HYBRID GAN AND SPECTRAL ANGULAR DISTANCE FOR CLOUD REMOVAL

Omid Ghozatlou[1], *Mihai Datcu*[1,2]

[1]Research Center for Spatial Information (CEOSpaceTech), University POLITEHNICA of Bucharest (UPB), Bucharest, Romania
[2]Earth Observation Center (EOC), German Aerospace Center (DLR), Oberpfaffenhofen, Germany

ABSTRACT

This paper aims to present a new algorithm to remove thin clouds and retain information in corrupted images without the use of auxiliary data. By injecting physical properties into the cycle consistent generative adversarial network (GAN), we were able to convert a cloudy multispectral image to a cloudless image. To recover information beneath clouds and shadows we create a synthetic multispectral space to obtain illumination invariant features. Multispectral vectors were transformed from Cartesian coordinates to Polar coordinates to obtain spectral angular distance (SAD) then we employed them as input to train the deep neural network (DNN). Afterward, the outputs of DNN were transformed to Cartesian coordinates to obtain shadow and cloud-free multispectral images. The proposed method, Hybrid GAN-SAD yields trustworthy reconstructed results because of exploiting transparent information from certain multispectral bands to recover uncorrupted images.

Index Terms— Cloud Removal, Generative Adversarial Networks (GANs), Polar Coordinates, Multispectral Satellite Images.

1. INTRODUCTION

Remote Sensing (RS) imagery data provides an excellent opportunity in Earth Observation (EO) to analyze and obtain information to understand the Earth's resources and physical phenomena parameters. However, clouds cover more than half of the Earth, according to statistics. This is a common issue with optical RS images causing information to be obscured by clouds and their associated shadows. As a result, we cannot capture reliable information from corrupted images unless we use clear sky images for the same time that they are not available. Therefore, one reasonable solution is improving the networks by leveraging trustworthy and transparent physical properties. Our proposed method is exploiting spectral angular distance (SAD) to train cycle-consistent adversarial networks with illumination invariant features that is illustrated in Figure 1.

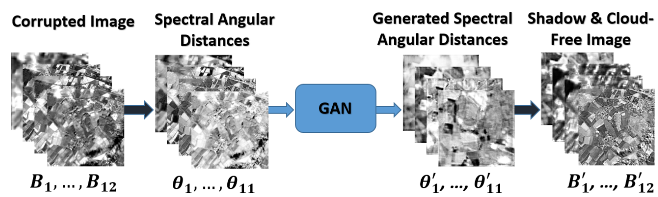

Fig. 1. Hybrid GAN-SAD removes cloud and shadows in corrupted images and recover underneath information in multispectral images using illumination invariant features in polar coordinate.

Since that, we need cloud-free images, many methods are proposed to detect and remove clouds from remote sensing images. By reviewing studies in cloud removal, we have grouped the methods into three categories. One category is multitemporal-based in which there are used multitemporal images of the same area. The drawback of this category is that on the one hand, the time interval of multitemporal images is long on the other hand the area is changing rapidly. For this reason, usually the accuracy of the reconstructed area is low. The second category is based on deep learning algorithms. Nowadays various kinds of generative adversarial networks (GANs) have been created for different applications. Therefore, many studies have been published for cloud removal using GANs. For example, [1] has proposed Multispectral conditional Generative Adversarial Networks (McGANs) for cloud removal from visible light RGB satellite images with multispectral images as inputs. The third category is based on multisensory data fusion in which is used auxiliary penetrable modalities. A helpful auxiliary data in this field consist in Synthetic Aperture Radar (SAR) images. Thanks to long-wavelength SAR can penetrate through clouds in different weather with different kinds of clouds. SAR-optical data fusion was proposed in [2] to remove clouds using cycle GAN. Also they computed cloud probability masks to specifically model cloud coverage explicitly while reconstructing cloud-covered information.

