IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10640766

A DEEP LEARNING APPROACH TO CLOUD AND SHADOW DETECTION IN MULTIRESOLUTION, MULTITEMPORAL AND MULTISENSOR IMAGES

Darwin Alexis Arrechea-Castillo [1,2], Yady Tatiana Solano-Correa [1,3],
Julián Fernando Muñoz-Ordoñez [4], Edgar Leonairo Pencue-Fierro [1]

[1] Universidad del Cauca, Calle 5 #4-70, 190001 Popayán, Colombia
[2] International Center for Tropical Agriculture, Km 17 Recta Cali–Palmira, 763537, Palmira, Colombia
[3] Universidad Tecnológica de Bolívar, Km 1 Vía Turbaco, 130010 Cartagena, Colombia.
[4] Corporación Universitaria Comfacauca, Calle 4 N° 8-30, 190003 Popayán, Colombia

ABSTRACT

Accurate detection of clouds and shadows present in optical imagery is important in remote sensing for ensuring data quality and reliability. This study introduces a deep learning model capable of generating precise cloud and shadows masks for subsequent filtering. Unlike other works in literature, this model operates efficiently across diverse temporalities, sensors, and spatial resolutions, without the need for any relative or absolute transformation of the original data. This versatility, to date unreported in the literature, marks a significant advancement in the field. The model utilizes data from PlanetScope, Landsat and Sentinel-2 sensors and is based on a simplified convolutional neural network (CNN) architecture, LeNet, which facilitates easy training on standard computers with minimal time requirements. Despite its simplicity, the model demonstrates robustness, achieving accuracy metrics over 96% in validation data. These results show the model potential in transforming cloud and shadow detection in remote sensing, combining ease of use with high accuracy.

***Index Terms*—** Cloud Detection, Cloud Shadow Detection, Deep Learning, Remote Sensing, MultiSensor.

1. INTRODUCTION

The precise mapping of clouds and their shadows in remote sensing (RS) is a critical aspect for ensuring the quality and reliability of satellite imagery [1]. Accurate detection of these features is essential for a variety of applications, including climate monitoring, weather forecasting, and land cover mapping [2]. Clouds and their shadows can significantly mask ground information, leading to inaccuracies in data interpretation. Recent advances in deep learning (DL) have shown promising results in addressing this challenge [2–5]. DL models excel in extracting both spatial and spectral features, which are crucial for effective cloud and shadow detection in satellite imagery. These models, have demonstrated their ability to handle complex cloud and cloud shadows segmentation tasks efficiently. However, existing models often face limitations when it comes to handling data from multiple sensors or under varying conditions [6]. There is a growing need for an algorithm capable of detecting clouds and cloud shadows across different spatial resolutions, regions, temporalities, and sensors, without requiring significant relative or absolute transformations of the original data. Transformations, typically used to standardize or normalize data, can sometimes alter the intrinsic properties of the dataset, leading to potential inaccuracies in the analysis. In this context, the use of PlanetScope imagery [7] for training DL models is strategic. PlanetScope offers a unique combination of temporal resolution and wide coverage, making it an ideal dataset for developing a versatile cloud and cloud shadow detection model. The challenge lies in creating a model that is not only highly accurate but also adaptable to the diverse conditions. These include multiple: spatial resolution, geographic areas, RS acquisition dates, and RS data sources, such as various Landsat and Sentinel-2 (S2) sensors. Additionally, the model demonstrates robust performance in detecting the three main types of clouds: cirrus, cumulus, and stratus. While there have been attempts to create models for similar purposes, many of these have fallen short in achieving the level of adaptability and accuracy required for comprehensive cloud and cloud shadow detection across diverse datasets [8–12]. The development of a model that overcomes these limitations while maintaining high accuracy and efficiency in processing, would mark a significant advancement in the field of RS.

