JAN 4-8 WACV 2024 WAIKOLOA HAWAII

# 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

# Outline

Challenges

Workflow

TPS

SeNCE

Experiments

Conclusion

# Challenges for Detection in Rain

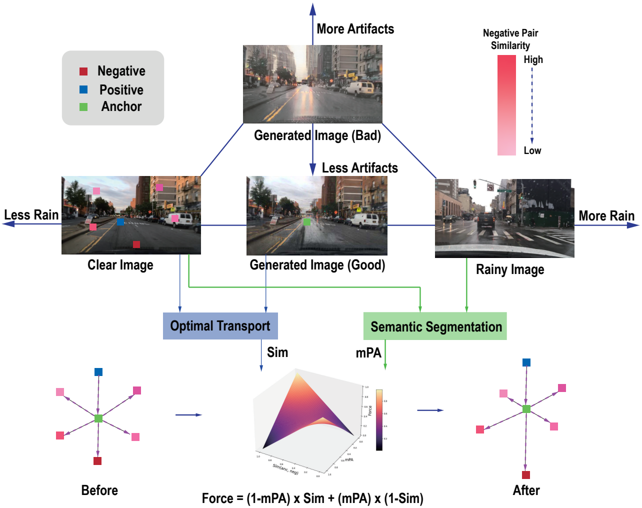
# Challenges for Rain Generation

# Challenges for Rain Generation

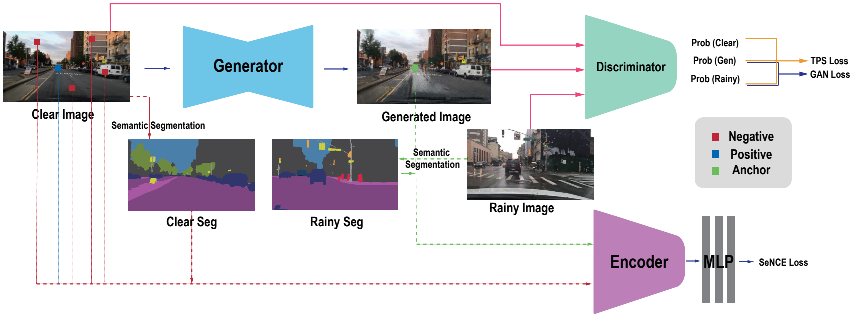
- No Proper Constraint: Artifacts and Distortions

- No Effective Strategy: Unrealistic Rain Amount

# Brief Workflow

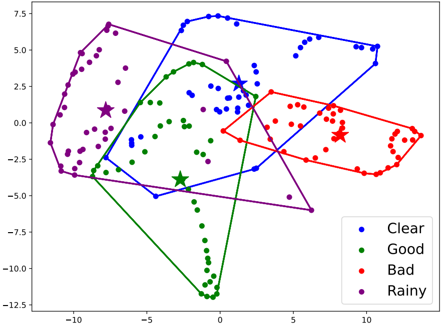


# Detailed Workflow



# TPS: Motivation

Clear (X) Good (Z+) Bad (Z-) Rainy (Y)




T-SNE

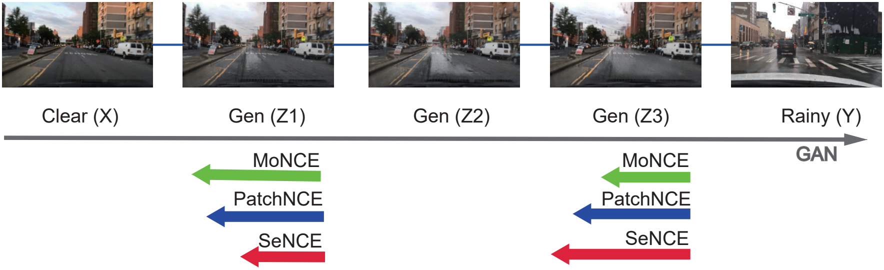
# TPS: Formulas

Triangular Probability Similarity (TPS) Loss:

$$\begin{aligned}\mathcal{L}_{\mathrm{TPS}}(\mathrm{X},\mathrm{Y},\mathrm{Z}) = \frac{1}{\mathrm{HW}}\sum_{\mathrm{i}=1}^{\mathrm{H}}\sum_{\mathrm{j}=1}^{\mathrm{W}}(&|\mathrm{D}(\mathrm{X})_{\mathrm{i,j}} - \mathrm{D}(\mathrm{Z})_{\mathrm{i,j}}| \\ &+ |\mathrm{D}(\mathrm{Y})_{\mathrm{i,j}} - \mathrm{D}(\mathrm{Z})_{\mathrm{i,j}}| \\ &- |\mathrm{D}(\mathrm{X})_{\mathrm{i,j}} - \mathrm{D}(\mathrm{Y})_{\mathrm{i,j}}|)\end{aligned} \tag{1}$$

Where $\mathrm{H}$ and $\mathrm{W}$ represent the height and width of the probability matrix $\mathrm{D}(\mathrm{Z})$.

# SeNCE: Motivation



# SeNCE: Formulas

Semantic Noise Contrastive Estimation (SeNCE)

$$\mathcal{L}_{\mathrm{SeNCE}}(X, Y, Z) = -\sum_{i=1}^{N} \log \frac{e^{\frac{x_i \cdot z_i}{\tau}}}{e^{\frac{x_i \cdot z_i}{\tau}} + Q(N-1)\sum_{\substack{j=1 \\ j \neq i}}^{N} w_{ij} \cdot e^{\frac{x_i \cdot z_j}{\tau}}} \tag{2}$$

Where

$$w_{ij} = \mathrm{softmax}\left(\frac{F(i,j)}{\beta}\right)_j \tag{3}$$

Where

$$F(i,j) = (1 - mPA(X,Y))\,(x_i \cdot z_j) + (mPA(X,Y))(1 - x_i \cdot z_j) \tag{4}$$

# Quantitative: Clear2Rainy (Rain Generation)

| Methods | BDD100K Dataset (clear → rainy) | | | | INIT Dataset (clear → rainy) | | | |
|---|---|---|---|---|---|---|---|---|
| | **Content↑** | **Style↑** | **KID↓** | **FID↓** | **Content↑** | **Style↑** | **MMD↓** | **ED↓** |
| UNIT [4] | 3.24 | 3.48 | 88.85 | **18.099** | 2.58 | 2.66 | 34.231 | 35.702 |
| MUNIT [2] | 2.44 | 2.80 | 189.12 | 26.538 | 2.80 | 2.72 | 34.425 | 36.458 |
| CUT [5] | 3.32 | 3.38 | 85.29 | 21.901 | 3.16 | 2.90 | 33.704 | 34.777 |
| QS-Attn [1] | 3.34 | 3.58 | 85.59 | 21.614 | 2.46 | 2.66 | 33.836 | 34.853 |
| MoNCE [8] | 3.10 | 3.30 | 75.66 | 18.595 | 2.18 | 2.24 | 33.579 | 34.814 |
| Ours | **3.58** | **3.70** | **72.19** | 19.341 | **3.42** | **3.04** | **33.535** | **34.774** |

Table 1: Quantitative comparison for image rain generation at BDD100K test clear and INIT test clear. KID is in $10^{-4}$, and MMD is in $10^{-5}$.

# Quantitative: Rainy2Clear (Deraining)

| **SAPNet trained on** | **BDD100K Dataset (rainy → clear)** | | | |
|---|---|---|---|---|
| | **Quality↑** | **Perf↑** | **MUSIQ↑** | **DBCNN↑** |
| Rain100H [7] | 3.20 | 2.08 | 53.267 | 46.585 |
| UNIT [4] | 3.37 | 3.49 | 56.036 | 45.912 |
| MUNIT [2] | 3.04 | 2.20 | 55.074 | 41.923 |
| CUT [5] | 2.88 | 3.92 | 59.974 | 54.491 |
| QS-Attn [1] | 2.94 | 3.22 | 61.064 | 57.311 |
| MoNCE [8] | 2.78 | 2.37 | 48.952 | 37.083 |
| Ours | **3.53** | **4.33** | **61.853** | **57.826** |

Table 2: Quantitative comparison for image deraining at BDD100K test rainy. SAPNet [9] is used for deraining, and it is trained on images from different rain generation methods.

