Cloud Removal from Satellite Images

Vaishnavi Pradip Mane
Department of Electronics and Telecommunication
MKSSS's Cummins College of Engineering for Women
Pune,India

Vaishnavi Raviraj Mane
Department of Electronics and Telecommunication
MKSSS's Cummins College Of Engineering for women
Pune, India

Aishwarya Anil Menon
Department of Electronics and Telecommunication
MKSSS's Cummins College of Engineering for women
Pune ,India

Aishwarya Purohit
Department of Electronics and Telecommunication
MKSSS's Cummins College of Engineering for women
Pune, India

Dr. Prof. Ashwini Deshpande
Department of Electronics and Telecommunication
MKSSS's Cummins College of Engineering for women
Pune ,India

Abstract— In the domain of satellite imagery analysis, the removal of cloud cover holds importance for accurate environmental monitoring and earth observation. Cloud occlusions can significantly hinder the extraction of ground information, limiting the efficacy of downstream analyses. Addressing this challenge, we present a novel approach leveraging Generative Adversarial Networks (GANs) with a UNet-based generator model for cloud removal. Our technique, applied to the dataset with cloud and cloud free images, a comprehensive spatiotemporal dataset covering diverse atmospheric conditions and geographic locations, demonstrates remarkable performance improvements over traditional methods. Through extensive experimentation, we showcase the effectiveness of our approach in generating high-quality, cloud-free images, thus facilitating more precise land cover classification and enhancing the utility of satellite imagery for environmental monitoring purposes.

Keywords— Cloud occlusions, Generative Adversarial Networks (GANs), UNet-based generator model

I . Introduction

Monitoring the surface of earth from space is crucial for addressing climate change, managing natural resources, and mitigating disasters. Remote Sensing imagery plays a vital role in these endeavors. However, cloud cover often obstructs the view, presenting a significant challenge. Satellites capture multispectral images across visible and infrared bands, providing a wealth of information. Roughly half of the Earth's surface is enveloped in dense cloud formations, with an extra 20% experiencing the presence of either cirrus clouds or thin layers of cloud cover. In some regions, cloud cover persists for months, hindering accurate observation. Numerous efforts have been made to address this issue, ranging from traditional methods to machine learning (ML) approaches.

Conventional methods for removing cloud occlusions typically rely on manual filtering techniques such as mean and median filters. These filters are applied to a large volume of images taken over a specific area to generate a background image. For example, a study utilized Sentinel-2 images captured every 6-7 days over a three-month period [6]. However, these composite approaches necessitate a substantial number of mostly cloud-free images collected over an unchanging landscape, limiting their practicality and applicability. Moreover, they are impractical in scenarios where the landscape undergoes gradual changes over time, as older images may not accurately represent the current environment. Additionally, these methods struggle to generate realistic ground representations in areas where cloud cover persists. An alternative approach involves utilizing cloud-detection algorithms to identify cloudy regions and fill or reconstruct them. Nonetheless, this inpainting technique overlooks sources of partial information, such as partially visible areas in shadow and the inherent transparency of clouds.

In the realm of domain translation, deep generative models have emerged as powerful tools, offering solutions to a variety of image transformation challenges. We are specifically focusing on the paired image translation approach for cloud removal in satellite imagery. This means that each cloudy satellite image directly corresponds to a cloud-free version of the same scene. To tackle this problem, we are adopting the Pix2Pix model, which has proven to be highly effective in similar tasks. The Pix2Pix model utilizes a conditional-GAN (cGAN) framework, combined with a U-Net generator architecture, to achieve state-of-the-art results across various domains. By leveraging this approach, we aim to develop a robust solution that accurately removes cloud occlusions from satellite images, ultimately enhancing their usability for environmental monitoring and earth observation purposes.

II . Literature Review

In paper [1] Recent research introduces a novel cloud removal technique for Sentinel-2 imagery, bypassing the need for explicit paired datasets. This method, solely reliant on visible range images, eliminates synthetic dataset creation for enhanced realism. While effective, it acknowledges the necessity of additional high wavelength imagery for thick cloud scenarios, highlighting a potential avenue for future exploration. Notably, reported results showcase significant performance improvements in synthetic test scenes, validating the approach's efficacy.

In paper [2], The research paper focuses on enhancing cloud detection and removal techniques for satellite imagery using Sentinel Hub and Generative Adversarial Networks (GANs). They successfully implemented cloud detection using existing methods and improved upon them. Additionally, they employed GANs to generate cloud-free images from cloudy satellite images, enhancing the overall image quality.

In paper [3], The research introduces a groundbreaking framework utilizing deep generative models to produce

979-8-3503-6052-3/24/$31.00 ©2024 IEEE

cloud-free images from cloudy satellite data. They create extensive paired spatial and spatiotemporal datasets from Sentinel-2 imagery, the largest of their kind to date. Their novel generative architecture, STGAN, effectively generates realistic cloud-free images, surpassing current models, even in challenging terrains with thick cloud cover.

