DURRNet: Deep Unfolded Single Image Reflection Removal Network With Joint Prior

Jun-Jie Huang[1]†, Tianrui Liu[1]†, Jingyuan Xia[1], Meng Wang[2] and Pier Luigi Dragotti[3]

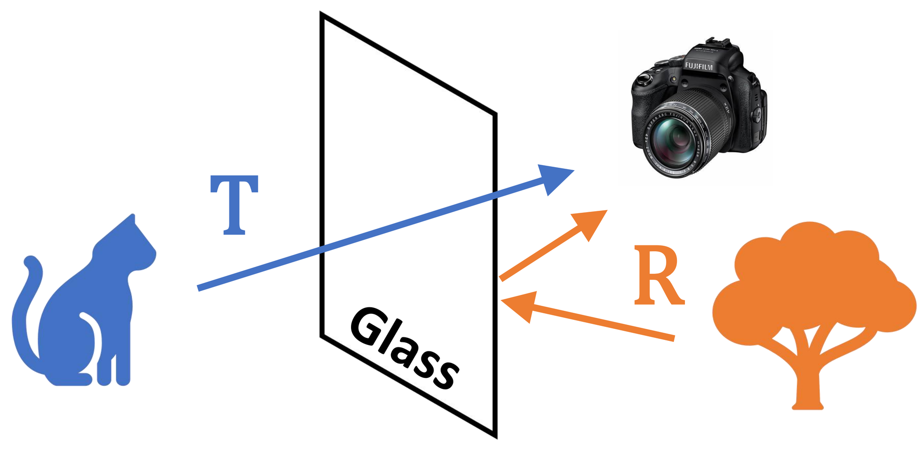
[1]College of Computer Science and Technology, National University of Defense Technology, China

[2]School of Computer Science and Information Engineering, Hefei University of Technology, China

[3]Department of Electrical and Electronic Engineering, Imperial College London, UK

Introduction

Single Image Reflection Removal is a typical Blind Source Separation problem. It aims to decompose a reflection-contaminated image **I** captured through a glass into a transmission image **T** and a reflection image **R**.

