference and non-reference datasets.

2. Pre-Experiemnt

Our pre-Experiment compares several different enhancement networks and adjusts them to the same network size. The performance of various enhancement models (one-step network model, two-step network model adjusting illuminance map and two-step coarse-to-fine network proposed in this paper) is compared horizontally. All networks are trained on 2080Ti GPU. The learning rate is set as 10^{-4}. The batch size is set to 4. We use Adam optimizer [24] with learning rate 1e-3, where β_1 and β_2 take the defaulte values of 0.9 and 0.999, respectively. We employ a paired dataset, which is called as LOLdataset [20] to train our pre-expriment.

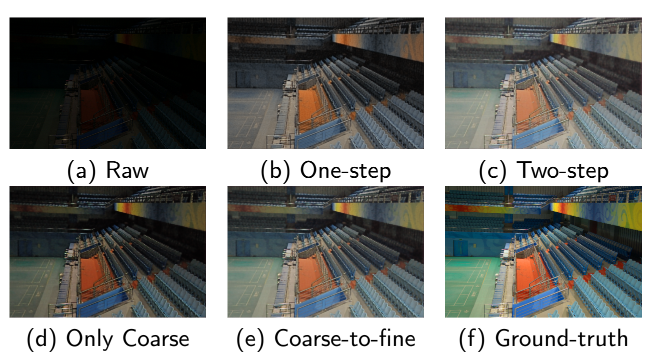


Fig. 2: Visual comparison on LOLdataset [20]. (a) shows the raw input. (b)-(e) show the results of One-step, Two-step, only coarse and our proposed coarse-to-fine. (f) shows the ground-truth.

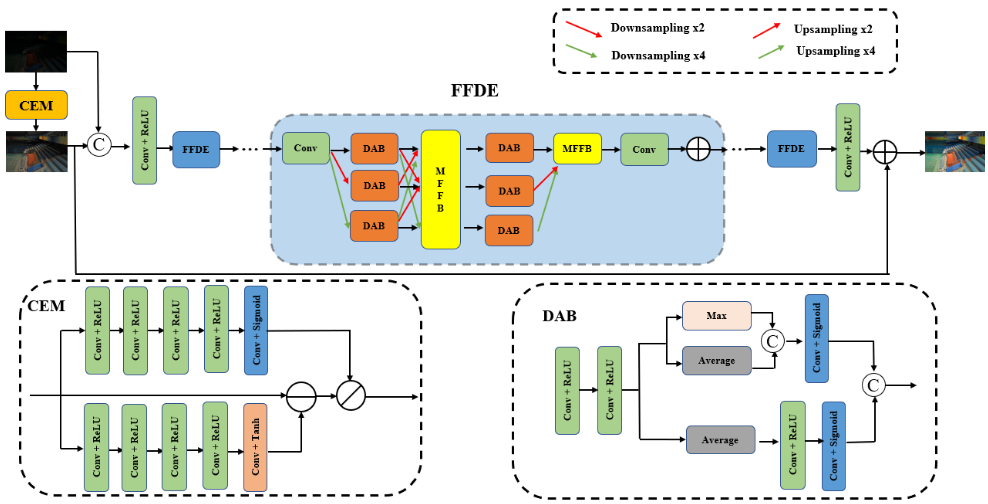


Fig. 3: The architecture of our proposed Coarse-to-fine Network (CTF-Net).

We can see from experimental results in Table 1 that under the same scale, the coarse-to-fine network outperforms one-step network by 0.7016dB (=16.9548-16.2532), and outperforms two-step methods by 0.4692dB (=16.9548-16.4856). Also, among all the enhancement results in Figure 2, the coarse-to-fine method achieves best visual result, and only coarse result fails to properly adjust the brightness of low-light image.

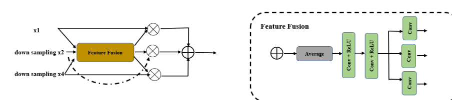


Fig. 4: The architecture of multi-level feature fusion block (MFFB).

