Using Generative Adversarial Networks for Cloud Removal in Satellite Imagery

Aleena Varghese
Christ (Deemed to be University)
*D*epartment of Computer Science
aleena.varghese@msam.christuniversity.in

***Abstract*—Satellite imaging is one of the most common uses for applications agricultural, urban planning and environmental monitoring to mention a few. Unfortunately, even the best-laid plans for aerial photography can be decimated by one thing: cloud cover. The novel way of cloud extraction from the satellite data that is demonstrated in this article, use a Generative Adversarial Network (GAN). For the betterment of cloud removal, ResNet based discriminator and a UNet-based generator are utilized in the suggested approach. To accurately train the networks, a new technique was also developed to introduce noise that resembles natural cloud patterns. The PSNR score, as a qualitative and quantitative index card that uses the PyTorch-based GAN methodology to verify different performances in traditional methods based on EuroSat.**

***Index Terms*—Satellite Imagery, Cloud Removal, Generative Adversarial Networks (GANs), UNet, ResNet, Perlin Noise, Image Reconstruction, Peak Signal-to-Noise Ratio (PSNR).**

I. Introduction

Satellite images: Pictures of Earth captured by imaging satellites, either privately or publicly owned (or operated). In this regard, Satellite Imagery is one of the most important and powerful tools we have for earth observation. The changes in soil, water, air and vegetation are traced together with the global human fingerprint. Human activities could be measured, identified and monitored using the satellite pictures. Some of the applications for satellite images are: meteorology and weather forecasting, fisheries, oceanography, agriculture, geology mapping regional planning forestry conservation landscape analysis environmental evaluation intelligence conflict education. In the field of education, they are used to supplement text, maps and graphs by textbooks or internet.

Among the different types of reach capabilities, satellite imaging is utilized in various industries such as agriculture, urban planning and environmental monitoring. Cloud cover, however poses a big issue as it masks significant information and makes data analysis more challenging (Goodfellow et al) [1]. Traditional cloud removal techniques often require complex pre-processing steps and are time-consuming, or we need human interventions that may not work every time[5]. Recently Generative Adversarial Networks (GANs) has shown great promise as an automated approach for cloud removal in picture generation and enhancement [7].

Satellite imagery is a very valuable resource for spatial analysis and satellite images in GIS as they are high resolution imaging of earth surface at different geo location. Beyond what can be embodied in traditional paper maps and satellite images, the technology of sensors aboard satellites has also made it possible for us to conduct detailed analysis about both natural and artificial environmental conditions which could not at all have been scrutinized from ground. Remote sensing change analysis is the important stage of satellite image processing to check on time basis changes in environment, urban development and other dynamic events. With deep learning techniques, which automate the highly precise detection of complex patterns and temporal changes, strengthens this study.

Space observation and monitoring of Earth is essential for managing natural resources, preventing disasters, addressing climate change, and addressing other critical challenges of the twenty-first century. Recently, there has been a surge in the popularity of Earth observation satellites, which orbit the planet and collect data on its surface. Cirrus or thin clouds cover 20% of the Earth's surface, whereas impenetrable clouds cover around 55% of the planet's surface[7]. For uses in urban planning, agriculture, and environmental monitoring, cloud removal in satellite imagery—which encompasses a range of traditional techniques—is crucial. Histogram matching reduces the impact of cloud cover by aligning cloudy photographs with clear ones, while radiometric correction adjusts image properties.

Cloudy photos can be aligned with clear ones using a technique called histogram matching, while images can be altered using radiometric correction to lessen the appearance of clouds. Multi-temporal techniques, such image fusion and change detection, employ many images to identify and correct cloud-covered regions. Using spectral unmixing to separate pixel fingerprints, clouds are eliminated and segregated. Interpolation and inpainting techniques estimate or fill in cloud-covered pixels by using surrounding data. Using historical data, physical and empirical models simulate light interactions or account for cloud effects. Statistical methods such as PCA and regression analysis decrease data dimensionality and examine correlations in order to mitigate the influence of clouds. However, these approaches could require a lot of work, a sizable dataset, or they might not work as well in complex scenarios.