2. RS DATA AND STUDY AREA DESCRIPTION

This study employs high spatial resolution satellite imagery from three different optical sensors. PlanetScope, commonly referred as $Planet$. PlanetScope imagery is renowned for its fine spatial (3m) and temporal (daily) resolutions, making it an ideal choice for detailed remote sensing analysis [13]. The dataset encompasses two distinct types of PlanetScope tiles. The first subset includes imagery from the year 2018 with a 4-band configuration, providing essential spectral informa-

Authorized licensed use limited to: Pontificia Universidad Javeriana. Downloaded on September 09,2024 at 13:48:11 UTC from IEEE Xplore. Restrictions apply.

tion for cloud and cloud shadow detection. The second subset consists of more recent, 2023, tiles featuring 8-band configuration, offering enhanced spectral detail. For testing purposes, the study further incorporates imagery from alternative remote sensing sources, specifically S2 and various Landsat sensors. These additional datasets encompass a diverse range of geographic areas and acquisition dates, providing a rigorous testing environment for the model's adaptability and accuracy. To build the training and validation data, two representative sub-basins in Colombia were used: the Las Piedras and Palacé river sub-basins [14]. In total, the entire study area adds up to around 711 km^2. These areas were strategically chosen due to their location within the Intertropical Convergence Zone, a region characterized by high cloud cover [14]. This choice ensures that the dataset includes a wide range of cloud and cloud shadow scenarios, crucial for developing a robust detection model. The total number of images and dataset used for this study are presented in Table 1.

Table 1: Raw dataset collected

	Date	#Tiles	#Bands
Las Piedras	Feb-19-2018	2	4
	Jul-02-2018	4	4
	Jan-03-2018	2	4
	Aug-06-2018	2	4
	Jul-07-2018	4	4
	Aug-31-2018	2	4
	Jan-26-2023	1	8
	Jan-04-2023	2	8
	Feb-05-2023	1	8
	Feb-07-2023	1	8
	Jan-25-2023	1	8
Palacé	Nov-06-2018	2	4

3. PROPOSED APPROACH FOR CLOUD AND CLOUD SHADOW DETECTION

This study presents a structured methodology optimized for multispectral PlanetScope imagery that is also suitable for detections in Landsat and S2 sensors. The approach is designed to tackle the complexity of cloud and cloud shadow detection across varying spatial and temporal resolutions, geographical regions, and sensor types. The workflow of the proposed methodology is detailed in Figure 1.

3.1. PlanetScope Raw Data and Preprocessing

Multispectral imagery from the Planet Explorer website [13] are chosen and downloaded. Following the acquisition of the imagery specified in Table 1, it was necessary to execute a cropping phase of the tiles to the designated study area. This step allows optimizing computational resources. By focusing solely on the area of interest, the demands on storage and

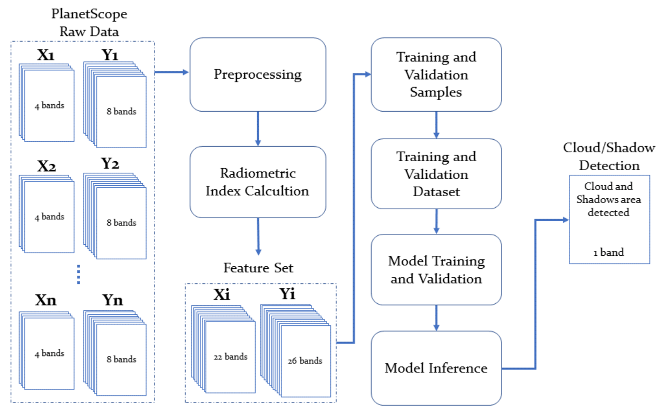


Fig. 1: General methodology.

processing cost required for subsequent data handling was reduced.