3. Method

In this section, we introduce our coarse-to-fine model and the objective function. As shown in Figure 3, we use coarse network for pre-enhancement, and the intermediate result is denoted as $\mathbf{I}_{coarse}$. In the second part, we extract the illumination map of low-light image with illumination estimation network, reduce noise in low-light image with noise estimation network, and use fine enhancement module to fine-tune the intermediate $\mathbf{I}_{coarse}$ to obtain final $\mathbf{I}_{fine}$.

Our coarse model follows robust Retinex theory with stacked multi-branch CNN structure for illumination and noise estimation. We notice that illumination map is very smooth, reflecting the intensity and distribution of brightness, while the reflectance map contains most high-frequency information. We make the assumption that the reflectance map can be treated as a coarse enhancement result, with certain extent of color distortion and loss of detail induced by removing noise and illumination map. To overcome such drawbacks, we utilize a refinement module to fine-tune the coarse result.

3.1. Coarse Enhancement Module (CEM)

Most low-light enhancement algorithms are based on Retinex algorithm, including RetinexNet [20], KinD[21] and R2RNet[22], which adjust the decomposed illumination map and combine it with reflectance map to get final enhancement result. One common problem regarding these methods is: decomposing input image to get illumination map and adjusting illumination map requires different networks, which greatly increases computational complexity. Besides, since the illumination map is very smooth, adjusting the illumination map may not sufficiently utilize the detail information. Inspired by DeepUPE and DSR, we discard the extra illumination adjustment; instead, we merge this process into illumination estimation and regularize the illumination and noise estimation with brightness control function over coarse enhancement result and use reflectance map as final coarse result. The detail of coarse enhancement module (CEM) is as follows:

$$\hat{\mathbf{T}} = f_{IEN}(\mathbf{I}) \tag{2}$$

where $\hat{\mathbf{T}}$ is the illumination map extracted by illumination estimation network (IEN), which is composed of one pooling payer, three convolutional layer with ReLU activation and one convolutional layer with Sigmoid activation as shown in Figure 3.

$$\hat{\mathbf{N}} = f_{NEN}(\mathbf{I}) \tag{3}$$

where $\hat{\mathbf{N}}$ is the noise extracted by noise estimation network (IEN), which is composed of three convolutional layer with ReLU activation and one convolutional layer with Tanh

Table 2
Quantitative performance comparison on different datasets with image quality metrics PSNR/SSIM.

Dataset	LIME	JIEP	RetinexNet	EnlightenGAN	DSR	FIDE
LOL	12.0606/0.5433	13.3208/0.5947	14.0158/0.5705	14.7164/0.5986	17.7959/0.6943	16.5923/0.6149
SICE	13.7134/0.5412	13.5401/0.5708	13.6132/0.5413	15.0584 /0.6092	16.1874/0.6282	14.6158/0.5421
MIT-Adobe FiveK	11.9876/0.5078	12.0587/0.5768	13.1567/0.5964	13.9875/0.6047	15.5472/0.6872	15.9687/0.6087
Dataset	MBLLEN	DeepUPE	R2RNet	ZERO-DCE	RUAS	Ours
LOL	18.1807/0.7291	14.5923/0.5422	18.8675/0.7854	15.4328/0.6305	16.4368/0.6817	**20.6854/0.8608**
SICE	17.0606/0.5433	13.5401/0.5708	17.6158/0.6505	14.7164/0.5986	16.7959/0.6643	**18.5923/0.7149**
MIT-Adobe FiveK	17.5672/0.7021	15.6871/0.6578	17.8821/0.6874	15.6547/0.6378	16.1879/0.6915	**20.0635/0.8471**

activation as shown in Figure 3.

$$\mathbf{I}_{coarse} = \left(\mathbf{I} - \hat{\mathbf{N}}\right) \oslash \left(\hat{\mathbf{T}} + \varepsilon\right) \tag{4}$$

where $\oslash$ represents element-wise division, and ε is set as 10^{-8} to avoid zero denominator.

3.2. Detail Refinement Module (DRM)

Compared with DeepUPE[16] and DSR[19] which treat reflectance map as final result, we use an extra refinement module to deal with loss of detail and color distortion, which are inevitable during the estimation process. Fine-enhancement module aims at refine the coarse intermediate result. During the first step, our coarse network simply removes noise and illumination map based on Retinex theory, failing to consider detail and color information. Thus in the second step, we design a feature-based detail refinement network, including both brightness refinement adjustment and detail enhancement.