To create GANs, two neural networks—the discriminator and the generator—engage in a competitive game. The discriminator provides input to help the generator generate better results by comparing the artificial data—such as images—that it generates with real data. The article explores the application of Generative Adversarial Networks (GANs) to satellite imagery cloud removal. The work investigates how to effectively remove clouds and other obstructions from satellite photos using GANs, enhancing the clarity and quality of the images. The creation of generative adversarial networks (GANs), which produce visuals using deep learning, is Goodfellow's most notable invention. With this method, the quality of an image can be competitively enhanced by using two neural networks.

The research makes several significant breakthroughs in addition to introducing a unique GAN-based methodology for cloud removal and proving its efficacy over traditional methods. It offers a comprehensive review of the performance of this GAN model, emphasizing improvements in image quality and cloud-free reconstruction.

II. LITERATURE REVIEW

For the purpose of training generative models without Markov chains, a GAN framework is presented. Evaluations of the generated samples, both qualitative and quantitative, show the potential of this approach. Deep belief networks, Boltzmann machines, and indirect graphical models are explored in detail. Different training criteria and generative models with latent variables are compared. There includes mention of a number of methods, including generative stochastic networks, noise-contrastive estimating, and score matching. Algorithm 1's convergence produces a trustworthy estimator of p data[1]. For pg pdata, the global optimum is effectively reached. It is observed that when both the discriminative and generative models have adequate capacity, pg approaches pdata. An aggressive method using gradient-based learning rule updates that incorporate momentum is used during the training phase. We investigate undirected graphical models, such as latent variable DBMs and RBMs. The methodology is extended to include learnt approximation inference, conditional generative models, and other improvements.

The categorization of land use using Sentinel-2 satellite image analysis is covered in the EuroSAT study. The study reviews a number of datasets and techniques for picture classification in land use studies. Using a unique dataset, the study reported an outstanding 98.57% classification accuracy.[2] When it came to categorization accuracy, the RGB band combination outperformed the other band combinations. Deep Convolutional Neural Networks (CNNs) were trained in the research either from scratch or by fine-tuning them with pre-existing networks. For the study, a new dataset comprising 13 spectral bands of Sentinel-2 satellite photos was created. Additionally, benchmarks employing sophisticated deep CNNs were provided by the study for analysis comparison.

When it comes to the removal of sparse clouds by transfer learning, YUV-GAN performs better than other methods. With its impressive performance in the cloud removal sector, YUV-GAN has become a leader in this industry. When faced with a shortage of training data, the model performs better when transfer learning is used. To break down the elements of the network architecture and training plan, an ablation research was carried out.[3] YUV-GAN's network architecture is specifically designed to eliminate sparse clouds from satellite pictures. The training strategy is based on transfer learning and is tailored for foggy images found in the real world.

For realistic computer-generated imagery (CGI), solid textures and stochastic effects are added using the Pixel Stream Editor. To help create CGI algorithms that are quick, realistic, and parallelized, a new paradigm is presented. Solid texture use enables the CGI programs to simulate many types of materials. A novel method for using scalar-valued functions to create stochastic forms is provided. The technique used ensures coherence and saves time by stopping at the pixel level. This paradigm goes beyond the limitations of conventional texture mapping techniques.[4]A simple method to handle rotational invariant and bandwidth constraints is described. The technique makes it possible to create realistic computer-generated imagery (CGI) of stars, fire, clouds, and water. At the pixel level, the algorithms demonstrate speed, realism, and parallelized. With nonlinear features in CGI design, the Pixel Stream Editor is intended. A paradigm for functional composition is put out to create both regular and stochastic structures. The algorithm runs independently at every sample point in order to maximize throughput. In CGI applications, texture and shape independence are achieved by the use of solid textures.

By using data augmentation, the U-Net architecture performs exceptionally well in the area of biomedical picture segmentation. In a modest amount of training time, this method produces outstanding results with a small number of annotated photos. Adding data augmentation turns out to be helpful in improving deep neural network training effectiveness. The U-Net design outperforms the sliding-window convolutional network in a number of segmentation tests. Additionally, the U-Net model achieves excellent results in segmentation tasks without requiring a lot of pre- or post-processing. When used for cell tracking and segmentation tasks, it produces better results. Furthermore, the U-Net framework uses an overlap-tile method to allow for the smooth segmentation of big pictures[5]. The U-Net model performs better than earlier approaches with an astonishingly low warping error of 0.0003529. It clearly outperforms other algorithms in segmentation tasks, demonstrating its supremacy. On 512x512 images, the U-Net model likewise demonstrates quick segmentation abilities, doing the job in less than a second. The key to maximizing the use of annotated samples in segmentation tasks is to leverage data augmentation. The U-Net architecture participates in challenges like the ISBI cell tracking challenge and is efficiently applied to tasks involving the segmentation of neural structures.

