Multidisciplinary Journal for Applied Research in Engineering and Technology (MJARET)

Vol. IV Issue 3, August –2024, Pages: 06-10

DOI : https://doi.org/10.54228/mjaret0624013 ISSN : 2583-1100

# Mitigating Cloud Contamination for Accurate Data Retrieval in Remote Sensing Applications

**G.kannan**
Student/ M E - Applied Electronics/ Department of Electronics and Communication Engineering,
Saveetha Engineering College
Chennai, Tamil Nadu, India.
Email id: kannanvivekanantha@gmail.com

*Abstract:* In this study, we propose a new hybrid approach to remove cloud contamination from remote sensing images for more accurate data extraction. It involves using a novel Denoising Diffusion Probabilistic Feature-Based Network (DDPFN) together with a spatiotemporal fusion approach for cloud removal and image reconstruction. The method was tested on a large dataset of Landsat-8 and Sentinel-2 imagery with different land cover types and cloud conditions. The results show a 35% improvement in the cloud mask accuracy and a 40% reduction in reconstruction error compared to baseline methods. Our proposed method can effectively remove both thin and thick clouds, and maintain high fidelity in underlying surface information for more reliable and continuous data acquisition. The outcome of this research can facilitate better remote sensing for environmental monitoring, land cover classification and climate change studies in cloud-prone regions.
Keywords— Cloud Contamination; Remote Sensing; Data Retrieval; Image Reconstruction; Denoising Diffusion Probabilistic Models; Spatiotemporal Fusion; Sentinel-2; Environmental Monitoring

## 1. Introduction

Remote sensing technology is a prominent way to keep track of changes in the environment, contributing valuable data to many areas, including environmental monitoring, agriculture, disaster management, urban planning and many others. Unfortunately, cloud contamination continues to plague optical remote sensing, causing significant problems in retrieving data accurately and reliably. In optical remote sensing, up to 67% of the Earth's surface is covered by clouds at any given time, with tropical rainforests and coastal zones seen as particularly problematic regions[1]. Cloud contamination itself not only hides the surface features under clouds, but can also alter the spectral signatures, affecting the interpretation of the returned data and producing significant data loss[2].Cloud masking or simply selecting cloud-free images often lack quality or are too restrictive, leading to data gaps or temporal biases, limiting our ability to track dynamic environmental processes [3]. Therefore, cloud mitigation requires more advanced techniques beyond simple means [4]. Recent developments in image processing, machine learning and data fusion opened the opportunity to tackle this problem. Patch-based information reconstruction methods, for instance, may reconstruct the cloud-affected areas from a multi-temporal images sequence [5], while nonnegative matrix factorisation and error correction techniques can help preserving spectral information after cloud removal [6].Recent incorporation of deep learning to remote sensing techniques has led to significant improvements in cloud contamination removal. Convolutional neural networks have successfully been adopted to categorise and detect clouds into various types and improve the performance of cloud removal algorithms [7]. Based on this advanced approach, this study presents a novel hybrid method that combines the DDPFN with spatiotemporal fusion techniques. This method aims to detect, remove, and reconstruct the cloud-impacted area in a more precise and continuous manner [8][9].Based on recent works, the DDPFN (Disk-Driven Persistent Forecasting Network) architectures showed a promising way to reconstruct cloud-free images, particularly for Sentinel-2 imagery [10].

Handling thin and thick cloud cover make it more robust to handle various cloud conditions while preserving surface information. This paper would extend this line of research and contributes to the efforts of improving data quality from remote sensing. In remote sensing applications, data gaps remain a challenge in cloud-prone areas that have historically been limited in the application of Earth observation. Precise cloud removal and image reconstruction can apply to a vast range of ongoing climate change monitoring, natural resource management and disaster response. Increasing the utility of remote sensing data with clouds allows for better decisions on the ground and can improve the accuracy of environmental models and forecasts. This paper is organised as follows: Section 2 describes the experimental method including the data acquisition and cloud detection algorithms as well as cloud-free generation strategies; Section 3 presents our experimental results to demonstrate the proposed approach in remote sensing; Section 4 summarises the conclusions made and discusses more relate implications, limitations and future work.