The mechanism of adjusting receptive fields can be incorporated in CNN by using multi-scale feature fusion. The detail refinement module (DRM) is composed of several feature fusion detail enhancement units (FFDE), each of which is composed of convolution layers,dual attention block (DAB) and multi-level feature fusion block (MFFB).

DRM learns enriched feature representations for image restoration and enhancement. DRM is based on a recursive residual design. In the core of DRM is the feature fusion detail enhancement uint (FFDE) whose main idea is dedicated to maintaining spatially-precise high-resolution representations through the entire network and the complimentary set of parallel branches provide better contextualized features. It also allows multi-level feature fusion block (MFFB) in order to consolidate the highresolution features with the help of low-resolution features, and vice versa. As illustrated in Figure 4, MFFB generates global feature descriptors are ultilized to recalibrate the feature maps (of different streams) followed by their aggregation.

While FFDE fuses information across multi resolution branches, we also need a mechanism to share information within a feature tensor, both along the spatial and the channel dimensions. Motivated by the advances of recent low-level vision methods based on the attention mechanisms , we propose the dual attention block (DAB) to extract features in the convolutional streams. The schematic of DAB is shown in Figure 3. The DAB suppresses less useful features and only allows more informative ones to pass further. This feature recalibration is achieved by using channel attention and spatial attention mechanisms.

3.3. Loss Functions

$$L = L_{coarse} + L_{fine} \tag{5}$$

where L represents the total loss function, L_{coarse} represents the loss function of physical prior based coarse enhancement module (CEM), and L_{fine} represents the loss function of detail refinement module (DRM). L_{coarse} contains three parts, L_{TV}, L_{noise} and L_{exp}:

$$L_{coarse} = \lambda_{TV} L_{TV} + \lambda_{noise} L_{noise} + \lambda_{exp} L_{exp} \tag{6}$$

$$L_{TV} = \frac{1}{H \times W} \left\| \nabla \hat{\mathbf{T}} \right\|_1 \tag{7}$$

$$L_{noise} = \frac{1}{H \times W} \left\| \hat{\mathbf{N}} \right\|_F^2 \tag{8}$$

where $\|\cdot\|_1$ represents ℓ_1 norm, and $\|\cdot\|_F$ represents Frobenius norm. H and W represent height and width of input image.

$$L_{exp} = \frac{1}{H \times W} \left\| \mathbf{I}_{coarse} - \theta_{exp} \right\|_1 \tag{9}$$

where θ_{exp} represents brightness adjustment parameter to control the local average brightness of the coarse enhancement result and is set as 0.6. $\mathbf{I}_{coarse}$ represents the result of coarse enhancement module (CEM).

$$L_{fine} = \lambda_{rec} L_{rec} + \lambda_{VGG} L_{VGG} + \lambda_{detail} L_{detail} \tag{10}$$

$$L_{rec} = \frac{1}{H \times W} \left\| \mathbf{I}_{fine} - \mathbf{I}_{gt} \right\|_1 \tag{11}$$

where $\mathbf{I}_{fine}$ represents the result of detail refinement module (DRM). $\mathbf{I}_{gt}$ indicates the ground-truth of the input image.

$$L_{VGG} = \frac{1}{H \times W} \left\| \phi\left(\mathbf{I}_{fine}\right) - \phi\left(\mathbf{I}_{gt}\right) \right\|_1 \tag{12}$$

