Residual-Conditioned Optimal Transport: Towards Structure-Preserving Unpaired and Paired Image Restoration

Xiaole Tang, Xin Hu, Xiang Gu, Jian Sun

School of Mathematics and Statistics, Xi'an Jiaotong University, China

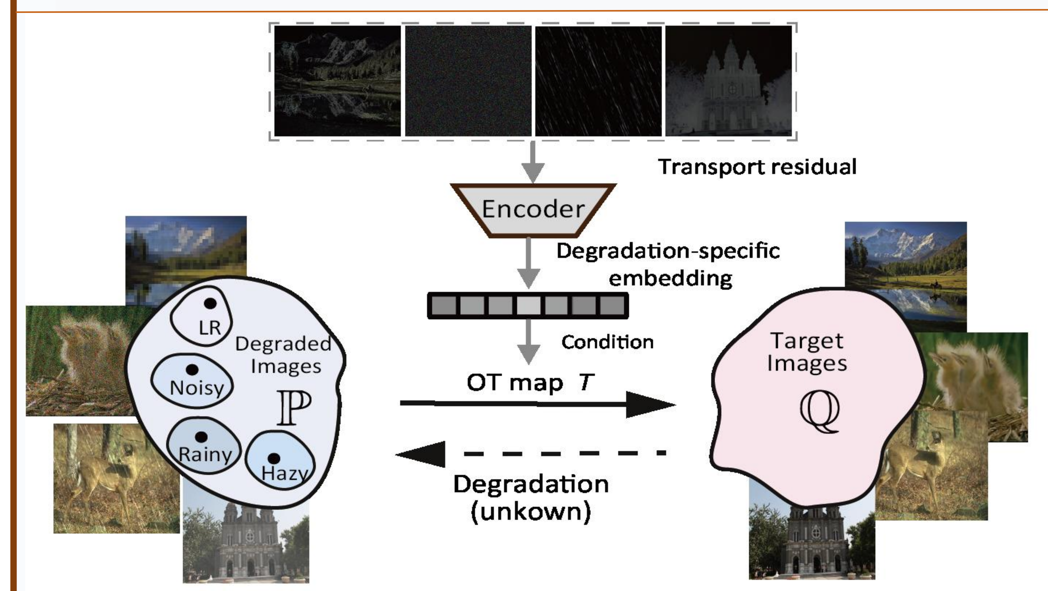
Motivation and Demo

Determinative methods (e.g., MPRNet, Restormer, and PromptIR):

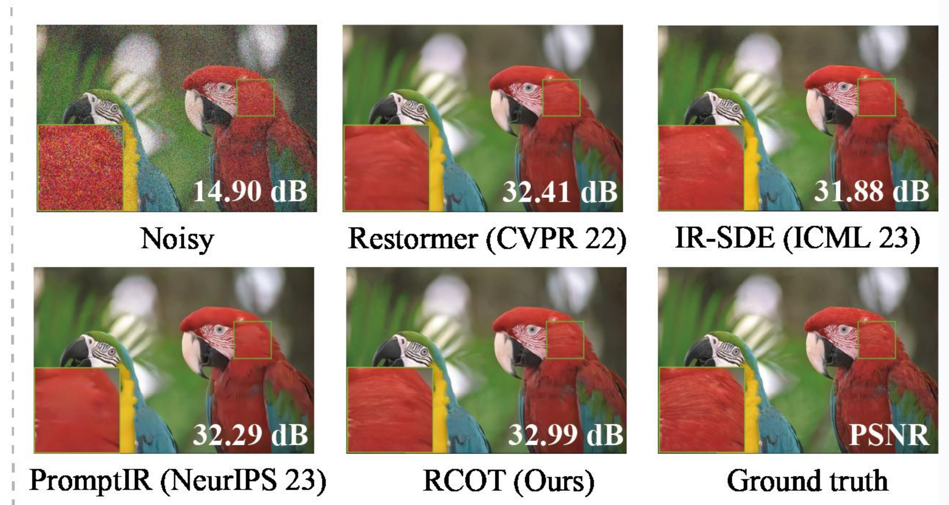
- fit **regression models** and minimize **distortion measures** (L1, L2, MSE, etc.)
- often capture the **"mean" results** of the potential targets, which may be vulnerable to **excessive smoothness and compromised structural details**.

Generation-based methods (e.g., IR-SDE, CycleGAN):

- often use the **degraded image as a condition** without including specific information about the degradation.
- can result in outputs with **remaining distortions** and **inaccurate structures**.



(a) The core idea of RCOT that uses transport residual embedding as condition for the optimal transport (OT) map.



(b) Denoising example: the RCOT approach produces results with better structural content.