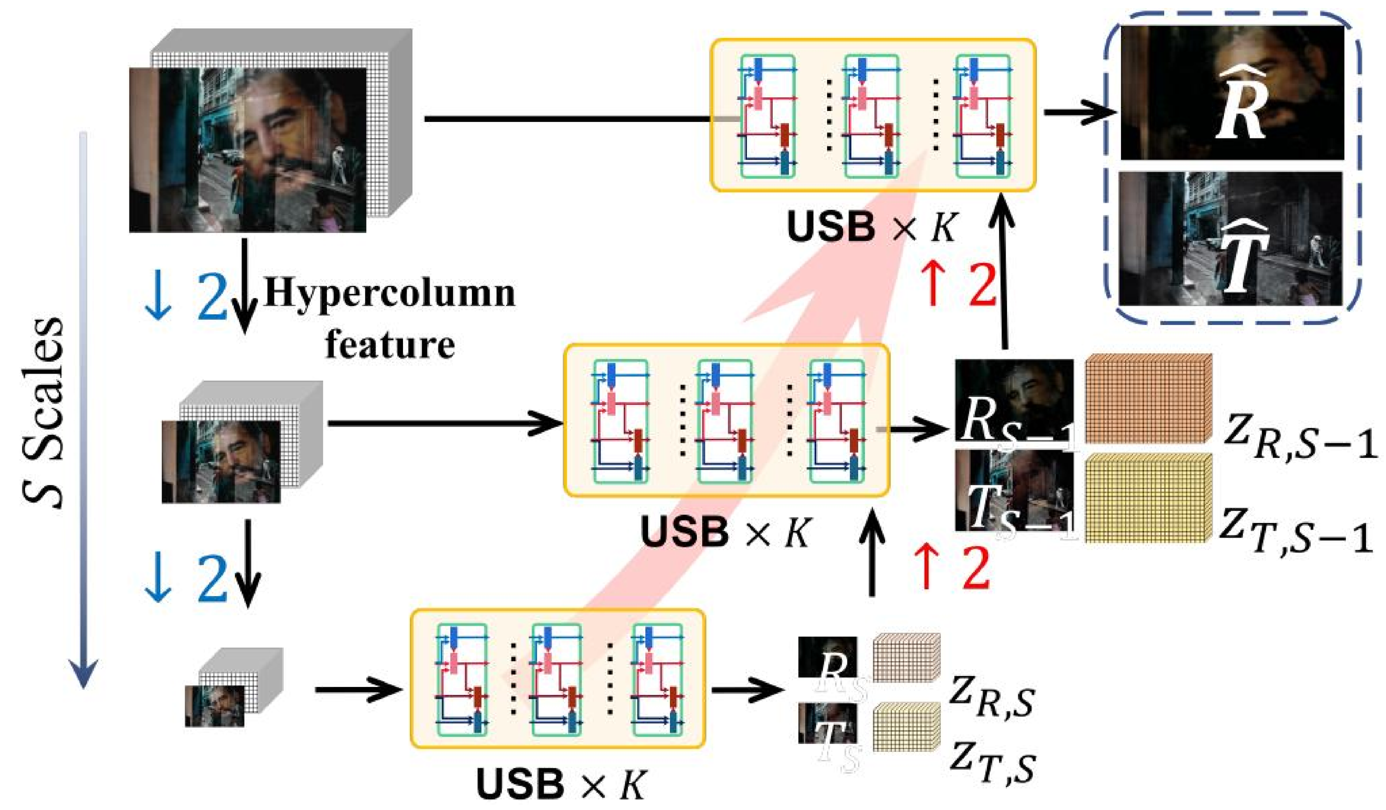


- Highly ill-posed problem
- Minimize shared image contents

In this paper, we propose **a principle approach** to design a novel **Deep Unfolded Reflection Removal Network**

✓ good interpretability

✓ good trade-off between performance and complexity



Proposed Method

☐ Model Formulation

The interplay between the shared contents of image layers are essential for effective separation, while is often overlooked by the existing deep architectures.

We propose a Convolutional Sparse Coding (CSC) based optimization formulation with a transform-based exclusion prior:

$$\min_{\mathbf{z}_T,\mathbf{z}_R} \frac{1}{2}\left\|\mathbf{I} - \sum_{i=1}^{N} \mathbf{D}_T^i \otimes \mathbf{z}_T^i - \sum_{i=1}^{N} \mathbf{D}_R^i \otimes \mathbf{z}_R^i\right\|_F^2 + \lambda_T p_T(\mathbf{z}_T)$$

$$+ \lambda_R p_R(\mathbf{z}_R) + \kappa\mathcal{E}\left(\sum_{i=1}^{N} \mathbf{D}_T^i \otimes \mathbf{z}_T^i, \sum_{i=1}^{N} \mathbf{D}_R^i \otimes \mathbf{z}_R^i\right)$$

Exclusion Prior: $\mathcal{E}(\mathbf{T},\mathbf{R}) = \sum_{m=1}^{M} \|(\mathbf{W}_m \otimes \mathbf{T}) \odot (\mathbf{W}_m \otimes \mathbf{R})\|_1$

where **W** denotes wavelet transform.