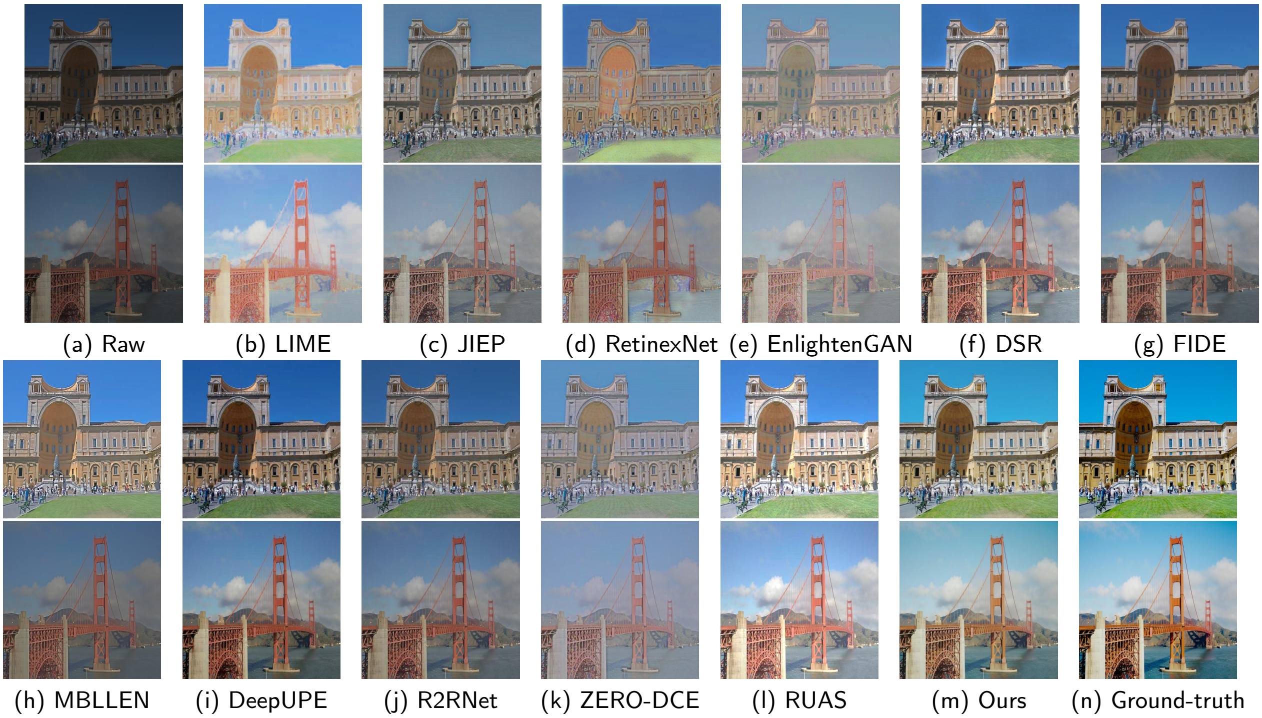


Fig. 5: Qualitative comparisons on MIT-Adobe FiveK dataset [28]. (a) raw inputs. (b) - (l) SOTA methods. (m) the results of our proposed method CTF-Net. (n) the Ground-truth.

where ϕ indicates the feature map obtained by VGG Network [25]. We use the well-behaved VGG Network as the content extractor [26]. Color distribution is hidden in higher-level dimension feature which can be extracted by pre-trained VGG-16 network. Meanwhile, the higher-level dimension feature consist brightness refinement information, which can guide the coarse enhancement to a better refinement result.

$$L_{detail} = 1 - SSIM(\mathbf{I}_{fine}, \mathbf{I}_{gt}) \tag{13}$$

where L_{detail} is introduced to preserve the proper image detail and avoid blurring or sharpening. We use the well-known image quality assessment algorithm SSIM [27] to preserve image detail.

4. Experiments

4.1. Implementation Details

We have implemented the proposed model in the Pytorch framework. The proposed method CTF-net is trained on an NVIDIA 2080Ti GPU. The batch size is set to 4. We use Adam optimizer [24] with learning rate 10^{-4}, where β_1 and β_2 take the defaulte values of 0.9 and 0.999, respectively. We employ a large-scale dataset, which is called as LSRW dataset [22] to train the proposed CTF-Net. During training, we randomly crop patches of resolutions 128×128. We train our method for 500 epochs. In this paper, we empirically set $\lambda_{TV}, \lambda_{noise}, \lambda_{exp} = \{0.1, 5000, 200\}$, $\lambda_{rec}, \lambda_{VGG}, \lambda_{detail} = \{1, 1, 2\}$.