# Quantitative: Object Detection in Rain

| Pretrained | Finetuned | mAP | mAP_50 | mAP_75 | mAP_s | mAP_m | mAP_l |
|---|---|---|---|---|---|---|---|
| COCO | None | 0.171 | 0.373 | 0.134 | 0.093 | 0.300 | 0.409 |
| | Clear | 0.231 | 0.492 | 0.190 | 0.151 | 0.366 | 0.454 |
| | UNIT | 0.215 | 0.462 | 0.178 | 0.113 | 0.404 | 0.546 |
| | MUNIT | 0.178 | 0.394 | 0.147 | 0.108 | 0.291 | 0.438 |
| | CUT | 0.243 | 0.512 | 0.203 | 0.149 | 0.410 | 0.548 |
| | QS-Attn | 0.248 | 0.513 | 0.213 | 0.151 | 0.401 | 0.574 |
| | MoNCE | 0.246 | 0.510 | 0.211 | 0.151 | 0.400 | 0.570 |
| | Ours | **0.262** | **0.526** | **0.237** | **0.158** | **0.419** | **0.646** |

Table 3: Quantitative comparison for Yolov3 [6] object detection at BDD100K test rainy. The Yolov3 object detector are pretrained on COCO [3], and finetuned on generated rainy images from different rain generation methods. mAP is computed on the most challenging 100 rainy images.

# Qualitative: Clear2Rainy (Rain Generation)

Clear

MUNIT [2]

CUT [5]

Ours

# Qualitative: Rainy2Clear (Deraining)

Rainy

Rain100H [7]

CUT [5]

Ours

# Qualitative: Object Detection in Rain


Pretrained


Clear


CUT [5]


Ours

# Qualitative: Day2Night


Clear

CUT [5]

QS-Attn [1]

Ours

# Qualitative: Clear2Snowy

Clear

CUT [5]

MoNCE [8]

Ours

# Conclusion

### Summary

- TPSeNCE: Unpaired Image-to-Image Translation
- Rain Generation benefits Deraining and Detection
- Generalize to Snowy and Night

### Contributions

- TPS: Suppress Artifacts and Distortions
- SeNCE: Optimize the Amounts of Generated Rain

### Future Works

- Extremely Heavy Rain
- Strong Light Sources
- Diffusion Models

# References I


Xueqi Hu, Xinyue Zhou, Qiusheng Huang, Zhengyi Shi, Li Sun, and Qingli Li.
Qs-attn: Query-selected attention for contrastive learning in i2i translation.
In *CVPR*, 2022.

Xun Huang, Ming-Yu Liu, Serge Belongie, and Jan Kautz.
Multimodal unsupervised image-to-image translation.
In *ECCV*, 2018.

Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick.
Microsoft coco: Common objects in context.
In *ECCV*, 2014.

Ming-Yu Liu, Thomas Breuel, and Jan Kautz.
Unsupervised image-to-image translation networks.
*NeurIPS*, 2017.

Taesung Park, Alexei A Efros, Richard Zhang, and Jun-Yan Zhu.
Contrastive learning for unpaired image-to-image translation.
In *ECCV*, 2020.

Joseph Redmon and Ali Farhadi.
Yolov3: An incremental improvement.
*arXiv*, 2018.

# References II


Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan.
Deep joint rain detection and removal from a single image.
In *CVPR*, 2017.

Fangneng Zhan, Jiahui Zhang, Yingchen Yu, Rongliang Wu, and Shijian Lu.
Modulated contrast for versatile image synthesis.
In *CVPR*, 2022.

Shen Zheng, Changjie Lu, Yuxiong Wu, and Gaurav Gupta.
Sapnet: Segmentation-aware progressive network for perceptual contrastive deraining.
In *WACV*, 2022.

# Thank you for your attention!

**This project is sponsored by General Motors Israel.**

**Scan the QR Code to access our Github webpage.**