☐ Optimization Algorithm

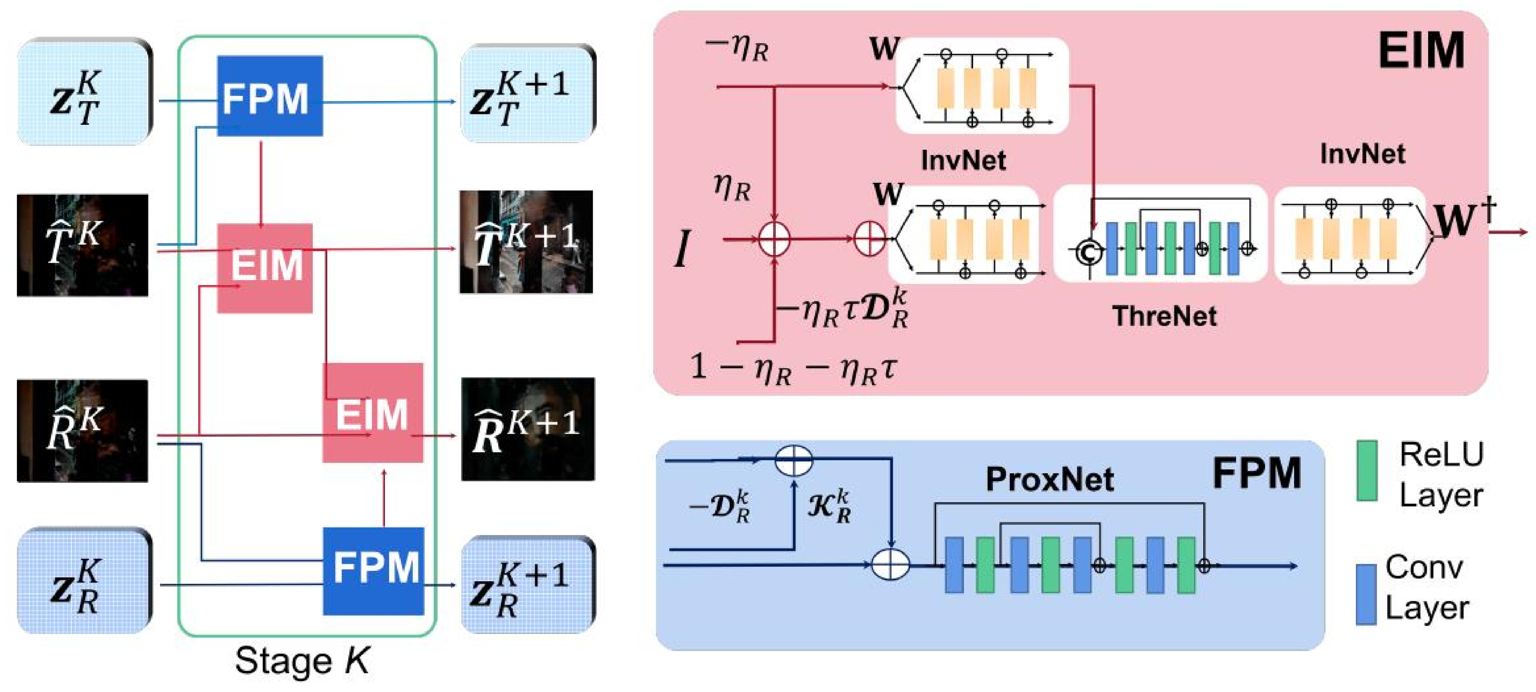
The optimization formulation is reformulated by Half-Quadratic Splitting (HQS) algorithm. Based on Proximal Gradient Descent (PGD), we design an iterative Feature and Auxiliary SeparaTion (i-FAST) algorithm with closed-form solutions for each sub-problem:

$$\begin{cases} \mathbf{z}_T^{k+1} = \mathrm{prox}_{\eta_1\lambda_T/\tau}\left(\mathbf{z}_T^k - \eta_1 \nabla f(\mathbf{z}_T^k)\right), \\ \mathbf{z}_R^{k+1} = \mathrm{prox}_{\eta_2\lambda_R/\tau}\left(\mathbf{z}_R^k - \eta_2 \nabla h(\mathbf{z}_R^k)\right), \\ \hat{\mathbf{T}}^{k+1} = \sum_{m=1}^{M} \mathbf{W}_m^\dagger \otimes \mathcal{S}_{\kappa|\mathbf{W}_m \otimes \hat{\mathbf{R}}^k|}\left(\mathbf{W}_m \otimes \phi(\hat{\mathbf{T}}^k)\right), \\ \hat{\mathbf{R}}^{k+1} = \sum_{m=1}^{M} \mathbf{W}_m^\dagger \otimes \mathcal{S}_{\kappa|\mathbf{W}_m \otimes \hat{\mathbf{T}}^{k+1}|}\left(\mathbf{W}_m \otimes \psi(\hat{\mathbf{R}}^k)\right), \end{cases}$$

☐ Proposed Deep Unfolded Network

We unfold and parameterize each iteration of i-FAST algorithm into a **Unfolded Separation Block (USB):**

- **Feature Prediction Module (FPM):** the proximal operators $\mathrm{prox}_{\eta\lambda/\tau}(\cdot)$ are parameterized by a learnable ProxNet which is a deep residual network with a global skip connection to learning from training dataset to capture the prior information
- **Exclusion Interaction Module (EIM):** we propose to overparameterize the fixed wavelet transform as a learnable **Invertible Neural Network** (InvNet) and to use a Thresholding Network (ThreNet) to replace the soft-thresholding operator.



