RainDiffuse: A Diffusion Framework for Deraining

Ziyue Lin, Runshi Yang, Yi Zhang
University of Toronto

Motivation

Rain severely degrades outdoor images in two distinct ways:

- **Raindrops on the camera lens** cause localized geometric distortions that blur, magnify, or refract background structures.
- **Fine rain streaks** introduce global high-frequency noise that reduces contrast and obscures important scene details.

When both appear together, the degradation becomes even more complex and challenging to reverse (Figure 1).

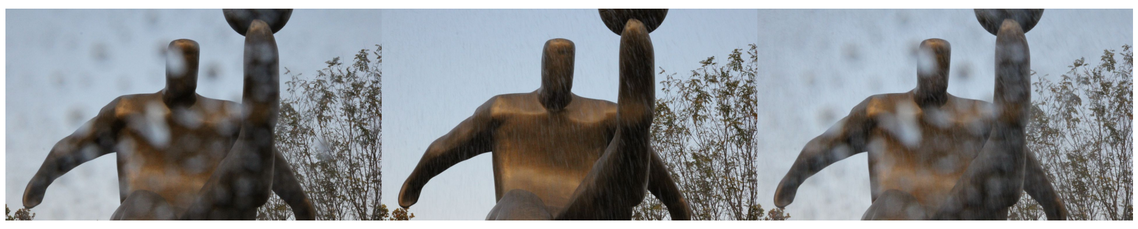

Figure 1: raindrop, rain streak, and combined raindrop + rain streak degradation.

Reliable deraining is crucial for a wide range of downstream applications, such as autonomous driving, video surveillance, robust object detection, and outdoor image enhancement, where rain artifacts can substantially impair visual understanding.

Related Work

- **Filter-Based Methods**
  Classical approaches use filter-based decomposition of high-frequency components.[1]
  Limitation: break down when noise is nonstationary or structured.
- **CNN-based Methods**
  Jointly detect rain regions and remove streaks within a single feed-forward network. [2, 3]
  Limitation: struggle with highly complex or overlapping rain patterns.
- **GAN-based Methods**
  Models such as AttentiveGAN [4] incorporate spatial attention to emphasize rain-affected areas and guide the generator.
  Limitation: GAN training is inherently unstable and often introduces hallucinated textures or structural artifacts.
- **Diffusion Models**
  DDPM/DDIM [5, 6] progressively denoise images using a learned noise schedule.
  Literature gap: rarely explored for deraining because Gaussian noise assumptions misalign with structured, directional rain.