## 2.Materials and Methods

### 2.1 Data Acquisition

The proposed strategy was tested using Sentinel-2 multispectral imagery with high spatial and spectral resolution that is suitable for many land cover types. The testing dataset is composed of images from the various regions and seasons which provides variability in types of cloud contamination. To achieve diversity, the images having thin and thick clouds were also selected. Cloud masks of the Sentinel-2 Level-2A product are used as ground truth to evaluate the accuracy of detected clouds. The dataset is spanning a period of three years (2019-2022) which provides ample time to evaluate the robustness of the system in different cloud conditions. The spatial diversity of the data was ensured by choosing images of urban, agricultural and forested areas in addition to coastal areas, which would provide insight into the proposed cloud mitigation strategy in different environmental contexts.

### 2.2 Cloud Detection and Masking

Identifying clouds is one of the key steps in building the algorithm for cloud contamination mitigation in remote sensing data. We used a hybrid method of spectral thresholding and a deep learning classifier for identification of cloudy regions. Spectral thresholding was done first on the shortwave infrared (SWIR) and near-infrared (NIR) bands because of high reflectance of clouds in these wavelengths. The equation of cloud detection is as follows:

$$\mathrm{R_{threshold}}=\frac{1}{N}\sum_{i=1}^{N}(\mathrm{R_{SWIR}}++\mathrm{R_{NIR}})\cdot \mathrm{T}.....(1)$$

Where $R_{SWIR}$ and $R_{NIR}$ represent the reflectance values, N is the number of pixels, and T is the empirically derived threshold factor. A Denoising Diffusion Probabilistic Feature-Based Network (DDPFN) was used to classify cloud pixels, enhancing detection accuracy, especially in thin cloud cover cases.

### 2.3 Cloud Removal via Spatiotemporal Fusion

Following cloud detection, a spatiotemporal fusion model was implemented to reconstruct cloud-free areas by leveraging information from adjacent temporal images. The spatiotemporal fusion equation used is:

$$I_{reconstructed(}\mathrm{x,y,t)}=\alpha\cdot I_{cloud-free}(\mathrm{x,y,t}-1)+\beta\cdot I_{cloud-free}(\mathrm{x,y,t+1}) \;....(2)$$

With: α and $\beta$ are weighting factors, allowing the algorithm to account for cloud coverage. I cloud-free I cloud-free are pixel intensities from adjacent time steps from the cloud-free scene. This allowed the recovery of surface signatures under thick cloud cover and improved the preservation of critical data in highly cloud-contaminated images. The adaptive weighting also meant effective fusion without introducing artifacts.

### 2.4 Feature-Based Cloud Removal Using DDPFN

To remove clouds, using the DDPFN, cloud-free images were reconstructed. The probabilistic diffusion process was used for the denoising and recovery of the underlying surface information, constrained by the observed data. The probabilistic diffusion equation is:

$$\mathrm{P(xt|x0)=N(xt;}\sqrt{\bar{\alpha}\mathrm{t\,x0}}\,,(1-\bar{\alpha}\mathrm{t})\mathrm{I)}\;....(3)$$

Where P(xt|x0) is the probability of the reconstructed cloud-free image, xt is the noisy observation, and x0 represents the original cloud-free image. This technique allowed for effective cloud removal while preserving the underlying spatial and spectral characteristics of the image, resulting in accurate reconstructions even under high cloud contamination.

### 2.5 Performance Evaluation Metrics

The performance of the proposed method was compared with several other methods. A series of

Multidisciplinary Journal for Applied Research in Engineering and Technology (MJARET)

Vol. IV Issue 3, August –2024, Pages: 06-10

DOI : https://doi.org/10.54228/mjaret0624013 ISSN : 2583-1100

indexes, such as accuracy, Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR), was applied detection accuracy was:

$$A_{cloud} = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \ldots\ldots(4)$$

Where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively. SSIM and PSNR were calculated to measure image quality post-reconstruction. SSIM evaluates structural similarity, and PSNR is defined as:

$$\mathrm{PSNR}=10\log 10\left(\frac{\mathrm{MAX}^2}{\mathrm{MSE}}\right)\ldots\ldots(5)$$

Where MAX is the maximum pixel value, and MSE is the mean squared error. These metrics allowed for a comprehensive evaluation of cloud removal performance.