Experimental Results

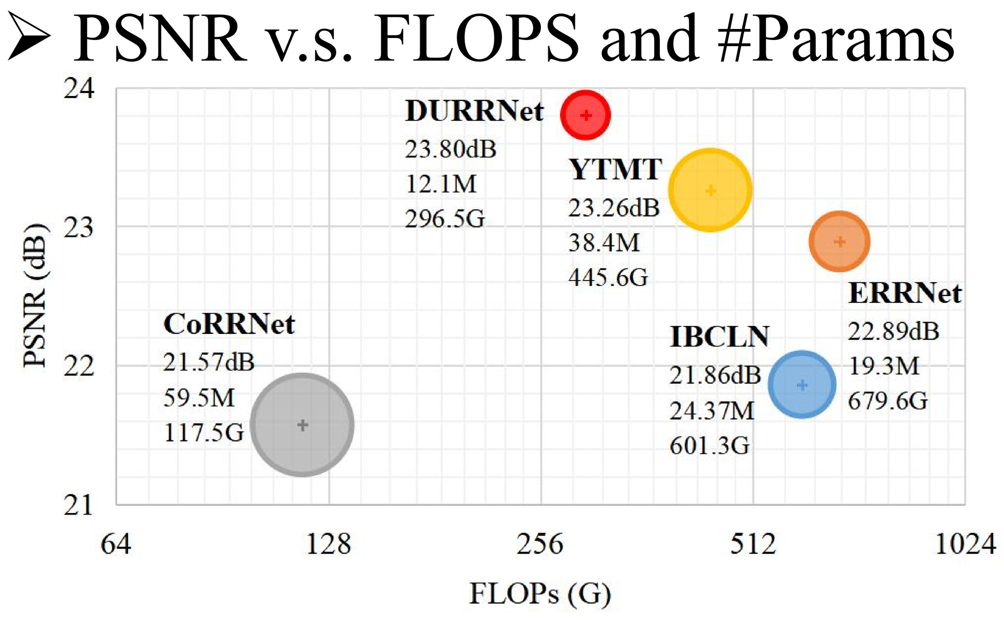
➢ Objective comparisons:

Dataset	Metrics	CEILNet [8]	Zhang *et al.* [10]	BDN [11]	IBCLN [14]	CoRRN [22]	ERRNet [13]	YTMT [17]	DURRNet (proposed)
Real20 (20)	PSNR (↑)	18.45	22.55	18.41	21.86	21.57	22.89	23.26	**23.80**
	SSIM (↑)	0.690	0.788	0.726	0.762	0.807	0.803	0.806	**0.814**
Nature (20)	PSNR (↑)	19.33	19.56	18.92	23.57	21.84	20.60	23.85	**24.24**
	SSIM (↑)	0.745	0.736	0.737	0.783	0.805	0.755	0.810	**0.812**

➢ Subjective comparisons

(a) Input (b) Zhang [10] (c) BDN [11] (d) IBCLN [14] (e) ERRNet [13] (f) YTMT [17] (g) DURRNet (h) Ground-truth

➢ PSNR v.s. FLOPS and #Params



➢ By combing the merits of model-based and learning-based paradigms, DURRNet achieves SOTA performance with fewer parameters and lower complexity.

Conclusions

✓ A novel model-inspired single image reflection removal network.

✓ Model-based optimization formulation which exploits the sparsity based image model with joint priors.

✓ By unfolding i-FAST, the proposed DURRNet is constructed using multiple scales of Unfolded Separation Blocks (USB).