New Technique

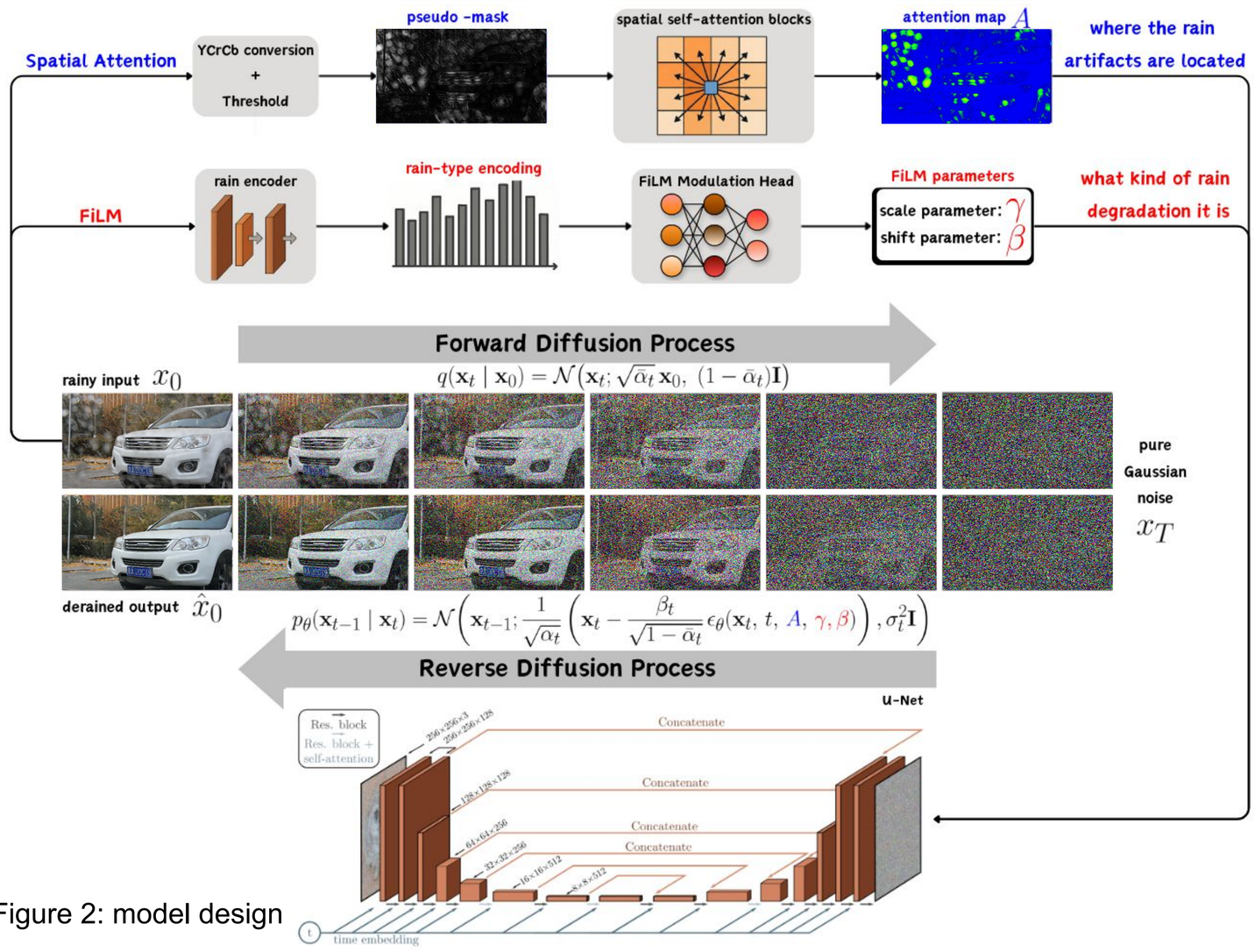

Figure 2: model design

Model Architecture

- **Spatial Attention** injects location information by producing an attention map that highlights rain-affected regions and amplifies or suppresses U-Net features accordingly.
- **Feature-wise Linear Modulation (FiLM)** injects rain-type information into the U-Net by applying feature-wise scale and shift to modulate its intermediate activations.
- **Diffusion U-Net** predicts the noise from a noisy input by combining timestep embeddings, spatial attention, and FiLM conditioning to guide the reverse diffusion process.

Loss Function

$$\mathcal{L} = \lambda_{\text{noise}}\mathcal{L}_{\text{noise}} + \lambda_{\text{rec}}\mathcal{L}_{\text{rec}} + \lambda_{\text{cons}}\mathcal{L}_{\text{cons}} + \lambda_{\text{seg}}\mathcal{L}_{\text{seg}}$$

- $\mathcal{L}_{\text{noise}} = \|\hat{\epsilon} - \epsilon\|_2^2$ fits to true Gaussian noise.
- $\mathcal{L}_{\text{rec}} = \|\hat{x}_0 - x_0\|_1$ ensures predicted image matches target.
- $\mathcal{L}_{\text{cons}} = \left\|\hat{x}_0^{(i)} - \hat{x}_0^{(j)}\right\|_1$ stabilizes results across samples.
- $\mathcal{L}_{\text{seg}} = \text{BCE}(m_{\text{pred}}, m_{\text{pseudo}}) + \text{Dice}(m_{\text{pred}}, m_{\text{pseudo}})$ supervises attention mask using pseudo-map, robust to imbalance.

