Supplementary Material for "Uncertainty-Based Thin Cloud Removal Network via Conditional Variational Autoencoders"

Haidong Ding, Yue Zi, and Fengying Xie(✉)

School of Astronautics, Beihang University, Beijing 100191, China
{dinghaidong, ziyue91, xfy_73}@buaa.edu.cn

1 Variational Lower Bound of the Proposed Model

In this section, we provide a derivation for variational lower bound of the proposed model (Equation (3) in the main paper):

$$\begin{aligned}
\log p_\theta(Y|X) &= \int\limits_z q_\phi(z|X)\log p_\theta(Y|X)dz \\
&= \int\limits_z q_\phi(z|X)[-\log p_\theta(z|X,Y) + \log p_\theta(Y,z|X)]dz \\
&= \int\limits_z q_\phi(z|X)[-\log \frac{p_\theta(z|X,Y)}{q_\phi(z|X)} + \log \frac{p_\theta(Y,z|X)}{q_\phi(z|X)}]dz \\
&= D_{KL}(q_\phi(z|X)||p_\theta(z|X,Y)) + E_{q_\phi(z|X)}[\log \frac{p_\theta(Y,z|X)}{q_\phi(z|X)}] \\
&\geq E_{q_\phi(z|X)}[-\log q_\phi(z|X) + \log p_\theta(Y,z|X)] \\
&= E_{q_\phi(z|X)}[-\log q_\phi(z|X) + \log(p_\theta(z|X)\cdot p_\theta(Y|X,z))] \\
&= E_{q_\phi(z|X)}[-\log q_\phi(z|X) + \log p_\theta(z|X)] + E_{q_\phi(z|X)}[\log p_\theta(Y|X,z)] \\
&= -D_{KL}(q_\phi(z|X)||p_\theta(z|X)) + E_{q_\phi(z|X)}[\log p_\theta(Y|X,z)]
\end{aligned}$$

2 A New Dataset for Thin Cloud Removal:T-CLOUD

There are few available datasets for thin cloud removal from remote sensing images, so the proposal of T-CLOUD can promote the research and development of thin cloud removal algorithms.

T-CLOUD is a large-scale benchmark dataset for single remote sensing image de-clouding. Our dataset is different from the existing dataset RICE-I [1] in the following points: (1) T-CLOUD is a large-scale dataset. Our dataset contains 2939 doublets of cloud images and their clear counterpart while RICE-I only contains 500 image pairs. The large-scale dataset can effectively improve the performance of the thin cloud removal algorithms. (2) The ground scenes in our dataset have much finer texture details. T-CLOUD includes many different

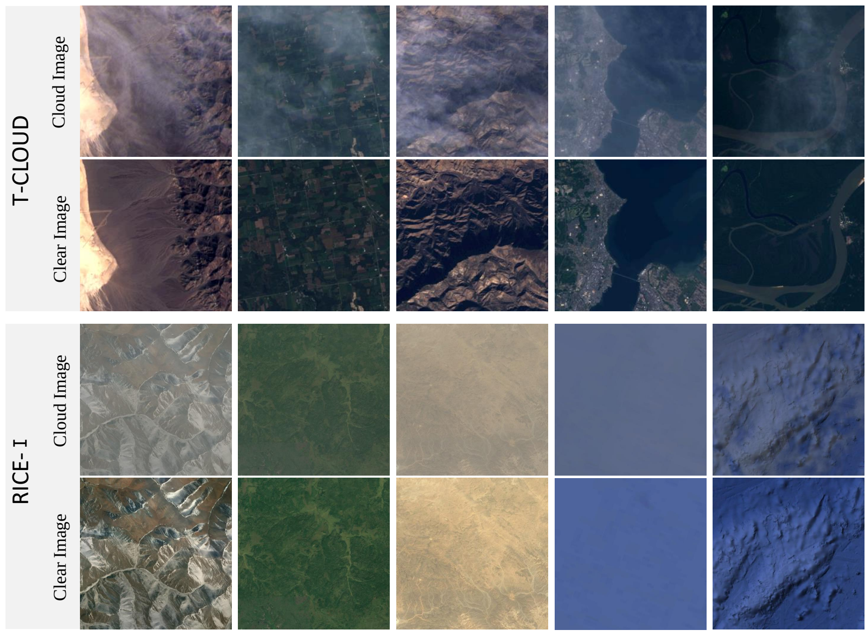

Fig. 1. Several examples in T-CLOUD.

ground scenarios such as cities, rivers, and deserts while RICE-I is relatively simple. (3) The thin clouds exhibited by T-CLOUD are non-homogeneous which is consistent with the characteristics of remote sensing images occluded by thin clouds. Figure 1 shows some visual comparisons between T-CLOUD and RICE-I, it can be observed that our dataset is more realistic. Table 1 summarizes the similarities and differences between the two datasets.