3.2. Radiometric Index Calculation and Feature Set Configuration

To construct a robust feature set providing comprehensive information to the Convolutional Neural Network (CNN), a total of 18 radiometric indices were computed. These indices were chosen for their ability to be derived from the minimum of four spectral bands - Red, Green, Blue, and Near-Infrared (RGB-NIR) - present in PlanetScope imagery, which typically consists of 4 and 8 bands. The selection of these specific indices ensures that the proposed methodology is not limited to PlanetScope data, but is broadly applicable to most remote sensing datasets, as they commonly include these four fundamental bands. The calculated indices include: ARVI, ATSAVI, BNDVI, CI, EVI, GCL, GDVI, GLI, IO, IPVI, NDTI, NDVI, NDWI, CRI550, SAVI, SIPI, SR, and SRWI. The feature sets are subsequently formed by combining the original spectral bands with these computed indices, resulting in enhanced images comprising 22 and 26 bands for the 4-band and 8-band images, respectively. This approach enriches the dataset, enabling the CNN to extract more nuanced information for cloud and cloud shadow detection tasks.

3.3. Model Training, Validation and Inference

For this study, a LeNet-based architecture model was developed. For training and validation, a dataset was built after marking some samples using a custom interface developed using Python with GDAL, QT, and OpenCV libraries. The training and validation process was initiated after artificially balancing the raw dataset shown in Figure 2, meaning that the number of instances in each class was equalized based on the class with the highest occurrence in both datasets built from the 4-band and 8-band images. The distribution of data was 80% for training and 20% for validation. It is important to re-

Authorized licensed use limited to: Pontificia Universidad Javeriana. Downloaded on September 09,2024 at 13:48:11 UTC from IEEE Xplore. Restrictions apply.

call that training data is selected only from PlanetScope data. But validation is performed by considering different sensors, i.e., PlanetScope (3m), Landsat (30m) and S2 (10m).

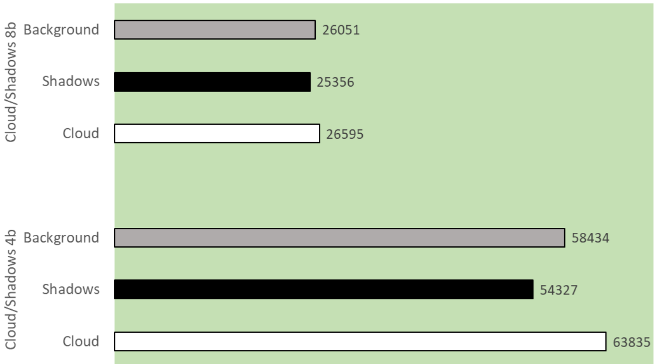


Fig. 2: Raw Dataset Distribution for PlanetScope with 4 and 8 bands.