In paper [4], This work introduces SEN12MS-CR-TS, a novel multimodal and multitemporal dataset addressing optical satellite image reconstruction and cloud removal challenges. It proposes two models: a 3-D convolution neural network predicting cloud-free images from cloudy sequences, and a sequence-to-sequence translation model generating cloud-free time series from cloud-covered ones, both evaluated on SEN12MS-CR-TS, demonstrating the dataset's significance and the advantages of multimodal and multitemporal information in remote sensing applications.

In paper [5], It introduces TaylorGAN, a technique enhancing GAN training by customizing loss functions for each network through Taylor expansions. By optimizing these functions via multiobjective evolution, TaylorGAN mitigates mode collapse and instability, notably improving image quality and performance metrics on image-to-image translation tasks, promising advancements in GAN applications.

In paper [6], The letter outlines a methodology for creating a seamless cloud-free composite of Africa for 2016 using Sentinel-2A data at 10-meter resolution provided by the European Space Agency. The approach combines two robust time series methods, namely the "darkest pixel" and the "maximum Normalised Difference Vegetation Index (NDVI)," previously used with AVHRR data. This method is crucial for initial processing in various land applications.

In paper [7], This research paper gives detailed information about Unet architecture.

In paper [8], This paper provides an Image Cloud Removing Dataset which contains RICE1 and RICE2 dataset with cloudy image and cloud free image.

III. METHODOLOGY

1. Data Collection and Preprocessing

Acquired a RICE dataset [8] consisting of 500 paired satellite images, where each image includes a cloudy version and its corresponding cloud-free version of the same scene. Resize the images from their original dimensions of 512x512 to 256x256 pixels to reduce computational complexity and memory requirements. Applied standard preprocessing techniques such as normalization and augmentation to enhance the dataset's diversity and improve the model's generalization ability.

2. Model Selection

Considering paired image translation tasks and focusing on models capable of learning complex mappings between cloudy and cloud-free images, We Choose the Pix2Pix model as the primary architecture for cloud removal. The Pix2Pix model is a conditional GAN (cGAN) with a U-Net generator, which has shown promising results in paired image translation tasks.

3. Model Architecture

Implementing the Pix2Pix model architecture, consisting of a generator and a discriminator which is trained simultaneously in an adversarial manner.

The generator is based on the U-Net architecture, which is known for its ability to preserve spatial information and capture fine details. The generator comprises an encoder-decoder structure with skip connections, facilitating the generation of high-quality cloud-free images while retaining essential features from the input cloudy images. We configured the generator to accept 256x256 cloudy images as input and output corresponding 256x256 cloud-free images.

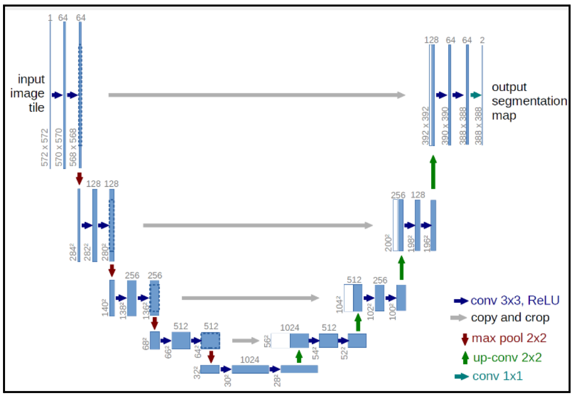

Fig. 1. Unet generator architecture[7]

We have Implemented a PatchGAN discriminator architecture, specifically tailored for the conditional Generative Adversarial Network (cGAN) framework. Trained the discriminator to discriminate between patches of real cloud-free images and patches of generated cloud-free images, providing detailed feedback to the generator during the training process. This fine-grained discrimination at the patch level enhances the generator's ability to produce high-quality cloud-free images that closely resemble the ground truth.

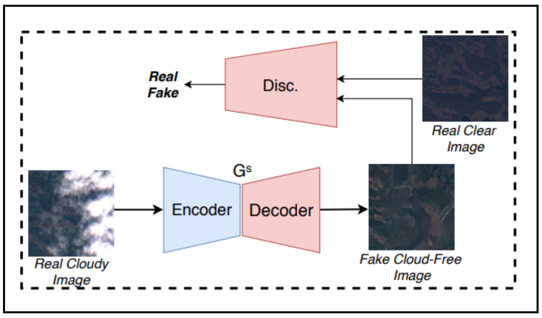

Fig. 2. Pipeline of architecture[3]

4. Training Procedure

The dataset comprising 500 paired satellite images is split into training and validation sets in an 80:20 ratio to ensure adequate data for both model training and validation. The Pix2Pix model is then trained on the training set, where batches of paired images are fed through the generator and discriminator networks. The generator aims to produce realistic cloud-free images from cloudy input images, while the discriminator distinguishes between real and generated images. Weight updates for both networks are computed based on losses using backpropagation and optimization algorithms such as Adam. Following training, the model's performance is assessed on the validation set to identify overfitting and fine-tune hyperparameters as needed, ensuring robust performance on unseen satellite images.

