TPSeNCE: Towards Artifact-Free Realistic Rain Generation for Deraining and Object Detection in Rain

Shen Zheng [1], Changjie Lu [2], Srinivasa Narasimhan [1]

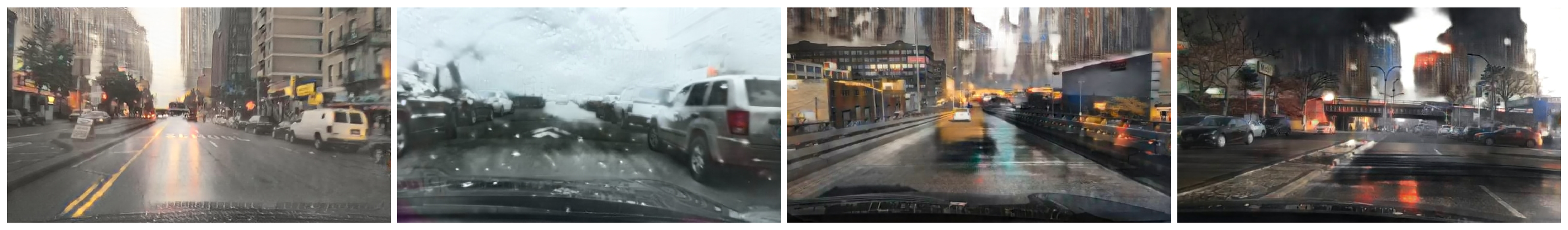
[1]Carnegie Mellon University, [2]University of Illinois Urbana-Champaign

JAN 4-8 WACV 2024 WAIKOLOA HAWAII

Backgrounds

- Rain is a common bad weather that significantly impair the vision systems.
- Rain drops and road reflections bring challenges for deraining and object detection in rain.

Challenges



- Rain generation algorithms have the potential to improve object detection in rain.
- Current methods yield undesirable rain amounts and many artifacts and distortions.

Contributions & Prospects

- Triangular Probability Similarity (TPS) for minimizing artifacts and distortions.
- Semantic Noise Contrastive Estimation (SeNCE) for optimizing the amounts of generated rain.
- Realistic rain generation benefits deraining and object detection in real rainy conditions.
- Future Works: (1) Extremely Heavy Rain (2) Strong Light Sources (3) Diffusion Models