Our approach (RCOT) models image restoration as an **OT problem**, and introduce the **transport residual as a degradation-specific cue for both the transport cost and transport map** for degradation-aware and structure-preserving restoration.

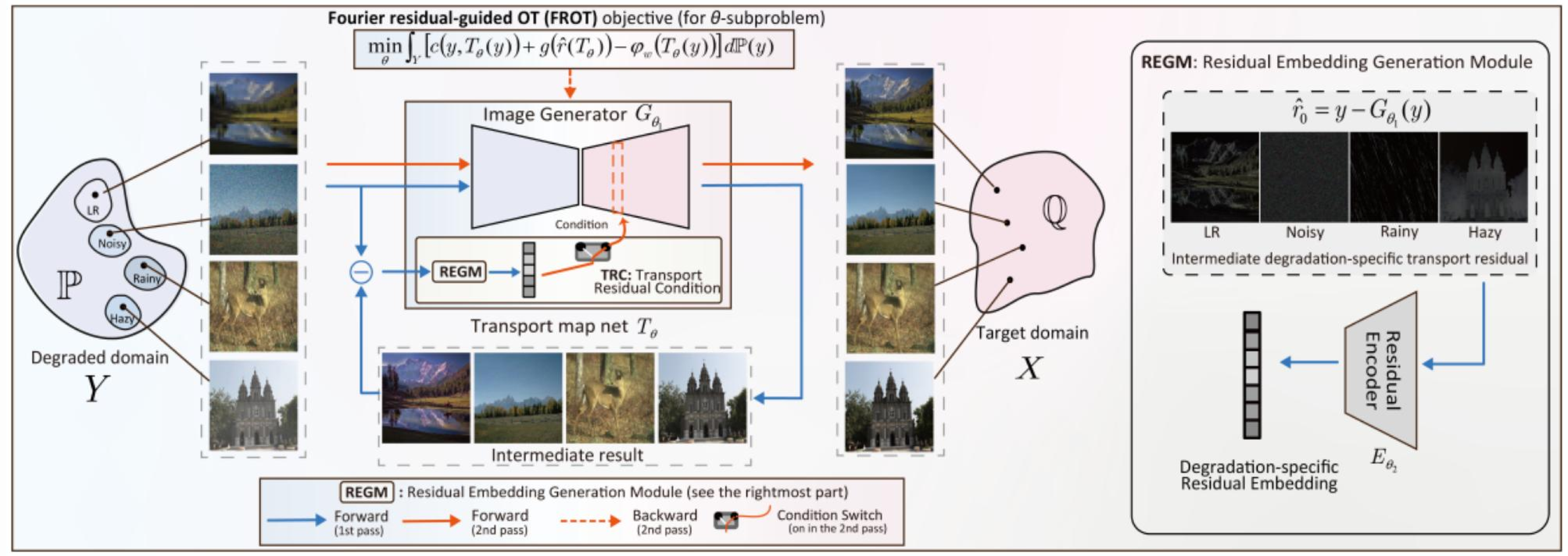
Contribution

Genearl Novelty: Residual-Conditioned Optimal Transport (RCOT)

+ Applications in Image Restoration

- RCOT: introduces transport residual as a degradation-specific cue into the OT framework (i.e., the transport cost and map);
- controls the transportation (restoration) with a two-pass transport residual condition (TRC) mechanism;
- tackles structure-preserving issues for multiple IR tasks;

Our Approach: Residual-Conditioned Optimal Transport (RCOT)



Key idea: incorporate the degradation-specific knowledge (from the residual or its embedding) into the transport cost, and more importantly, into the transport map via a two-pass transport residual conditioning mechanism.

Fourier Residual-guided OT Objective

Kantorovich Prime

$$\mathrm{FROT}(\mathbb{P},\mathbb{Q}) \triangleq \inf_{\pi\in\Pi(\mathbb{P},\mathbb{Q})} \int_{X\times Y} \tilde{c}(y,x)d\pi(y,x).$$

$$\tilde{c}(y,x) = c(x,y) + g(r) \qquad r = y - x$$

Kantorovich Dual

$$\mathrm{FROT}(\mathbb{P},\mathbb{Q}) = \sup_{\varphi}\inf_{T}\Big\{\mathcal{L}(T,\varphi) \triangleq \int_X \varphi(x)d\mathbb{Q}(x) + \int_Y [c(T(y),y) + g(\hat{r}(T)) - \varphi(T(y))]\,d\mathbb{P}(y)\Big\}$$

Two-pass RCOT Map

The first pass unconditionally generates an intermediate result along with the estimated residual $\hat{r}_0$. The second pass then restores the refined result conditioned on the residual embedding $E_{\theta_2}(\hat{r}_0)$.