Unlike most computer vision algorithms for dehazing which are based on image enhancement, we inject physical properties into deep neural network to reconstruct the contaminated regions. Therefore, we use 12 Sentinel-2 bands in order to benefit spectral reflectance of multispectral images.

In this paper, we extend the proposed model in [3] based on cycle-consistent GAN to remove cloud from corrupted multispectral images. The most significant advantages of this method are the elimination of the need for paired images (cloudy/cloudless) and the use of an auxiliary modality that penetrates clouds. In contrast to [3] we exploit transparent information in the dataset by translating pixel values into polar coordinates, and then we train the network with illumination invariant features to reduce the impact of clouds and shadows. The proposed method, Hybrid GAN-SAD not only achieves notable results but also increases the trustworthiness of the network by utilizing reliable physical properties.

2. METHODOLOGY

The Hybrid GAN-SAD method is combing using polar coordinate transformation and training the cycle-consistent GAN by illumination invariant features to translate the cloudy images to the cloudless images that includes two mapping functions.

2.1. Spectral Angular distances (SAD)

We transform pixel values of Bottom of Atmosphere (BOA) reflectance images into the polar coordinates for both cloudy and cloudless domains. We train the networks based on illumination invariant features. We employ transformation of radiance values into polar coordinates that in [4] has been used for the scalable color descriptor. The polar feature space enables the network to recover information of the corrupted area and increase the robustness of the network. Figure 1 illustrates outline of the Hybrid GAN-SAD that each corrupted image includes 12 separate spectral bands that are shown as $B_1, B_2, ..., B_{12}$; and after polar coordinate transformation, we achieve 11 spectral angular distances that are defined as $\theta_1, \theta_2, ..., \theta_{11}$. The angular distances are calculated as follows:

$$\begin{gathered}\theta_1 = arctan\frac{\sqrt{B_{12}^2 + B_{11}^2 + ... + {B_3}^2 + {B_2}^2}}{B_1} \\ \theta_2 = arctan\frac{\sqrt{B_{12}^2 + B_{11}^2 + ... + {B_3}^2}}{B_2} \\ ... \\ \theta_{10} = arctan\frac{\sqrt{B_{12}^2 + B_{11}^2}}{B_{10}} \\ \theta_{11} = arctan\frac{B_{12}}{B_{11} + \sqrt{{B_{12}}^2 + {B_{11}}^2}}\end{gathered} \tag{1}$$

the inputs of networks include feature matrices with 128×128×11 dimensions for both cloudy and cloudless. The outputs of the GAN that are shown in Figure 1 as generated spectral angular distances are $\theta'_1, \theta'_2, ..., \theta'_{11}$. Finally, we obtain shadow and cloud-free multispectral images by employing the inverse of equation 1 that is formulated by the following equation to convert outputs of the GAN to Cartesian coordinates:

$$\begin{gathered}B'_1 = \rho cos(\theta'_1) \\ B'_2 = \rho sin(\theta'_1)cos(\theta'_2) \\ B'_3 = \rho sin(\theta'_1)sin(\theta'_2)\cos(\theta'_3) \\ ... \\ B'_{11} = \rho sin(\theta'_1)...sin(\theta'_{10})cos(\theta'_{11}) \\ B'_{12} = \rho sin(\theta'_1)...sin(\theta'_{10})sin(\theta'_{11})\end{gathered} \tag{2}$$

where ${B_1}', {B_2}', ..., {B_{12}}'$ are shadow and cloud -free spectral images and ρ is radial coordinate.