5. Evaluation

Evaluated the trained Pix2Pix model on a to assess its performance in removing cloud occlusions from satellite images. We have calculated the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) metrics to quantitatively measure the similarity between the generated and ground truth images. The PSNR value obtained is 27.54, indicating the quality of the generated images compared to the ground truth. Additionally, the SSIM value is 0.273, which also contributes to evaluating the fidelity of the generated images. Furthermore, qualitative analysis is conducted by visually inspecting the generated images to ensure they are free of cloud occlusions and preserve the essential features of the scene.

IV. Result

We present visual comparisons between the original ground truth images and the predicted images generated by our model. Through these comparisons, we aim to provide a clear visualization of the performance of our model in removing cloud occlusions from satellite images. Additionally, quantitative evaluation metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are calculated to assess the similarity between the predicted and ground truth images. Notably, our model achieves a PSNR value of 27.54 and an SSIM value of 0.273, indicating high fidelity and strong resemblance between the predicted and ground truth images. These results underscore the effectiveness of our model in producing cloud-free images that faithfully preserve the essential features of the scenes captured by satellite imagery.

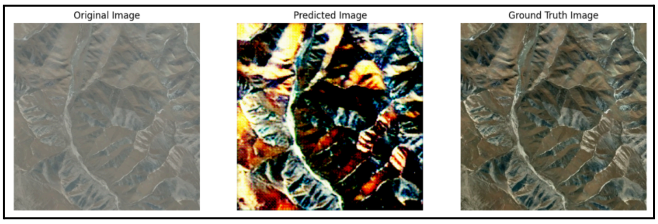

Fig. 3. Visualization of output

V. Conclusion

Our project employs the Pix2Pix model integrated with Generative Adversarial Networks (GANs) and the UNet architecture to successfully generate high-quality cloud-free images from cloudy satellite imagery. Leveraging this combined approach, we achieve promising results in effectively removing cloud occlusions from satellite images, thereby enhancing their usability for various applications in environmental monitoring and earth observation. The utilization of Pix2Pix with GAN and UNet demonstrates its efficacy in producing visually appealing and accurate cloud-free images, surpassing traditional methods and contributing to advancements in satellite image processing. These findings underscore the potential of our methodology to address challenges related to cloud cover in satellite imagery, ultimately paving the way for improved analysis and interpretation of satellite data for scientific research and practical applications.

VI. References

[1] Singh, Praveer, and Nikos Komodakis. ”Cloud-gan: Cloud removal for sentinel-2 imagery using a cyclic consistent generative adversarial network.” IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2018.

[2] Chauhan, R., Singh, A., & Saha, S. (2021). “Cloud Removal from Satellite Images”.Computing Science (Multimedia), University of Alberta, Edmonton, Canada.

[3] Sarukkai, V., Jain, A., Uzkent, B., & Ermon, S. (2020). "Cloud Removal in Satellite Images Using Spatiotemporal Generative Networks." In the Computer Science Department, Stanford University.

[4] Ebel, P., Xu, Y., Schmitt, M., & Zhu, X. X. (2020). "SEN12MS-CR-TS: A Remote-Sensing Data Set for Multimodal Multitemporal Cloud Removal." IEEE Transactions on Geoscience and Remote Sensing

[5] GAN Architecture from “Science Dictionary” https://www.sciencedirect.com/topics/computer-science/generative-adversarial-networks

[6] Fabrizio Ramoino, Florin Tutunaru, Fabrizio Pera, and Olivier Arino. Ten-meter sentinel-2a cloud-free composite—southern africa 2016. Remote Sensing, 9(7):652, 2017.

[7] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. Unet: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pages 234–241. Springer, 2015.

[8] Lin, Daoyu, et al. ”A remote sensing image dataset for cloud removal.” arXiv preprint arXiv:1901.00600 (2019).

Authors

Vaishnavi Pradip Mane

Final year student pursuing B.Tech. in Electronics and Telecommunication Engineering (ENTC) at Cummins College of Engineeri

Vaishnavi Raviraj Mane

Final year , Pursuing BTech in Electronics and Telecommunication from MKSSS Cummins college of Engineering for Women's Pune

Aishwarya Anil Menon

Final year , Pursuing BTech in Electronics and Telecommunication from MKSSS Cummins college of Engineering for Women's Pune

Aishwarya Purohit

Final year , Pursuing BTech in Electronics and Telecommunication from MKSSS Cummins college of Engineering for Women's Pune