Proposed Method

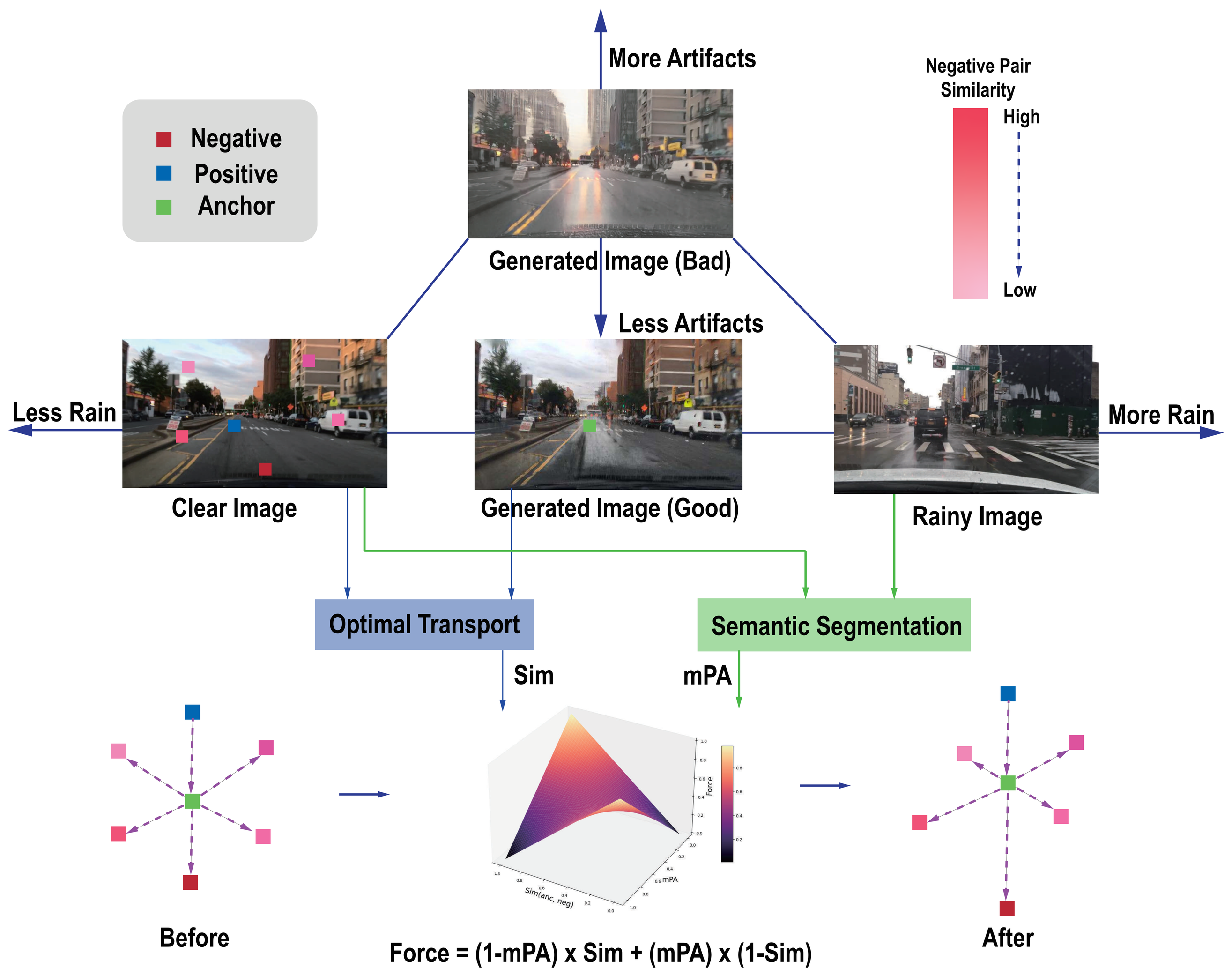


Figure 1. TPS aligns the generated rainy image with the clear and rainy image in the discriminator manifold to suppress artifacts and distortions, while SeNCE adjusts the pushing force of the negative patches based on their feature similarities with the anchor patch and refines that force with the semantic similarity between clear and rainy.