$$\hat{r}_0 = y - G_{\theta_1}(y), \quad T_\theta(y) = G_{\theta_1}(y|E_{\theta_2}(\hat{r}_0)).$$

Experiments

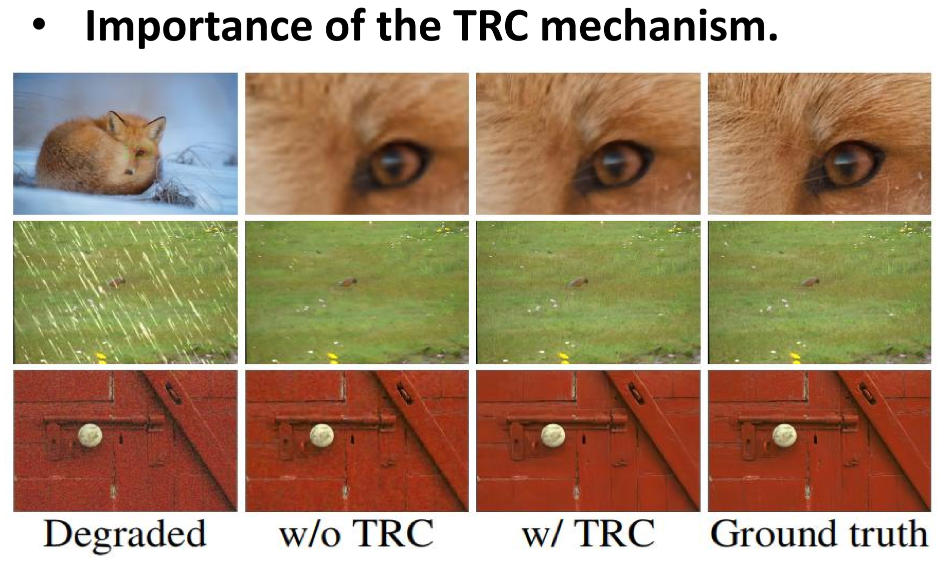
Ablation Study

- **Effect of the loss components.**

Loss	PSNR ↑	SSIM ↑	FID ↓
$\mathcal{L}_{\mathrm{FROT}}$	27.60	0.772	69.43
supervised ℓ_2	27.69	0.779	78.69
$\mathcal{L}_{\mathrm{FROT}}+\ell_2$	**28.25**	**0.799**	**56.60**

FROT shows decent performance, working alone and with the L_2 loss. Integrating FROT and supervised L_2 provides a significant gain to our model.

- **Importance of the TRC mechanism.**

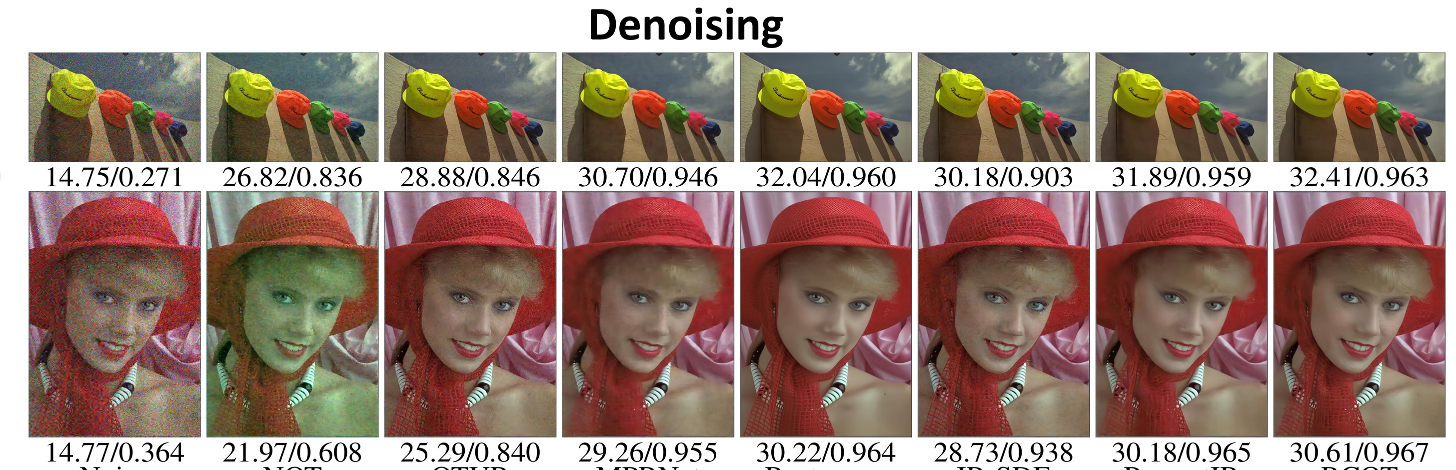


Method	SR on DIV2K	Deraining on Rain100L	Denoising on Kodak24	Average
w/o TRC	27.24/0.785/6.72	34.20/0.955/28.63	28.45/0.790/64.20	29.97/0.843/32.47
w/ TRC	**28.41/0.804/1.22**	**37.27/0.978/8.67**	**29.53/0.828/33.13**	**31.72/0.870/14.34**