Table 1. Comparisons with RICE-I dataset. Our dataset contains much more image pairs than RICE-I.

Dataset	Amount	Content of Image	Image Size	Cloud Shape	Type
RICE-I	500	cloud/cloud-free	512 x 512	homogeneous	pair
T-CLOUD	2939	cloud/cloud-free	256 x 256	non-homogeneous	pair

3 Additional Experimental Results

In this section, we show additional experimental results on the proposed dataset compared to several existing methods. As shown in Figure 2, our proposed algorithm restores more image details and performs better visual effects.

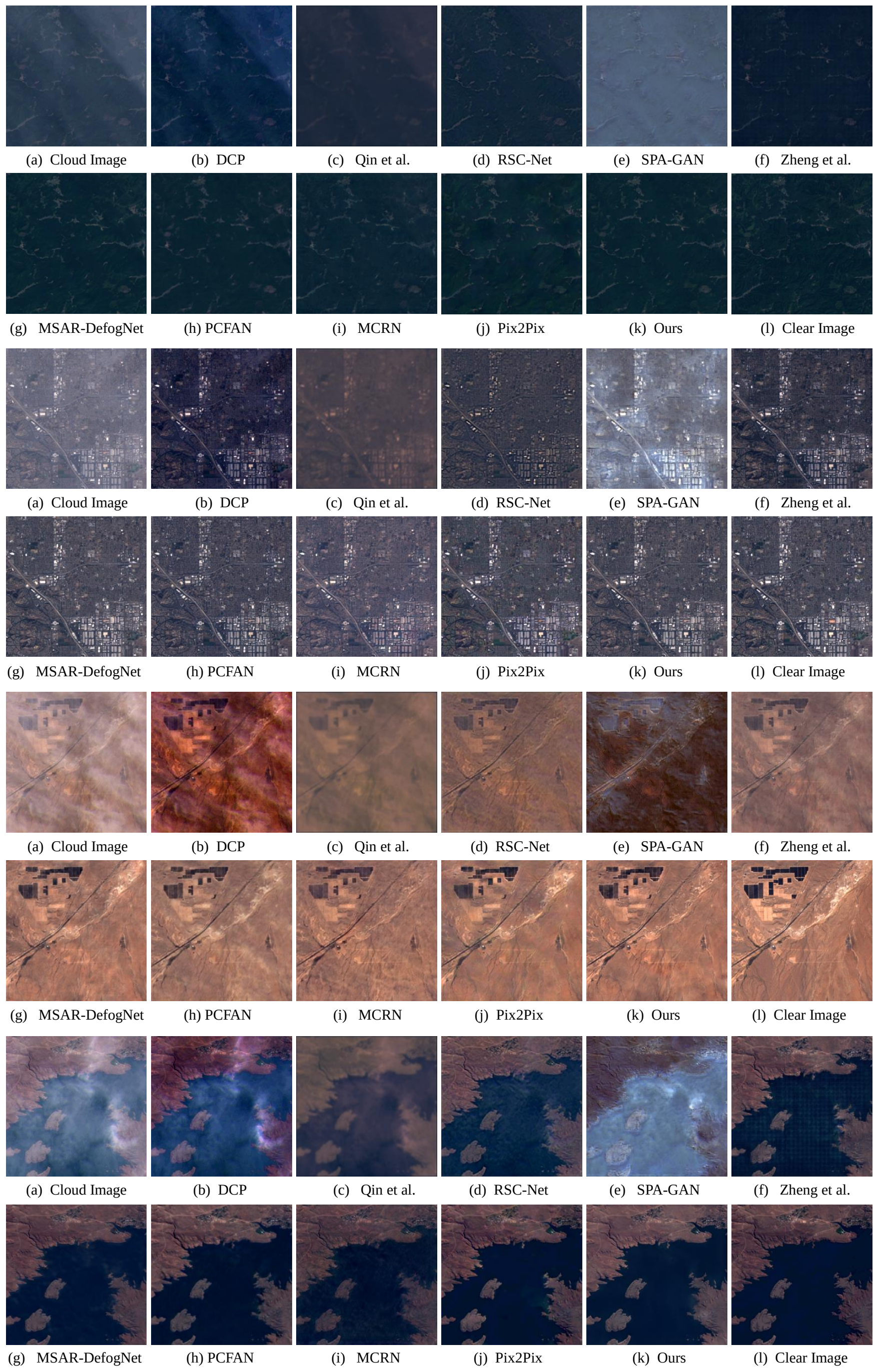

Fig. 2. Visual comparisons on the real-world cloud images. Zoom in for a better view.

References

1. Lin, D., Xu, G., Wang, X., Wang, Y., Sun, X., Fu, K.: A remote sensing image dataset for cloud removal. arXiv preprint arXiv:1901.00600 (2019)