2.2. Cycle-Consistent Generative Adversarial Networks

The cycle-consistent GAN has been presented in [5] for learning to translate an image from source domain to a target domain in absence of paired examples. Our training dataset consists of 470 patches for cloudy domain ($c \in C$) and 1500 patches for cloud-free domain ($f \in F$). The generator G_{cf} translates images from cloudy images to the cloud-free domain and the generator G_{fc} translates cloud-free images to the cloudy domain. In addition, there are two discriminators D_c and D_f that distinguish real images from generated images by the mentioned generators. Figure 2 (left) illustrates the mapping function from a cloudy image to cloud-free. The translated image is used as input for the second generator G_{fc} to minimize cycle loss function (L_{cyc}) and vice versa for mapping the cloud-free into the cloudy that is shown in the Figure 2 (right). Furthermore, we feed the cloudy source images into the second generator to obtain an identical image is used to minimize the identity loss function (L_{idt}).

According to [5] the cycle loss function and identity loss function are formulated as:

$$L_{cyc} = ||G_{fc}(G_{cf}(c)) - c||_1 + ||G_{cf}(G_{fc}(f)) - f||_1 \tag{3}$$

$$L_{idt} = ||G_{cf}(c) - c||_1 + ||G_{fc}(f) - f||_1 \tag{4}$$

in addition, the negative log-likelihood objective by the least square performs more stably during training and generates higher quality results. We use the least square error function that is formulated as:

$$L_{adv} = [D_c(G_{cf}(c)) - 1]^2 + [D_f(G_{fc}(f)) - 1]^2 \tag{5}$$

by combining equations (3), (4), and (5), we get the final loss function formula

$$L_{all} = \lambda_{cyc}L_{cyc} + \lambda_{idt}L_{idt} + \lambda_{adv}L_{adv} \tag{6}$$

where all weights of the loss functions are set similarly to [5] i.e. $\lambda_{cyc} = 10, \lambda_{idt} = 1, \lambda_{adv} = 5$.

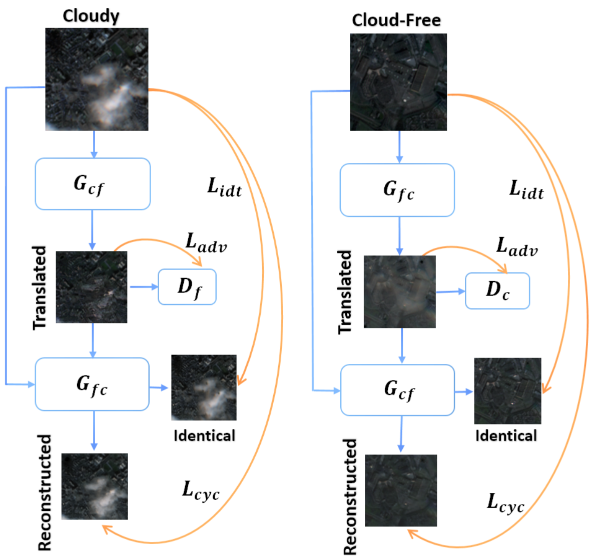

Fig. 2. Block diagram of cyclic GAN model. Two mapping networks and associated adversarial discriminators that are trained simultaneously to encourage more indistinguishable translation. To minimize the error of mappings, we use three different loss functions include L_{cyc}, L_{idt}, and L_{adv}.

3. RESULTS

Our dataset is composed of four scenes of Paris to compare our model with the model in [3]. We extracted the level-2A Sentinel-2 data acquired between February and March 2020. The land cover area includes urban area, agricultural and waterbody. We cropped them into 2400 patches by size 128×128 pixels for each band. For evaluation, we use the images of the same areas with 16 days temporal variation.The model was trained on an NVIDIA Quadro GV100 GPU with 32 GB memory. We use open-source code for cycle GAN that has been implemented using Keras library. We extend it to use for remote sensing data not only three visible bands but also Sentinel-2 12 bands.

We present the results of spectral angular distances obtained using polar coordinates transformation in Figure 3 in which are illustrated some samples of feature images in polar coordinate corresponding to multispectral band. The first row shows B4 (Red channel) samples including clouds and shadows and the second row shows the same scenes represented by angular distances. The pixel values of multispectral bands mapped in polar coordinate angles in range $[0, \frac{\pi}{2}]$.