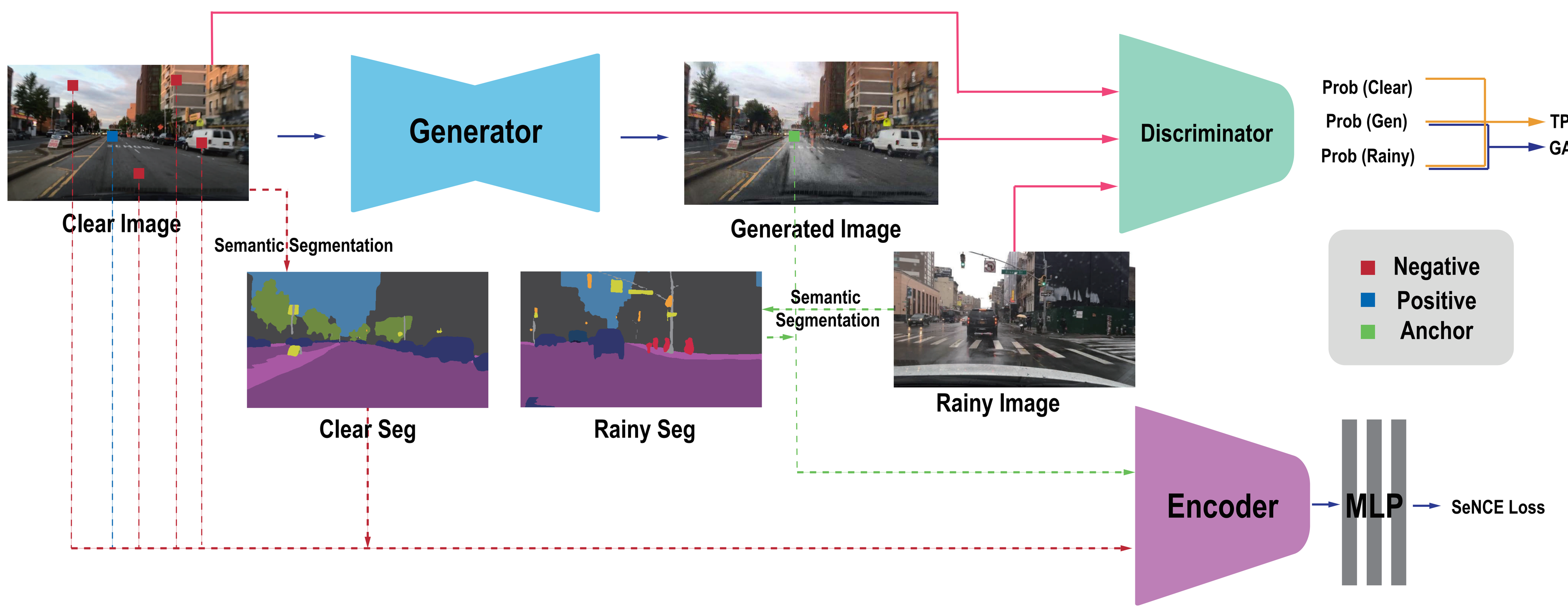


Figure 2. TPSeNCE uses a generator to translate clear images to rainy ones, a discriminator with TPS and GAN losses, and an encoder that embeds patches from both clear and generated images. MLPs process these patches contrastively to output SeNCE loss, guided by semantic segmentation maps.

Proposed Method (Continue)

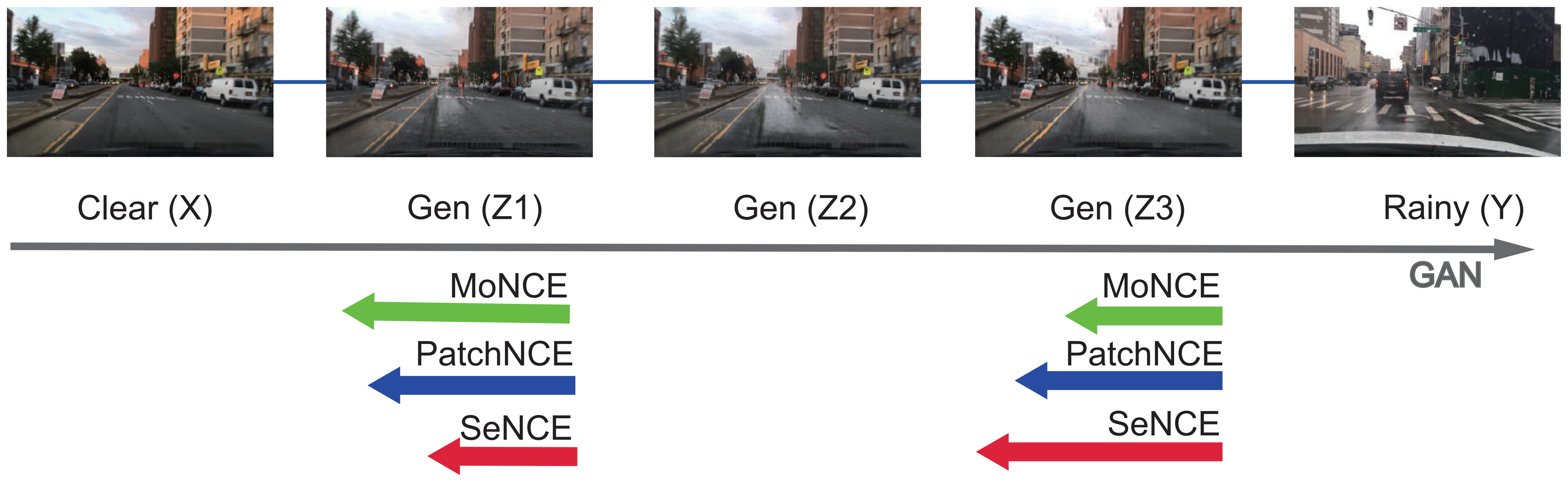


Figure 3. SeNCE outperforms PatchNCE and MoNCE in optimizing the amount of rain to produce realistic rainy images. The length of the arrow here represents the magnitude of the NCE losses.

Mathematical Details

Suppose X is the clear image, Y is the rainy image, and Z is the generated rainy image.

1. Triangular Probability Similarity (TPS)

$$\mathcal{L}_{\text{TPS}}(X,Y,Z) = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W}\left(|D(X)_{i,j} - D(Z)_{i,j}| + |D(Y)_{i,j} - D(Z)_{i,j}| - |D(X)_{i,j} - D(Y)_{i,j}|\right) \quad (1)$$