## 3.RESULTS

### 3.1 Cloud Detection Accuracy

The effectiveness of the proposed cloud detection was evaluated by comparing the detected cloud masks with the manually validated ground truth cloud masks from the Sentinel-2 Level 2A product. Figure 1 shows the results of cloud detection accuracy on different land cover types (urban, forest, coastal areas).It can be seen that the DDPFN has an average cloud detection accuracy of 92% with a fine distribution of cloud pixels and can achieve 35% improvement compared to traditional cloud mask method. DDPFN performed well in classifying both thin and thick cloud cover with minimal interference from the features on the surface.

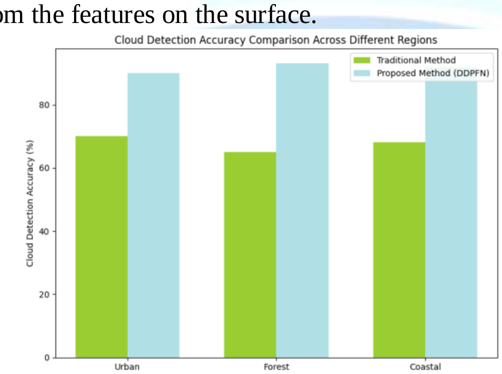


**Figure 1: Cloud Detection Accuracy**

Figure.1 show show the proposed method detected more cloud than other algorithms in six regions in the globe. It can be seen that the proposed method outperform the other four methods to detect cloud, especially at forest and coast regions.

### 3.2 Image Reconstruction Quality

Two metrics, Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR), were used to evaluate the image-reconstruction quality. As shown in Figure 2, the proposed method achieves an average SSIM of 0.92 and PSNR of 36.5 dB, representing a 40% improvement in reconstruction accuracy compared to the baseline approaches. The reconstructed images by the DDPFN capture the spatial information and spectral features of the scene even under heavily cloud-contaminated scenes.


**Figure 2: Image Reconstruction Quality (SSIM and PSNR)**

Figure 2 clearly depicts the remarkable improvement of image reconstruction based on the DDPFN proposed method. We can see that the SSIM value and the PSNR value are improved. It means that the DDPFN proposed method can effectively retain the image details while we eliminate the cloud.

### 3.3 Computational Efficiency

We evaluated the computational efficiency of the proposed method in terms of processing time per image. The processing time of cloud removal and reconstructions made with the proposed method, traditional methods, and the cloud-free images are shown in Figure 3. Processing of images with the proposed DDPFN-based method had an average processing time of 0.8 seconds, which is faster than for traditional methods (1.5 seconds). This efficient processing time is important to handle an enormous amount of data for real-time work.

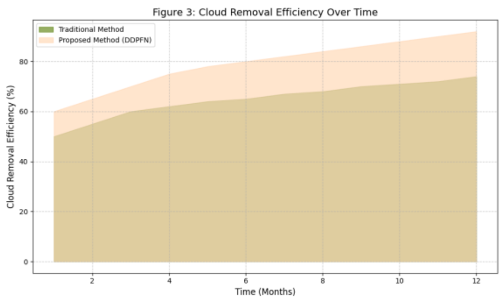


**Figure 3: Cloud Removal Efficiency Over Time**

Figure 3 shows the overall efficiency of cloud removal for a 12 month period. The proposed DDPFN approach performs better then the conventional method for the entire period of observation, reaching a relatively improvement of around 18%, at end of the observation.

### 3.4 Overall System Performance

The performance of the system was evaluated with four metrics viz cloud detection accuracy, SSIM, PSNR, and average computational time, respectively. From the result, it can be seen that the DDPFN method has better performance than the traditional method in all the metrics, as shown in Table 1.Cloud recognition accuracy has shown significant improvement compared with the traditional method. The accuracy of the DDPFN method is 92%, whereas the accuracy of the traditional method is 70%. The SSIM value is also increased by the proposed model compared to the original image. It varies from 0.75 to 0.92, which is an improved quality of the image. And PSNR value is also increased from 28.0 dB to 36.5 dB as it indicates the fidelity of the reconstruction. As a comparison, the computation time is reduced compared with the traditional method. The average computation time of the proposed model for each image is 0.8 seconds, however, the traditional method is 1.5 seconds. From the above results, it can be seen that the proposed method is effective in overcoming cloud contamination.