To get higher accuracy in picture recognition, deep residual learning is presented. Residual functions are used to explore optimization problems in deep networks. The system wins both the COCO 2015 and ILSVRC contests. In deep learning models, shortcut links are used to expedite optimization. In order to re-frame problems at several scales and accelerate convergence, multi-grid approaches are applied. For classification challenges, residual representations like VLAD and Fisher Vector show great potential. Applying residual nets to the ImageNet test dataset yields an error rate of 3.57%. ResNets overcome optimization challenges with efficacy and show increasing accuracy with depth[6]. These nets have smaller reactions and are typically located nearer zero. A residual learning framework with shortcut connections for deep neural networks is established. Plain residual nets follow particular design ideas and are inspired by VGG nets. To address the deterioration problem, identity mapping and a customized residual function F are included. To aid in learning processes, building blocks are produced using residual mapping F(x, Wi).

RF signals are used to recover cloud-obscured pixels from satellite photos using a neural network. This method outperforms baseline techniques by 8dB in Peak Signal-to-Noise Ratio (PSNR), which will subsequently help with monitoring of wildfires, floods, and agriculture, among other sectors. This is accomplished by an optimization process that uses a matrix completion objective in conjunction with damped interpolation[7]. The system makes use of multi-spectral data from Sentinel-1 and Sentinel-2 as well as Synthetic Aperture Radar (SAR) and SpaceEye. According to experimental results, the system achieves an 8dB improvement in PSNR over baseline approaches. The neural network framework is intended for the reconstruction of multi-spectral images devoid of clouds, and the optimization technique is concentrated on accomplishing the matrix completion objective and damped interpolation.

III. METHODOLOGY

The Generative Adversarial Network (GAN) structure used for cloud removal in satellite imagery is the main focus of the research technique for the model in question. There are multiple stages to the methodology: preparation of the data, production of noise, construction of the model architecture, training, and evaluation.

Preparing Data Preparing the dataset for training and assessment is the first stage. The first thing the code sample does is organize data from satellite images using a directory structure. Two sets of photographs are categorized: all-inputs for the cloudy images with additional synthetic clouds and all-targets for the ground-truth cloud-free images. To guarantee a sequential naming scheme for simpler processing, the photos are renamed after being copied from the source dataset.

Generation Noise Perlin noise is used in satellite photography to imitate clouds. The generate-perlin-noise function generates a noise map with attributes that regulate the appearance and dispersion of noise patterns, including scale, octaves, persistence, and lacunarity. The generate-clouds function creates a realistic cloud pattern by utilizing many Perlin noise layers. The overlay-clouds function is then used to combine these clouds with the original photos, adjusting transparency to produce artificially clouded images for training.

Model Architecture Neural network models are utilized in the implementation of the generator and discriminator, which are the two main components of the design.

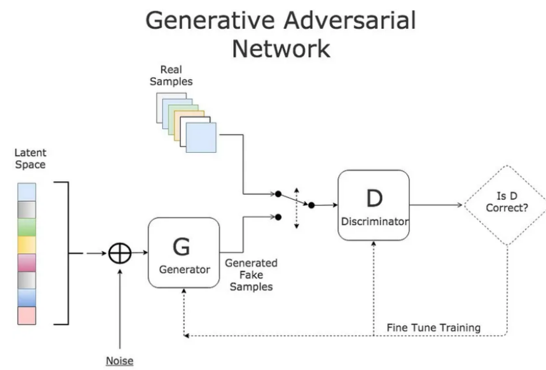

Fig. 1. GAN Model

*a) 1.**Generator (UNet):**The generator is built on top of the well-known U-Net architecture, which excels at problems involving image-to-image translation. The encoder-decoder structure is made up of many convolutional layers. The decoder uses skip connections to combine features from previous layers while increasing the size back to the original resolution after the encoder gradually reduces the image size while extracting features. This design facilitates the reconstruction of high-quality images and the capture of specific information. Cloudy inputs are converted into cloud-free images by the final layer, which produces an image with three color channels.:*

*b) 2.**Discriminator (ResNet):** A Residual Network (ResNet) is used as the discriminator to distinguish between artificially created and real images. It is made up of many residual blocks that are trained to recognize minute variations between generated and actual cloud-free images. By discovering residual mappings, residual blocks assist in preserving the gradient flow and enhancing model performance.:* **Training Process** An adversarial training process is used in the GAN configuration for the generator and discriminator. The idea is to train the discriminator to discern between created and genuine images, and to train the generator to produce realistic images free of clouds.