2. Semantic Noise Contrastive Estimation (SeNCE)

$$\mathcal{L}_{\text{SeNCE}}(X,Y,Z) = -\sum_{i=1}^{N}\log\frac{e^{\frac{x_i \cdot z_i}{\tau}}}{e^{\frac{x_i \cdot z_i}{\tau}} + Q(N-1)\sum_{j=1, j\neq i}^{N} w_{ij}\cdot e^{\frac{x_i \cdot z_j}{\tau}}}$$

$$w_{ij} = \text{softmax}\left(\frac{F(i,j)}{\beta}\right)_j \quad (2)$$

$$F(i,j) = (1 - mPA(X,Y))(x_i \cdot z_j) + (mPA(X,Y))(1 - x_i \cdot z_j)$$

Qualitative Comparisons

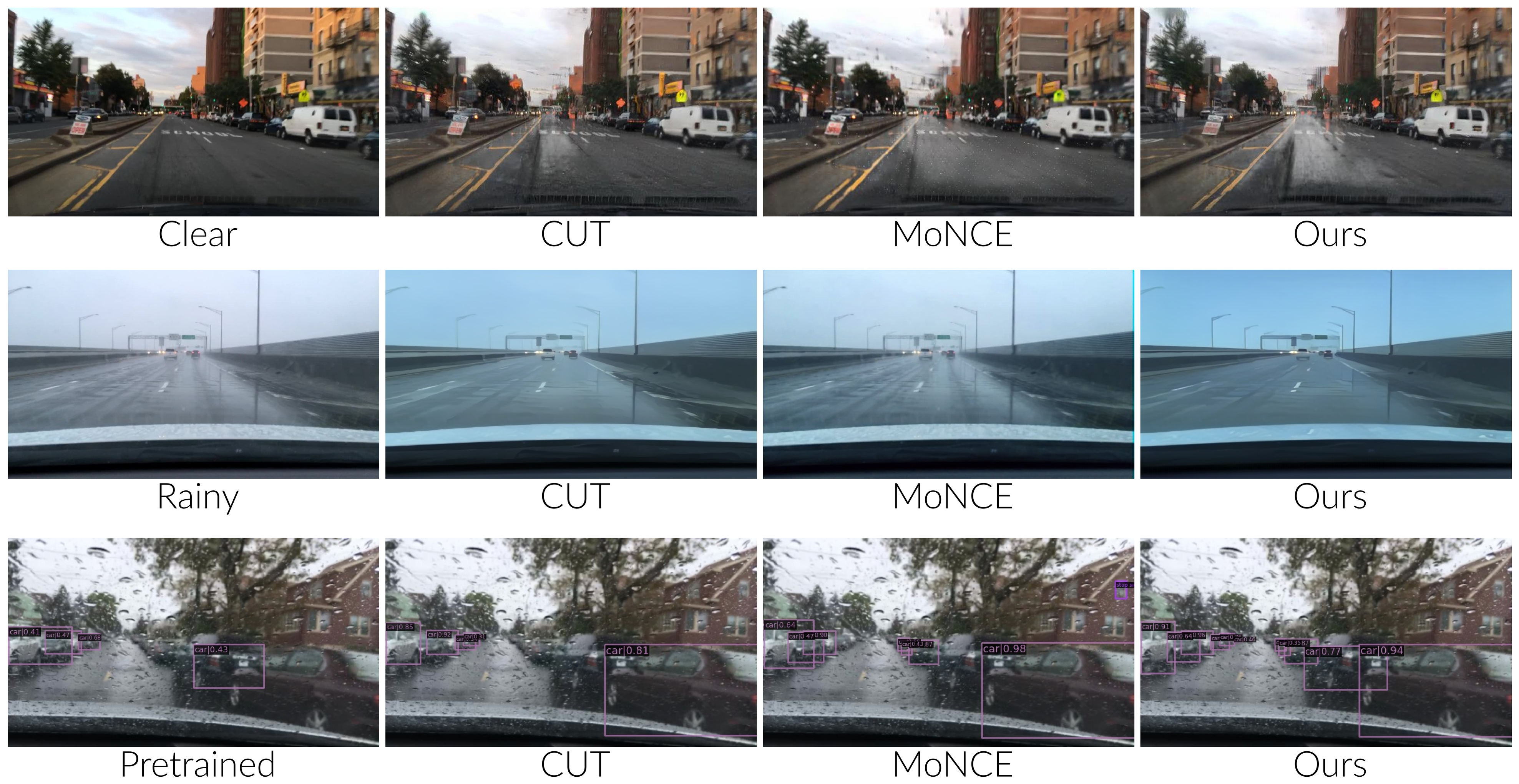


Figure 4. Qualitative comparison on rain generation (top), deraining (middle), and detection (bottom).

Figure 5. Qualitative comparison on clear2snowy (top), and day2night (bottom) translation.

Acknowledgements

This project is supported by NSF CPS Grant CNS-2038612, a DOT RITA Mobility21 Grant 69A3551747111 and by General Motors Israel.