4.2. Baselines

To evaluate the performance of the proposed method on enhancing low-light images, we quantitatively and visually compare our method to 11 state-of-the-art enhancement methods with available codes, including LIME [4], JIEP [5], Retinex-Net [20], EnlightenGAN [17], DSR [19], FIDE [14],MBLLEN [15], DeepUPE [16], R2RNet [22],ZERO-DCE [13] and RUAS [29].We use PSNR and SSIM for quantitative measurement.

4.3. Datasets and Metrics

The most commonly used low-light image datasets fall into two categories. The first category is full reference datasets, including LOL [20], SICE [30], MIT-Adobe FiveK [28]. SICE is a multi-exposure image dataset. Each sequence has 3 to 18 low-contrast images of different exposure levels. As for convience, we choose the first three images of each sequence in SICE dataset Part2 and resize them to 512×512. It is worth mentioning that in this paper, the results of expert C are used as the ground-truth of MIT-Adobe FiveK data set. In this paper, we resize the images in MIT-Adobe FiveK to 512×512. The second category is unreferenced data sets, including LIME [4], NPE [31], MEF [32], DICM [33], VV [1]. In this paper, we select four images from each of the above five non-reference datasets, and a total of 20 pictures form a mixed non-reference dataset (**MNR dataset**). For the non-reference datasets, we choose NIQE [34] and ARSIM [35] as measurement metrics. PSNR and SSIM [27] are used for the full-reference datasets.

[1] https://sites.google.com/site/vonikakis/datasets

Table 3
NIQE/ARISM scores on the MNR dataset. The best results are highlighted in bold. Lower NIQE and lower ARISM indicate better image quality.

Methods	LIME	JIEP	Retinex-Net	DeepUPE	ZERO-DCE	DSR	FIDE	CTF-Net
NIQE	3.9365	3.8876	4.1236	3.5678	3.6371	3.6812	3.5431	**3.4025**
ARSIM	2.8715	2.8261	2.8921	2.7862	2.7963	2.7986	2.7843	**2.7126**

Table 4
Ablation study on the loss fuctions. The best results in terms of PSNR/SSIM are highlighted in bold. w/o reprsents without.

Metrics	w/o L_{TV}	w/o L_{noise}	w/o L_{exp}	w/o L_{rec}	w/o L_{VGG}	w/o L_{detail}	Baseline
PSNR	20.0987	20.1583	19.8784	18.0165	19.7164	19.4959	**20.6854**
SSIM	0.7937	0.8405	0.8113	0.7463	0.7986	0.6943	**0.8608**

Table 5
Ablation study on the different modules. The best results in terms of PSNR/SSIM are highlighted in bold. w/o represents without. DRM represents dettail refinement module. CEM represents physical prior based coarse enhancement module.

CEM		✓	✓
DRM	✓		✓
PSNR	19.7542	16.3543	**20.6854**
SSIM	0.7693	0.6548	**0.8608**

4.4. Visual comparisons

The qualitative comparison is shown in Figure 5-7. On MIT-Adobe FiveK dataset, LIME, RUAS and EnlightenGAN leads to over-exposure, and the detail information is poorly preserved. DSR achieves relatively good contrast, however, the enhanced images suffer color distortion. Other methods lead to under-exposure. By contrast, our method better restore the color of the original image, and the brightness of both two images is suitably adjusted. On MNRF dataset, RetinexNet and DSR lead to over-exposure, and DSR shows an unnatural brightness contrast between foreground and background sky. On SICE dataset, the color and brightness of our enhancement result are closest to those of the ground-truth, while other methods suffer from over or under exposure and color distortion.

4.5. Quantitative comparisons

We move on to discuss the quantitative results. Table 1 shows quantitative performance of different methods on three datasets. Our method achieves best performance on all the three datasets, and significantly outperforms other SOTA methods regarding both PSNR and SSIM.