1.**Discriminator Training:** The discriminator is trained to distinguish between produced and genuine images. Binary cross-entropy (BCE) between the predicted labels and the genuine labels (real or fake) is used to calculate the discriminator's loss.By using backpropagation from the discriminator loss, the discriminator modifies its weights.

2.**Generator Training:** The generator is designed to create visuals that are convincing enough to trick the discriminator into thinking they are real. The generator's loss consists of two components: L1 loss (which promotes similarity between generated and target images) and BCE loss (which encourages the generator to create images categorized as real). The generator learns how to create crisp, cloud-free images by combining these losses.

Evaluation In the evaluation phase, a subset of the dataset is used to gauge how well the trained generator performs. Images

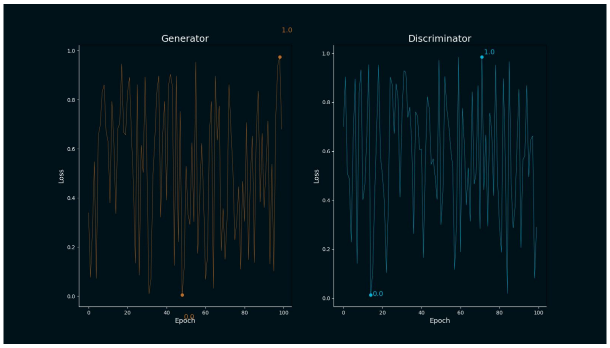

Fig. 2. Loss

from this subset were not viewed by the model during training. To measure the quality of generated images in comparison to ground-truth images, metrics like PSNR (Peak Signal-to-Noise Ratio) are computed. By comparing the input, target, and output images side by side and computing PSNR values to assess the fidelity of cloud removal, the model's performance is demonstrated.

The methodology describes a step-by-step process for creating and assessing a GAN-based model to remove clouds from satellite data. It includes rigorous training and evaluation procedures, the creation of synthetic clouds, sophisticated neural network architecture, and data processing. By efficiently eliminating clouds from satellite photos, the suggested methodology seeks to improve their quality, which will help a number of applications in urban planning, agriculture, and environmental monitoring.

IV. RESULTS

Fig:2 Loss consists of two plots that show the training loss of a Generative Adversarial Network (GAN) over 100 epochs. These plots are essential for understanding how well the GAN is learning to generate realistic images.

1. Generator Loss Plot (Left):

- The loss value of the generator, which shows how well it is producing realistic images, is represented on the Y-axis.The number of training epochs, which ranges from 0 to 100, is represented by the X-axis.
- The generator's loss during the training phase is depicted in the plot with notable variations. Due to the ongoing adversarial process between the discriminator and generator, these variations are typical of GAN training. While the discriminator attempts to enhance its capacity to discriminate between real and created images, the generator seeks to produce images that the discriminator is unable to identify from genuine ones. The oscillations in the loss levels are frequently caused by this antagonistic dynamic.

2. Discriminator Loss Plot (Right):

- Similar to the generator plot, the Y-axis represents the loss value of the discriminator.
- The X-axis represents the number of training epochs.

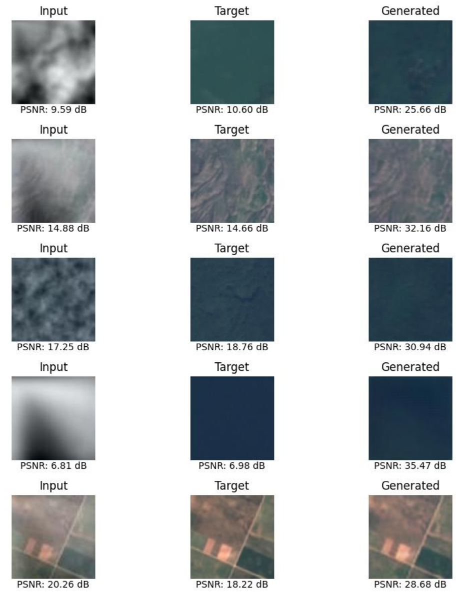

Fig. 3. Training results

- Over the course of the 100 epochs, there are also noticeable variations in the discriminator's loss. The discriminator's failure to accurately distinguish between produced and genuine images is reflected in its loss. Training increases the difficulty of the discriminator's duty as the generator gets better, which causes variations in its loss.