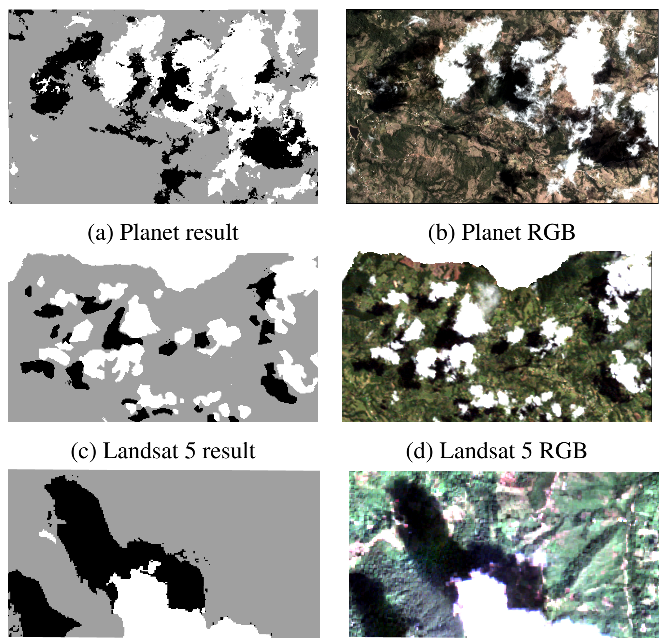
4. PRELIMINARY RESULTS AND DISCUSSIONS

Table 2 presents the quantitative validation results for the models derived from the constructed datasets. The 22-band feature set model, based on the 4-band images, and the 26-band feature set model, based o the 8-band images, both exhibit robust performance in terms of the accuracy metric. It is observed that the model trained with the 8-band images achieves higher accuracy, attributable to its access to a more broader spectral information. This expanded feature set (26 instead of 22) enhances the model's capacity to discern between the background, clouds, and cloud shadows more effectively. It is noteworthy that both models were trained in less than two hours. Specifically, the training for the 22-band based model was completed in 1.85 hours, and the training for the 26-band based model was completed in 1.96 hours on a system equipped with an Intel® Core™ i9-9900K CPU @ 3.60GHz (16 CPUs), ~3.6GHz and a NVIDIA GeForce RTX 3090 GPU. The final CNN configuration consisted of four convolutional layers with varying filter counts, kernel sizes set uniformly at 3x3, ReLu activation functions, and "same" padding to maintain dimensionality. Strides were set at one across all convolutional layers. Pooling layers alternated between max and average, with the introduction of dropout at strategic points to mitigate overfitting, specifically set at 0.25 for all but the third convolutional layer where it was omitted.

Upon quantitative validation, the implemented models were then subjected to inference on previously unseen images. These images varied in spatial resolution (3, 10, and 30 meters per pixel), covered different regions across Colombia, were taken on different dates, and originated from multiple sensors, including Landsat 5, 7, 8, and 9. Consistently favorable outcomes were observed across all scenarios. Figure 3 showcases a selection of results for PlanetScope, Landsat, and S2 data, where the left panel shows in white the clouds and on black their corresponding shadows. These multiple-case validations confirms the models' robustness and adaptability to diverse datasets and conditions. Furthermore, it is significant to highlight that these favorable results were achieved without the need for any absolute or relative transformations on the original data. This implies that the models' efficacy is not contingent on altering the raw satellite images through methods such as normalization, standardization, or contrast adjustment, which are commonly employed to homogenize data before analysis. However, a simple scaling of data between 0 and 1 (reflectance range) is the sole modification prior to the deep learning training, facilitating efficient convergence.

Table 2: Validation results of the cloud and cloud shadows detection models

Model	Validation Data Accuracy	Trained Epochs	Best Epoch
4 bands	96.88%	110/350	90
8 bands	99.56%	101/350	81



(a) Planet result (b) Planet RGB

(c) Landsat 5 result (d) Landsat 5 RGB

(e) Sentinel 2 result (f) Sentinel 2 RGB

Fig. 3: Results of the model inference and corresponding RGB images for selected test areas.

Authorized licensed use limited to: Pontificia Universidad Javeriana. Downloaded on September 09,2024 at 13:48:11 UTC from IEEE Xplore. Restrictions apply.

5. CONCLUSIONS

In the conducted research, a DL model for cloud and cloud shadow detection in multiresolution, multitemporal and multisensor images was developed. The presented model excels in adaptability, efficiency, and accuracy for cloud and cloud shadow detection across a range of remote sensing scenarios. The foundational dataset for this model was sourced from PlanetScope imagery. Its capability to function effectively without the necessity of data transformations marks a significant advancement in the field of remote sensing. The model's streamlined and efficient training process, completed in under two hours, underscores its practicality, particularly in real-world applications where computational resources are generally limited. This efficiency is essential for applications that require timely analysis and limited computational expenditure. Overall, this research contributes a groundbreaking approach to cloud and cloud shadow detection. It represents a substantial step forward in remote sensing technologies and their applications, offering improved accuracy and efficiency in areas such as environmental monitoring and agricultural management, where reliable satellite imagery are crucial.

6. ACKNOWLEDGEMENTS

This research received support from the "Bioeconomic Strengthening for Social and Productive Reactivation," and the "Water Security and Sustainable Development Hub." These projects were funded by the science ministry in Colombia and the UK Research and Innovation's Global Challenges Research Fund (GCRF) [grant number: ES/S008179/1]. The authors also express their gratitude to the Universidad del Cauca, Corporación Universitaria Comfacauca and Universidad Tecnológica de Bolívar for their support during the development of this research. The authors would like to thank Planet Labs Inc. for providing the images used in this study under the "Education and Research Program".