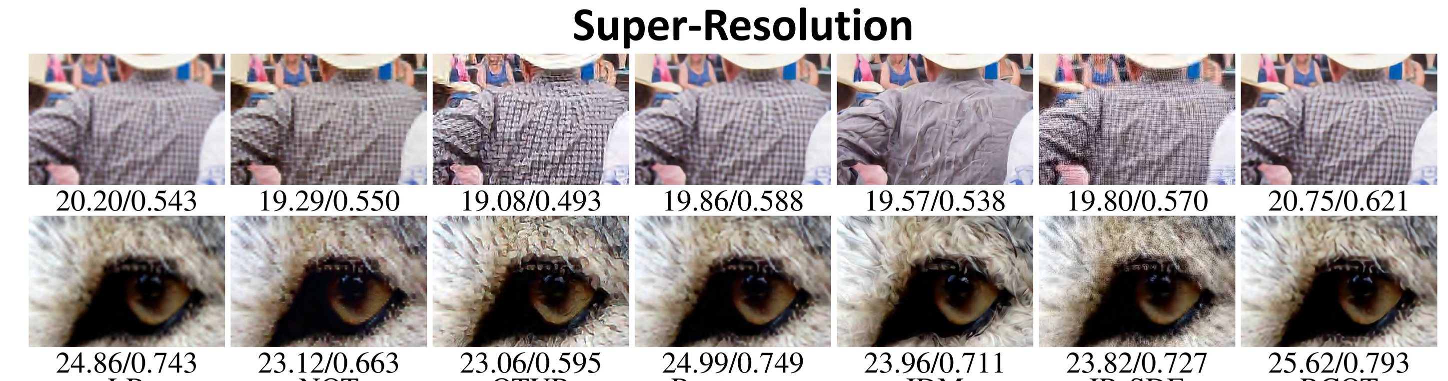
The proposed TRC module yields an average gain of 1.75 dB of PSNR value over the basic model. With the TRC, the model restores images with better structures.

Main Results

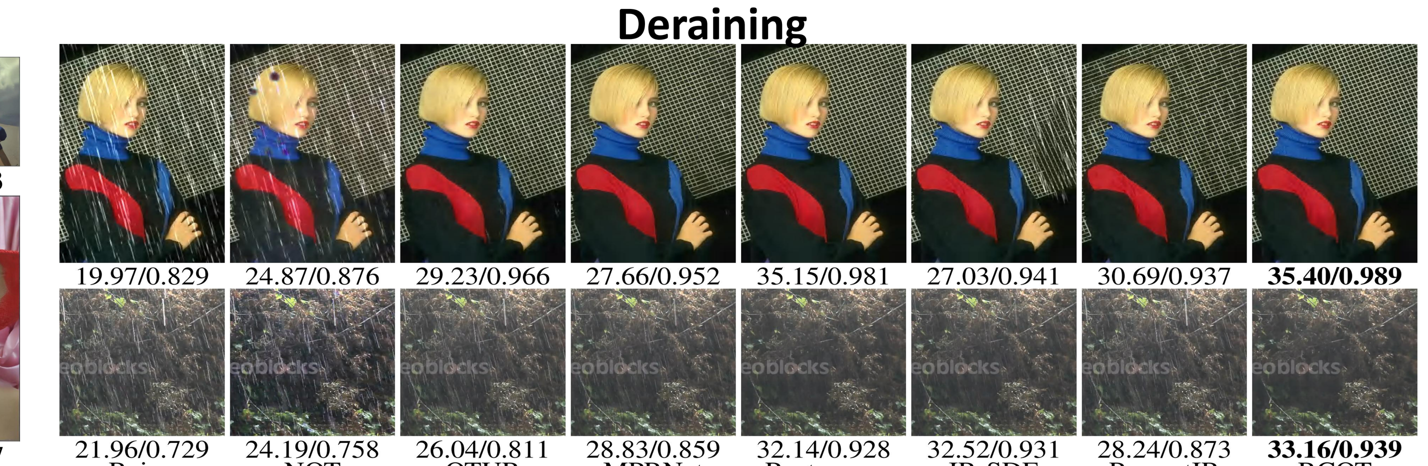
Denoising



Super-Resolution



Deraining



Dehazing