The 38th Annual AAAI Conference on Artificial Intelligence

FEBRUARY 20-27, 2024 | VANCOUVER, CANADA
VANCOUVER CONVENTION CENTRE – WEST BUILDING

DeVigNet: High-Resolution Vignetting Removal via a Dual Aggregated Fusion Transformer With Adaptive Channel Expansion

Shenghong Luo[1*], Xuhang Chen[123*], Weiwen Chen[12], Zinuo Li[12], Shuqiang Wang[2†], Chi-Man Pun[1†]

[1] University of Macau
[2] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
[3] Huizhou University
[*] These authors contributed equally.
[†]Corresponding Author

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Table of Contents

um 澳大

Introduction

Vignetting is a common optical degradation that results in a gradual decrease in brightness toward the edges of an image. It occurs due to multiple factors such as lens characteristics, filter presence, aperture settings, focal length settings, etc.

There are mathematical and prior-based methods available for vignetting removal (Zheng et al. 2008, 2013; Lopez-Fuentes, Oliver, and Massanet 2015). Nevertheless, these approaches have limitations.

1.These approaches ideally assume that the optical center is located at the center of the image, which may not be valid in real-world scenarios.

2.These methods can demonstrate bias under certain conditions and frequently necessitate extensive parameter adjustments to achieve optimal performance.

3.These parameters are highly sensitive to high-resolution images, often leading to inferior outcomes.

4.Significant challenge arises from the absence of ground truth in the datasets used for evaluation, which contributes to subjective assessments of the experimental results.

Related Work

Vignetting Removal

A limited number of studies in the field of traditional vignetting removal has proposed methods that are based on mathematical principles, statistical analysis, and prior knowledge. Additionally, vignetting datasets often consist of low-quality images or lack the necessary characteristics for effectively assessing vignetting removal algorithms. Unfortunately, at present, there is no accessible dataset exclusively designed for vignetting removal that provides reliable ground truth for objective evaluation.

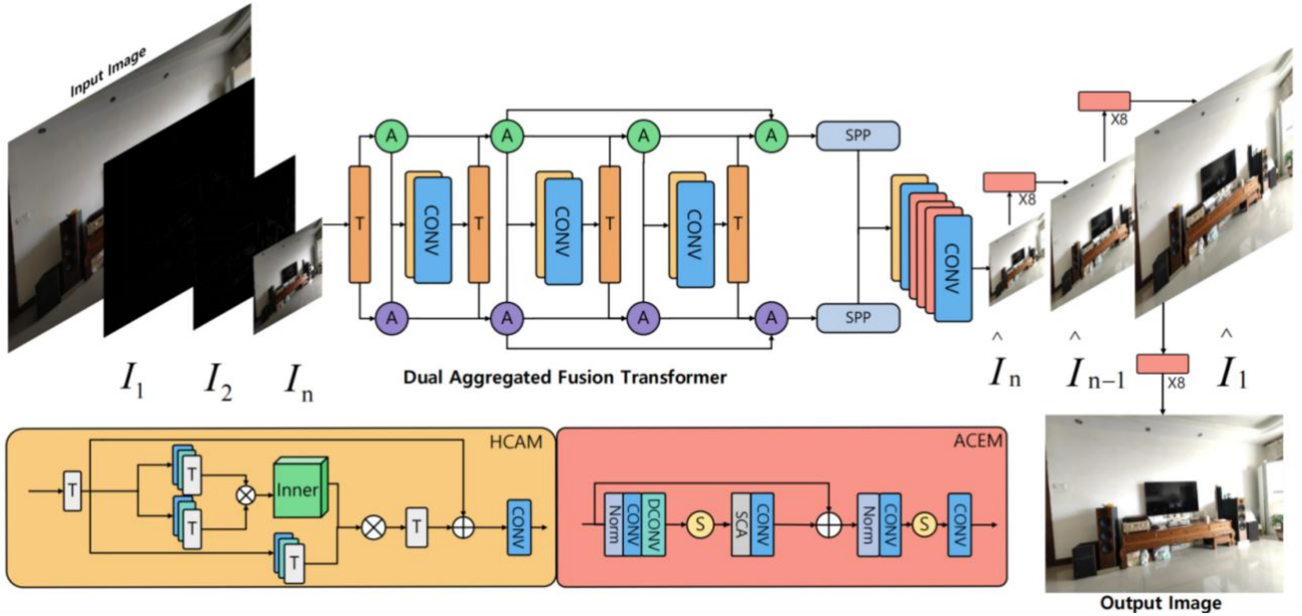
Low Light Enhancement

Traditional methods for Low-Light Image Enhancement often refers to the Retinex theory or histogram equalization, which is low accuracy.