For quantitative results, we assess the image quality by calculating the PSNR and SSIM. As a reference, we use clear sky patches from 19 March 2020 as a ground truth. Table 1 compares the average PSNR and SSIM for the considered strategies. In addition, the last column (Non-process) describes average of SSIM and PSNR that calculated between cloudy test dataset and ground truth. In all the cases, the haziness is removed more or less since the values are higher than SSIM: 0.7079 and PSNR: 22.6796 for Non-process. In terms of efficacy, RGB+IR thanks to the use of the penetrable near-infrared band, obtained the best results in both quantitative assessments, and then our proposed method using SAD has better results than full bands and the proposed method in [3].

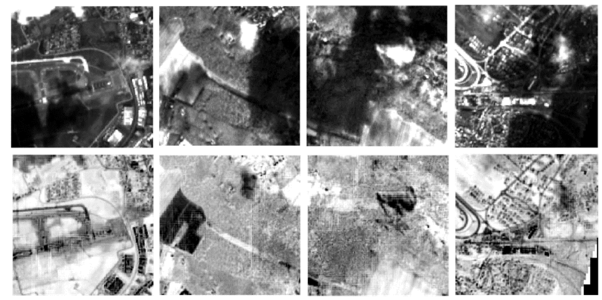

Fig. 3. Row I: Images corrupted by shadow and cloud in B4 Sentinel-2 and Row II: The angular distance (θ_4) corresponds to the above images

	RGB	RGB+IR	full bands	Hybrid GAN-SAD	Non-process
SSIM	0.7158	0.7558	0.7203	0.7398	0.7079
PSNR	26.2113	27.8445	26.0306	27.5246	22.6796

Table 1. Quantitative assessment (SSIM and PSNR) of the results of three different methods, proposed method using polar coordinates and test cloudy dataset.

For qualitative comparison of results, we train our network using three different strategies. First we use only RGB visible (B_4, B_3, B_2) bands of training dataset, second RGB and Near-Infrared (B_8) and finally full bands. Figure 4 shows the results of network using each methods. For visualization, we use only visible bands and increase the brightness of the figure for more clarity. On the first row there are samples of test images, on the second row is shown output of model in [3] then we added Near-Infrared bands and we got third row results. The forth row represents results using full bands of dataset. Furthermore, results of network that trained by converted images to polar coordinate (proposed method) are illustrated in fifth row. Finally, the last row shows cloud-free patches with 16 days different interval as a ground truth. As you can see, haziness could be removed easily in all methods (three last columns). Near-Infrared thanks to be thin cloud penetrable outperforms other methods for dehazing. Our proposed model has not only acceptable effect on cloud removal but also gain considerable impact to remove shadows and recover information beneath it that rest of other methods are disable to remove shadows (columns 2-7). By looking closer we find out corrupted patches by dense cloud is problem and recover all information of background without using auxiliary data such as SAR is impossible.