7. REFERENCES

[1] Z. Li, H. Shen, Q. Weng, Y. Zhang, P. Dou, and L. Zhang, "Cloud and cloud shadow detection for optical satellite imagery: Features, algorithms, validation, and prospects," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 188, pp. 89–108, 2022.

[2] R. Gupta and S. J. Nanda, "Cloud detection in satellite images with classical and deep neural network approach: A review," *Multimedia Tools and Applications*, vol. 81, no. 22, pp. 31847–31880, 2022.

[3] A. Francis, P. Sidiropoulos, and J.-P. Muller, "Cloud-FCN: Accurate and Robust Cloud Detection for Satellite Imagery with Deep Learning," *Remote Sensing*, vol. 11, no. 19, p. 2312, 2019.

[4] M. Khoshboresh-Masouleh and R. Shah-Hosseini, "A Deep Learning Method for Near-Real-Time Cloud and Cloud Shadow Segmentation from Gaofen-1 Images," *Computational Intelligence and Neuroscience*, vol. 2020, p. e8811630, 2020.

[5] K. Xu, K. Guan, J. Peng, Y. Luo, and S. Wang, "Deep-Mask: An algorithm for cloud and cloud shadow detection in optical satellite remote sensing images using deep residual network," 2019.

[6] H. Zhai, H. Zhang, L. Zhang, and P. Li, "Cloud/shadow detection based on spectral indices for multi/hyperspectral optical remote sensing imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 144, pp. 235–253, 2018.

[7] D. P. Roy, H. Huang, R. Houborg, and V. S. Martins, "A global analysis of the temporal availability of PlanetScope high spatial resolution multi-spectral imagery," *Remote Sensing of Environment*, vol. 264, no. 112586, p. 21, 2021.

[8] Z. Li, H. Shen, Q. Cheng, Y. Liu, S. You, and Z. He, "Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors," *Isprs Journal of Photogrammetry and Remote Sensing*, vol. 150, pp. 197–212, 2019.

[9] S. Mahajan and B. Fataniya, "Cloud detection methodologies: Variants and development—a review," *Complex & Intelligent Systems*, vol. 6, no. 2, pp. 251–261, 2020.

[10] N. Ma, L. Sun, C. Zhou, and Y. He, "Cloud Detection Algorithm for Multi-Satellite Remote Sensing Imagery Based on a Spectral Library and 1D Convolutional Neural Network," *REMOTE SENSING*, vol. 13, no. 16, p. 3319, 2021.

[11] D. Montero, C. Aybar, M. D. Mahecha, F. Martinuzzi, M. Söchting, and S. Wieneke, "A standardized catalogue of spectral indices to advance the use of remote sensing in Earth system research," *Scientific Data*, vol. 10, no. 1, p. 197, 2023.

[12] X. Xiang, K. Li, B. Huang, and Y. Cao, "A Multi-Sensor Data-Fusion Method Based on Cloud Model and Improved Evidence Theory," *Sensors*, vol. 22, no. 15, p. 5902, 2022.

[13] PLANET.COM, "Planet Imagery Product Specifications," 2022.

[14] D. A. Arrechea-Castillo, Y. T. Solano-Correa, J. F. Muñoz-Ordóñez, E. L. Pencue-Fierro, and A. Figueroa-Casas, "Multiclass Land Use and Land Cover Classification of Andean Sub-Basins in Colombia with Sentinel-2 and Deep Learning," *Remote Sensing*, vol. 15, no. 10, p. 2521, 2023.

Authorized licensed use limited to: Pontificia Universidad Javeriana. Downloaded on September 09,2024 at 13:48:11 UTC from IEEE Xplore. Restrictions apply.