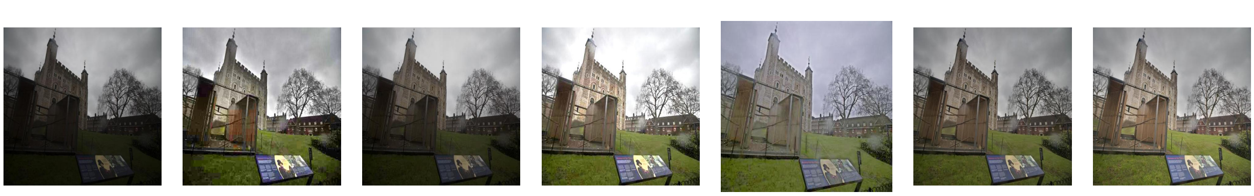
4.6. Ablation Study

We quantitatively evaluate the effectiveness of different modules and the loss functions setting in our CTF-Net model on the LOL dataset. The results are shown in Table 3 and 4.

- Contribution of Each Loss: We study the effect of each loss function by eliminating one loss term each time. As shown in Table 4, removing every loss term results in significant performance degradation, which validates the rationality of our loss function.

(a) Raw (b) RetinexNet (c) DSR (d) FIDE (e) DeepUPE (f) ZERO-DCE (g) Ours

Fig. 6: Qualitative comparisons on MNR dataset. (a) raw input. (b) - (f) SOTA methods. (g) results of our proposed method CTF-Net.



(a) Raw (b) DSR (d) FIDE (e) RUAS (c) ZERO-DCE (f) Ours (g) Ground-truth

Fig. 7: Qualitative comparisons on SICE dataset [30]. (a) raw input. (b) - (e) SOTA methods. (f) results of our proposed method CTF-Net. (g) the Ground-truth.

- Effectiveness of Each Module: As shown in Table 5, both CEM (coarse enhancement module) and DRM (detail refinement module) have a positive role in the experimental results, and removing either part will make the final results worse. Especially, the DRM greatly improves the image enhancement results with its efficiency in detail enhancement, brightness adjustment and color correction.

5. Conclusion

In this paper we discuss the superiority of coarse-to-fine methods over conventional one-step image enhancement methods and illumination adjustment two-step image enhancement methods. For the first time, we propose a two-step coarse-to-fine network to decompose the enhancement task into two subtasks: brightness adjustment and refinement. For the first task, we construct an unsupervised Retinex-based network to remove noise and illumination map from raw input image, and for the second task, we construct a multi-level feature-based supervised network, which adjusts brightness and color of the coarse result, and perform detail enhancement for better visual effect. Experiments on several datasets show that our method significantly outperform SOTA methods regarding all the metrics.

Acknowledgment

This research was funded by the National Natural Science Foundation of China, Grant No. 62073334.

CRediT authorship contribution statement

Fengji Ma: Conceptualization, Validation, Methodology, Software, Writing - original draft. **Junyi Chai:** Conceptualization, Validation,Writing - review & editing. **Haitao Li:** Validation, Formal analysis, Writing - review & editing. **Jinping Sun:** Project administration, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. A. Stark, Adaptive image contrast enhancement using generalizations of histogram equalization, IEEE TIP 9 (5) (2000) 889–896.

[2] C. Lee, C. Lee, C.-S. Kim, Contrast enhancement based on layered difference representation of 2d histograms, IEEE TIP 22 (12) (2013) 5372–5384.

[3] E. H. Land, The retinex theory of color vision, Scientific american 237 (6) (1977) 108–129.

[4] X. Guo, Y. Li, H. Ling, Lime: Low-light image enhancement via illumination map estimation, IEEE TIP 26 (2) (2016) 982–993.

[5] B. Cai, X. Xu, K. Guo, K. Jia, B. Hu, D. Tao, A joint intrinsic-extrinsic prior model for retinex, in: ICCV, 2017, pp. 4000–4009.

[6] X. Ren, W. Yang, W.-H. Cheng, J. Liu, Lr3m: Robust low-light enhancement via low-rank regularized retinex model, IEEE TIP 29 (2020) 5862–5876.

[7] M. Liu, L. Tang, S. Zhong, H. Luo, J. Peng, Learning noise-decoupled affine models for extreme low-light image enhancement, Neurocomputing 448 (2021) 21–29.

[8] Y. Wang, H. Wang, C. Yin, M. Dai, Biologically inspired image enhancement based on retinex, Neurocomputing 177 (2016) 373–384.

[9] F. Ma, J. Chai, H. Wang, Two-dimensional compact variational mode decomposition-based low-light image enhancement, IEEE Access 7 (2019) 136299–136309.