Recently, the utilization of these learning-based approaches gain traction as a prevalent solution for enhancing low-light images. However, conventional methods heavily depend on assumptions based on physics, which are not consistently accurate. Additionally, current LLIE methods are not well-suited for addressing the issue of vignetting.

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Methodology



The model consists of three main components:

Dual Aggregated Fusion Transformer: It is a neural network designed specifically for handling low-frequency information in images. As the foundational architecture, the Fusion Transformer employs multiple attention mechanisms to capture various points of focus, enabling the model to effectively incorporate both local and global information within its representation. Fusion Transformer significantly boosts the expressive capacity of the Transformer network.

Adaptive Contrast Enhancement Module: It is a module to make Laplacian pyramid reconstruction for optimal edge information in vignetting removal results in high-frequency.

Hierarchical Channel Attention Module: This module is dedicated to the hierarchical fusion of features and the acquisition of learnable correlations across diverse layers. The primary function of it is to compute and apply attention weights to the input feature map, resulting in the refinement and enhancement of vignetting feature.

Dataset

VigSet is first vignetting dataset that includes high-resolution vignetting images along with the corresponding vignetting-free ground truth.
In well-lit situations, vignetting is generally not noticeable. Therefore, we use an ND filter to reduce illumination and capture photos with vignetting. The center light, which passes through an ND filter, travels a shorter distance compared to the light at the edges. This difference in distance contributes to the occurrence of vignetting.

To ensure that there is no displacement between each pair of photos and to enhance data diversity, it is crucial to avoid dynamic objects such as swaying tree leaves, moving vehicles, and glass reflections on specific objects that are susceptible to motion. Therefore, we select multiple distinct indoor environments for the data collection.

To reduce device vibration caused by manual camera shutter presses, we use remote shutter control.

Experiments and Results

Dataset	VigSet
Train Set	803
Test Set	180

The LOL and MIT-Adobe FiveK, we adopt the experimental settings delineated in (Wang et al. 2023), ensuring consistency with prior research.

Detailed implementation and data available at https://github.com/CXH-Research/DeVigNet

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Results

Methods	512 × 512				1024 × 1024				2048 × 2048			
	PSNR ↑	SSIM ↑	MAE ↓	LPIPS ↓	PSNR ↑	SSIM ↑	MAE ↓	LPIPS ↓	PSNR ↑	SSIM ↑	MAE ↓	LPIPS ↓
Input	12.08	0.59	60.04	0.18	12.08	0.58	60.04	0.18	12.08	0.58	60.04	0.19
RIVC	13.08	0.59	55.17	0.18	13.08	0.59	55.17	0.18	13.08	0.59	55.17	0.19
SIVC	14.65	0.62	43.71	0.17	14.65	0.61	43.69	0.17	14.65	0.60	43.70	0.18
LIME	12.42	0.41	51.29	0.41	12.42	0.39	51.29	0.40	12.41	0.37	51.32	0.42
MSRCR	11.20	0.39	60.43	0.44	11.20	0.37	60.43	0.45	11.20	0.35	60.44	0.46
NPE	15.72	0.51	38.60	0.33	15.72	0.49	38.59	0.32	15.72	0.47	38.60	0.33
WV-SRIE	18.84	0.60	26.67	0.22	18.84	0.58	26.67	0.23	18.84	0.56	26.67	0.24
PM-SRIE	19.45	0.66	24.81	0.16	19.45	0.64	24.81	0.17	19.45	0.62	24.81	0.19
JieP	18.93	0.58	26.33	0.24	18.93	0.55	26.32	0.25	18.93	0.54	26.32	0.27
KID	14.73	0.71	44.01	0.18	14.74	0.71	43.95	0.22	14.73	0.71	44.00	0.31
DSLR	19.37	0.65	24.10	0.16	19.37	0.64	24.07	0.20	19.35	0.62	24.10	0.29
ELGAN	16.32	0.73	37.77	0.10	16.32	0.72	37.76	**0.11**	16.31	0.72	37.77	**0.12**
RUAS	15.54	0.60	36.93	0.22	15.54	0.57	36.92	0.24	15.54	0.56	36.92	0.25
Zero-DCE	16.28	0.58	34.77	0.26	16.28	0.57	34.77	0.26	16.28	0.55	34.78	0.26
Zero-DCE++	16.82	0.55	32.31	0.20	16.82	0.52	32.32	0.21	16.81	0.51	32.34	0.23
Uformer	20.95	0.77	21.32	0.19	20.60	0.77	22.80	0.25	20.67	0.77	22.69	0.28
Ours	**22.96**	**0.79**	**15.82**	**0.09**	**22.94**	**0.78**	**15.84**	**0.11**	**22.94**	**0.77**	**15.85**	0.13