Experimental Results

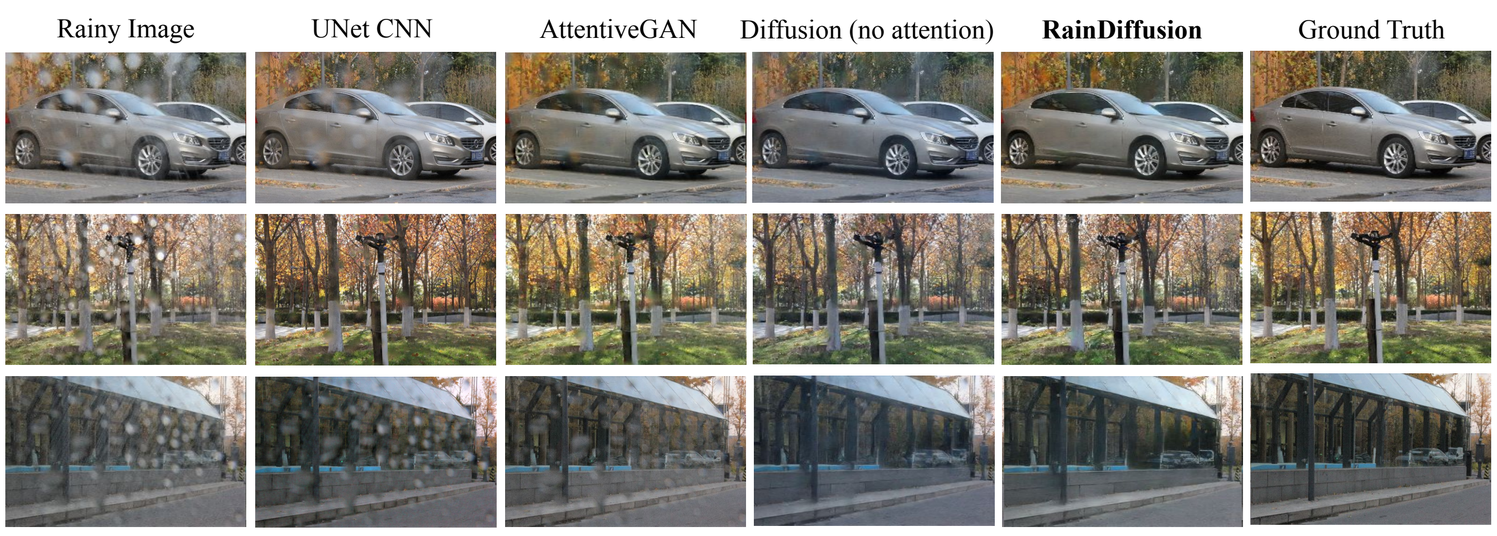

Figure 3: qualitative results

Table 1: PSNR (dB) comparison across datasets

Method	Raindrop	Rainstreak	Mixture
UNet-CNN	20.15	21.74	18.88
AttentiveGAN	21.60	23.77	20.72
Diffusion (no attention)	22.46	**24.66**	21.05
Attentive Diffusion	**22.57**	24.46	**21.61**

Table 2: SSIM comparison across datasets

Method	Raindrop	Rainstreak	Mixture
UNet-CNN	0.5768	0.6108	0.5229
AttentiveGAN	0.7019	0.7561	0.6449
Diffusion (no attention)	0.7255	**0.7858**	0.6690
Attentive Diffusion	**0.7378**	0.7723	**0.6808**

Figure 4: quantitative results

References

[1] Kang et al., *Automatic Single-Image-Based Rain Streaks Removal via Image Decomposition*, IEEE Trans. Image Processing, 2012.
[2] Fu et al., *Clearing the Skies: A Deep Network Architecture for Single-Image Rain Removal*, TIP 2017.
[3] Yang et al., *Deep Joint Rain Detection and Removal from a Single Image*, CVPR 2017.
[4] Qian et al., *AttentiveGAN for Raindrop Removal*, CVPR 2018.
[5] Ho et al., *Denoising Diffusion Probabilistic Models*, NeurIPS 2020.
[6] Song et al., *Denoising Diffusion Implicit Models*, ICLR 2021.