[10] W. Xia, E. Chen, S. Pautler, T. Peters, Laparoscopic image enhancement based on distributed retinex optimization with refined information fusion, Neurocomputing.

[11] X. Jia, X. Feng, W. Wang, L. Zhang, An extended variational image decomposition model for color image enhancement, Neurocomputing 322 (2018) 216–228.

[12] M. Li, J. Liu, W. Yang, X. Sun, Z. Guo, Structure-revealing low-light image enhancement via robust retinex model, IEEE TIP 27 (6) (2018) 2828–2841.

[13] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, R. Cong, Zero-reference deep curve estimation for low-light image enhancement, in: CVPR, 2020, pp. 1780–1789.

[14] K. Xu, X. Yang, B. Yin, R. W. Lau, Learning to restore low-light images via decomposition-and-enhancement, in: CVPR, 2020, pp. 2281–2290.

[15] F. Lv, F. Lu, J. Wu, C. Lim, Mbllen: Low-light image/video enhancement using cnns., in: BMVC, 2018, p. 220.

[16] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, J. Jia, Underexposed photo enhancement using deep illumination estimation, in: CVPR, 2019, pp. 6849–6857.

[17] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, Z. Wang, Enlightengan: Deep light enhancement without paired supervision, IEEE TIP 30 (2021) 2340–2349.

[18] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: MICCAI, Springer, 2015, pp. 234–241.

[19] Z. Jiang, H. Li, L. Liu, A. Men, H. Wang, A switched view of retinex: Deep self-regularized low-light image enhancement, Neurocomputing 454 (2021) 361–372.

[20] C. Wei, W. Wang, W. Yang, J. Liu, Deep retinex decomposition for low-light enhancement, arXiv preprint arXiv:1808.04560.

[21] Y. Zhang, J. Zhang, X. Guo, Kindling the darkness: A practical low-light image enhancer, in: ACM MM, 2019, pp. 1632–1640.

[22] J. Hai, Z. Xuan, R. Yang, Y. Hao, F. Zou, F. Lin, S. Han, R2rnet: Low-light image enhancement via real-low to real-normal network, arXiv preprint arXiv:2106.14501.

[23] M. Zhu, P. Pan, W. Chen, Y. Yang, Eemefn: Low-light image enhancement via edge-enhanced multi-exposure fusion network, in: AAAI, Vol. 34, 2020, pp. 13106–13113.

[24] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.

[25] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.

[26] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al., Photo-realistic single image super-resolution using a generative adversarial network, in: CVPR, 2017, pp. 4681–4690.

[27] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE TIP 13 (4) (2004) 600–612.

[28] V. Bychkovsky, S. Paris, E. Chan, F. Durand, Learning photographic global tonal adjustment with a database of input/output image pairs, in: CVPR 2011, IEEE, 2011, pp. 97–104.

[29] R. Liu, L. Ma, J. Zhang, X. Fan, Z. Luo, Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement, in: CVPR, 2021, pp. 10561–10570.

[30] J. Cai, S. Gu, L. Zhang, Learning a deep single image contrast enhancer from multi-exposure images, IEEE TIP 27 (4) (2018) 2049–2062.

[31] S. Wang, J. Zheng, H.-M. Hu, B. Li, Naturalness preserved enhancement algorithm for non-uniform illumination images, IEEE TIP 22 (9) (2013) 3538–3548.

[32] K. Ma, K. Zeng, Z. Wang, Perceptual quality assessment for multi-exposure image fusion, IEEE TIP 24 (11) (2015) 3345–3356.

[33] C. Lee, C. Lee, C.-S. Kim, Contrast enhancement based on layered difference representation, in: ICIP, 2012, pp. 965–968.

[34] N. Hautiere, J.-P. Tarel, D. Aubert, E. Dumont, Blind contrast enhancement assessment by gradient ratioing at visible edges, Image Analysis & Stereology 27 (2) (2008) 87–95.

[35] K. Gu, G. Zhai, W. Lin, X. Yang, W. Zhang, No-reference image sharpness assessment in autoregressive parameter space, IEEE TIP 24 (10) (2015) 3218–3231.