Results

Methods	LOL				MIT-Adobe FiveK			
	PSNR ↑	SSIM ↑	MAE ↓	LPIPS ↓	PSNR ↑	SSIM ↑	MAE ↓	LPIPS ↓
Input	7.77	0.20	99.80	0.42	12.26	0.61	55.98	0.22
LIME	16.76	0.44	30.59	0.31	13.30	0.75	52.12	0.09
MSRCR	13.17	0.45	52.71	0.33	13.31	0.75	50.82	0.12
NPE	16.97	0.47	32.89	0.31	17.38	0.79	31.22	0.10
WV-SRIE	11.86	0.49	65.56	0.24	18.63	0.84	26.27	0.08
PM-SRIE	12.28	0.51	63.28	0.23	19.70	0.84	23.42	0.07
JieP	12.05	0.51	64.34	0.22	18.64	0.84	26.42	0.07
RetinexNet	16.77	0.42	32.02	0.38	12.51	0.69	52.73	0.20
KID	17.65	**0.77**	31.40	**0.12**	16.20	0.79	35.16	0.11
DSLR	14.98	0.60	48.90	0.27	20.24	0.83	22.45	0.10
ELGAN	17.48	0.65	34.47	0.23	16.00	0.79	36.37	0.09
RUAS	16.40	0.50	39.11	0.19	17.91	0.84	33.12	0.08
Zero_DCE	14.86	0.56	47.07	0.24	15.93	0.77	36.36	0.12
Zero_DCE++	14.75	0.52	45.94	0.22	14.61	0.42	39.25	0.16
Uformer	18.55	0.73	28.91	0.23	21.92	**0.87**	17.91	**0.06**
Ours	**21.33**	0.76	**19.30**	0.16	**23.10**	0.84	**15.43**	0.16

Results

(a) Input (b) RIVC (c) SIVC

(d) ELGAN (e) Ours (f) Target

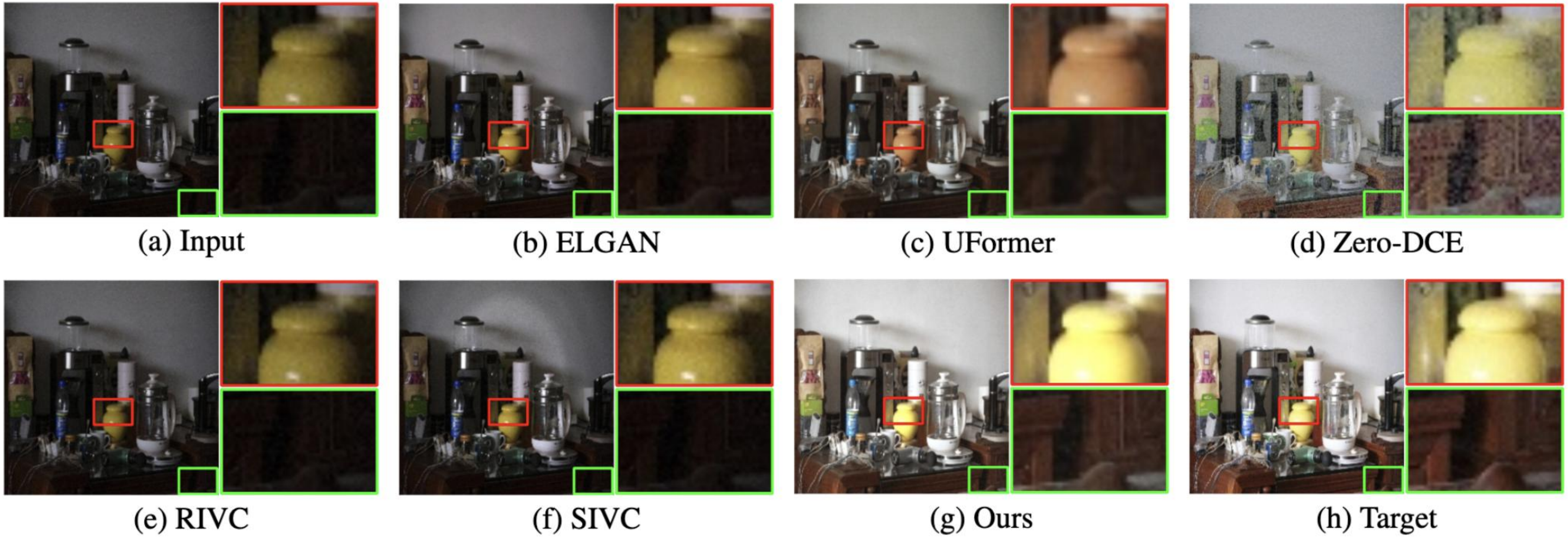


Figure 5: Visual comparison of various Vignetting Removal and LLIE methods on the VigSet dataset is presented. The figure clearly illustrates the presence of noticeable vignetting in methods (b), (f), and (e). Color degradation or distortion issues are apparent in methods (c) and (d).