The image fig:3 Training Results, showcases the input images, target images, and the generated images by the GAN, along with their Peak Signal-to-Noise Ratio (PSNR) values.

1) **1. Input:**

- These are the original images fed into the GAN for processing.
- Each input image is accompanied by its PSNR value, which indicates the quality of the image in terms of signal-to-noise ratio. Higher PSNR values generally indicate better image quality.

1) **2. Target:**

- These are the ground truth images that the GAN aims to replicate. They serve as the reference standard for the generated images.
- The PSNR values for the target images are also provided for comparison.

1) **3. Generated:**

- These images are the output from the generator after being trained.
- The PSNR values for the generated images are displayed, indicating the similarity between the generated images and the target images. Higher PSNR values suggest that the generated images are more similar to the target images.

The PSNR values and visual comparison give information about the GAN's performance. The quality and precision of the created images in mimicking the genuine ones can be evaluated by comparing the input, target, and generated images side by side. The PSNR values provide further quantification for this comparison, which aids in assessing how well the GAN renders realistic images.

In conclusion, Figs. 2 and 3 taken together offer a thorough picture of the GAN's performance and training procedure. The test images and their PSNR values show how well the GAN generates realistic images, while the loss graphs show the adversarial dynamics between the generator and discriminator.

V. CONCLUSION

The paper presents a novel use of Generative Adversarial Networks (GANs) for cloud removal in satellite data. The suggested method greatly enhances the quality of satellite images, increasing their utility in environmental monitoring, urban planning, and agricultural by utilizing a ResNet-based discriminator and a UNet-based generator.

After a thorough training phase with the PyTorch framework and the use of Perlin noise to create realistic cloud patterns, the GAN-based model outperformed conventional cloud removal strategies. The findings of the Peak Signal-to-Noise Ratio (PSNR) demonstrate how well the model produces clear, high-quality images that closely resemble the ground reality.

The thorough analysis, which includes qualitative and quantitative evaluations, validates the model's capacity to improve the usability and clarity of satellite photos. The model's stability and reliability in a variety of settings are highlighted by the visual comparisons and PSNR measures, highlighting its potential for practical applications.

In summary, satellite image processing has evolved tremendously as a result of the combination of cutting-edge neural network designs with creative noise generating approaches. In addition to offering a workable method for removing clouds, the report establishes the foundation for further research into deep learning-based satellite imagery enhancement. Subsequent investigations may examine the applicability of this technology to alternative types of image obstructions and the possibility of removing clouds in real time from satellite imaging systems.

REFERENCES

[1] Goodfellow, Ian Pouget-Abadie, Jean Mirza, Mehdi Xu, Bing Warde-Farley, David Ozair, Sherjil Courville, Aaron Bengio, Y.. (2014). Generative Adversarial Networks. Advances in Neural Information Processing Systems. 3. 10.1145/3422622.

[2] P. Helber, B. Bischke, A. Dengel and D. Borth, "EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 7, pp. 2217-2226, July 2019, doi: 10.1109/JSTARS.2019.2918242.

[3] Wen, Xue Pan, Zongxu Hu, Yuxin Liu, Jiayin. (2021). Generative Adversarial Learning in YUV Color Space for Thin Cloud Removal on Satellite Imagery. Remote Sensing. 13. 1079. 10.3390/rs13061079.

[4] Ken Perlin. 1985. An image synthesizer. In Proceedings of the 12th annual conference on Computer graphics and interactive techniques (SIGGRAPH '85). Association for Computing Machinery, New York, NY, USA, 287–296. https://doi.org/10.1145/325334.325247

[5] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham.

[6] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi:10.1109/CVPR.2016.90.

[7] Zhao, Mingmin Olsen, Peder Chandra, Ranveer. (2021). Seeing Through Clouds in Satellite Images.