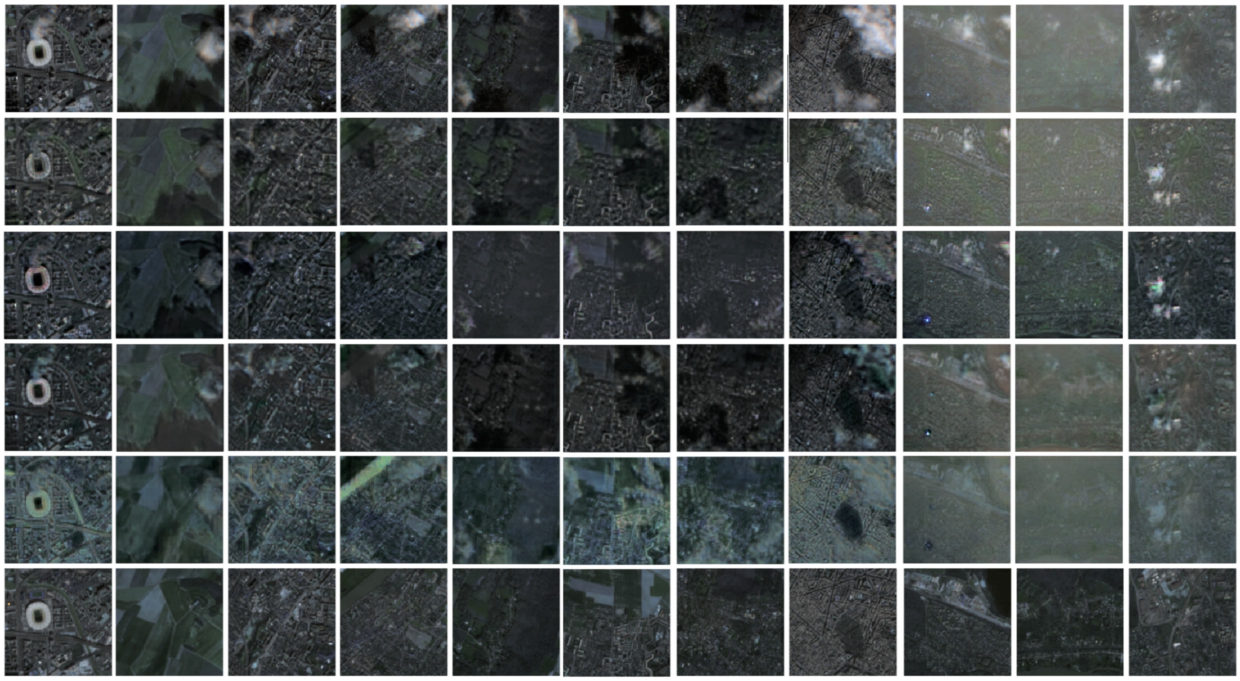

Fig. 4. Qualitative results, Row I: cloudy test samples, Row II: results of RGB [3], Row III: results of RGB+IR, Row IV: results of full bands, Row V: results of Hybrid GAN-SAD, Row VI: ground truth.

4. CONCLUSION

By comparing the results with clear sky images, we recognize that the model correctly detects clouds and tries to translate urban texture to the corrupted area. However, it is not enough because we need real information beneath clouds and shadows. Therefore, we proposed modified deep neural networks by injecting physical properties to achieve trustworthy results. As we recover real information of the background that has been blocked by the clouds, we can trust the outputs of the model. In summary, we conclude: First, the land cover distribution of the training dataset affects the learning of the model. Second, using angular distances provides physical information and an excellent solution to remove shadows by considering recover trustworthy information.

5. ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860370. Authors would like to thank Prof. Daniela Coltuc for her helpful scientific comments.

6. REFERENCES

[1] Kenji Enomoto, Ken Sakurada, Weimin Wang, Hiroshi Fukui, Masashi Matsuoka, Ryosuke Nakamura, and Nobuo Kawaguchi, "Filmy Cloud Removal on Satellite Imagery with Multispectral Conditional Generative Adversarial Nets," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2017-July, pp. 1533–1541, 2017.

[2] Patrick Ebel, Andrea Meraner, Michael Schmitt, and Xiao Xiang Zhu, "Multisensor Data Fusion for Cloud Removal in Global and All-Season Sentinel-2 Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 1, pp. 1–13, 2020.

[3] Praveer Singh and Nikos Komodakis, "Cloud-GAN: Cloud removal for sentinel-2 imagery using a cyclic consistent generative adversarial networks," *International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 2018-July, pp. 1772–1775, 2018.

[4] Florin Andrei Georgescu, Dan Răducanu, and Mihai Datcu, "New MPEG-7 Scalable Color Descriptor Based on Polar Coordinates for Multispectral Earth Observation Image Analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 987–991, 2017.

[5] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.