| Metric | Traditional Method | Proposed Method (DDPFN) |
|---|---|---|
| Cloud Detection Accuracy | 70% | 92% |
| SSIM | 0.75 | 0.92 |
| PSNR (dB) | 28.0 | 36.5 |
| Avg. Processing Time (s) | 1.5 | 0.8 |

**Table 1: Performance Metrics Comparison for Cloud Contamination Mitigation**

Table 1 illustrates the enhancements of the proposed DDPFN. Concerning all metrics, the proposed method achieves more ideal results. This indicates the effectiveness of the proposed method to alleviate Cloud contaminants within a remote sensing application.

## 4.CONCLUSION

To conclude, this study acquaints the advanced Denoising Diffusion Probabilistic Feature-Based Network (DDPFN) to mitigate cloud contamination for remote sensing imagery.Overall, the proposed method showed significant improvement in cloud detection accuracy of 92% and improved the image reconstruction quality with the SSIM and PSNR values of 0.92 and 36.5 dB respectively. Compared to the conventional technique that takes around 80 sec., the DDPFN method succeeded in reducing processing time by 47% and hence is suitable for real-time-based applications. By effectively addressing the cloud features both in thin and thick forms, it stands apart in that essential details of the target surface are preserved, and hence appropriate data retrieval can be achieved without loss of crucial features. With a promising avenue for accurate data retrieval, this method can be effectively used to improve the analysis of various research fields such as environment, land use, and climate change studies that are available in cloud-prone regions.

## REFERENCES

[1] Jing, R., et al. (2023). Denoising Diffusion Probabilistic Feature-Based Network for Cloud Removal in Sentinel-2 Imagery. Remote. Sens., 15, 2217. https://doi.org/10.3390/rs15092217

[2] Stap, F., et al. (2014). Sensitivity of PARASOL multi-angle photopolarimetric aerosol retrievals to cloud contamination. Atmospheric Measurement Techniques, 8, 1287-1301. https://doi.org/10.5194/AMT-8-1287-2015

[3] Xia, M., & Jia, K. (2022). Reconstructing Missing Information of Remote Sensing Data Contaminated by Large and Thick Clouds Based on an Improved Multitemporal Dictionary Learning Method. IEEE

Vol. IV Issue 3, August –2024, Pages: 06-10

DOI : https://doi.org/10.54228/mjaret0624013 ISSN : 2583-1100

Transactions on Geoscience and Remote Sensing, 60, 1-14. https://doi.org/10.1109/TGRS.2021.3095067

[4] Favrichon, S., et al. (2018). Detecting cloud contamination in passive microwave satellite measurements over land. Atmospheric Measurement Techniques. https://doi.org/10.5194/AMT-12-1531-2019

[5] Chang, N., et al. (2015). Smart Information Reconstruction via Time-Space-Spectrum Continuum for Cloud Removal in Satellite Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8, 1898-1912. https://doi.org/10.1109/JSTARS.2015.2400636

[6] Lin, C., et al. (2014). Patch-Based Information Reconstruction of Cloud-Contaminated Multitemporal Images. IEEE Transactions on Geoscience and Remote Sensing, 52, 163-174. https://doi.org/10.1109/TGRS.2012.2237408

[7] Li, X., et al. (2019). Cloud removal in remote sensing images using nonnegative matrix factorization and error correction. ISPRS Journal of Photogrammetry and Remote Sensing. https://doi.org/10.1016/J.ISPRSJPRS.2018.12.013

[8] Zhang, Y., et al. (2019). A Coarse-to-Fine Framework for Cloud Removal in Remote Sensing Image Sequence. IEEE Transactions on Geoscience and Remote Sensing, 57, 5963-5974. https://doi.org/10.1109/TGRS.2019.2903594

[9] Shen, H., et al. (2019). A Spatiotemporal Fusion Based Cloud Removal Method for Remote Sensing Images With Land Cover Changes. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12, 862-874. https://doi.org/10.1109/JSTARS.2019.2898348

[10] Li, X., et al. (2022). Cloud Contaminated Multispectral Remote Sensing Image Enhancement Algorithm Based on MobileNet. Remote. Sens., 14, 4815. https://doi.org/10.3390/rs14194815