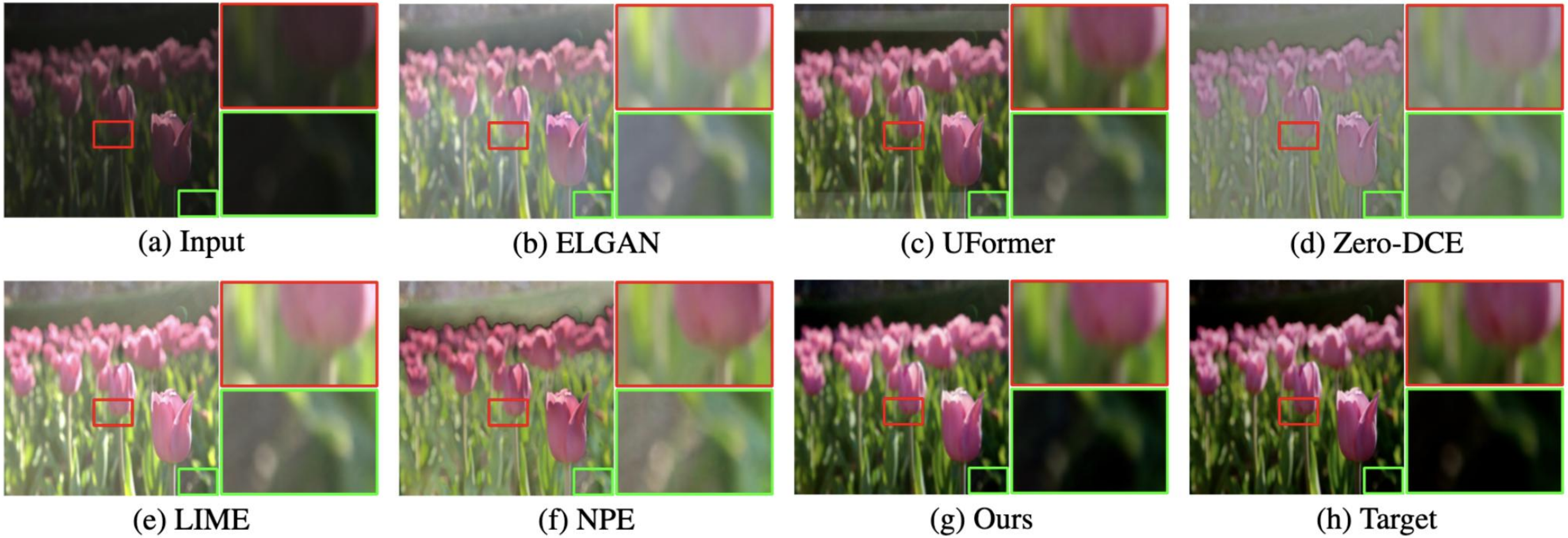


Figure 6: A visual comparison of different LLIE methods is conducted on the MIT-Adobe FiveK dataset. (b), (e), and (f) exhibit overexposure in the entire image. Both (d) and (e) exhibit either blurriness or distortion, respectively.

Conclusion

In this paper, we introduce Vigset, the first large-scale high-resolution vignetting removal dataset with ground truth images. Vigset comprises 983 pairs of images captured under different lighting conditions and in various scenes. Additionally, we propose a novel method called DeVigNet, specifically designed for vignetting removal on this dataset. It includes three components: The Dual Aggregated Fusion Transformer, the Adaptive Channel Expansion Module and the Hierarchical Channel Attention Module. By utilizing the Laplacian pyramid, DeVigNet performs vignetting removal on the color information in the high-frequency and low-frequency domains of the image, thereby achieving optimal results. DeVigNet effectively eliminates vignetting effects in images, demonstrating superior performance compared to existing methods in terms of both quality and quantity for vignetting removal.

澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

Acknowledgment

This work was supported in part by the Science and Technology Development Fund, Macau SAR, under Grant 0087/2020/A2 and Grant 0141/2023/RIA2, in part by the National Natural Science Foundations of China under Grant 62172403, in part by the Distinguished Young Scholars Fund of Guangdong under Grant 2021B1515020019, in part by the Excellent Young Scholars of Shenzhen under Grant RCYX20200714114641211.

Thank You!

Avenida da Universidade, Taipa, Macau, China
Tel : (853) 8822 4369
Email : cmpun@umac.mo Website : https://www.cis.um.edu.mo/~cmpun/