Detecting Forged Sentinel-2 Images Through Parallax-Based Cloud Analysis

Matthieu Serfaty, Quentin Bammey, Tina Nikoukhah, Rafael Grompone von Gioi, Carlo de Franchis

Detecting Forged Sentinel-2 Images Through Parallax-Based Cloud Analysis

Matthieu Serfaty[1], Quentin Bammey[1], Tina Nikoukhah[1], Rafael Grompone von Gioi[1], and Carlo de Franchis[1,2]

[1] Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centre Borelli, 91190 Gîf sur Yvette, France
[2] Kayrros SAS, France

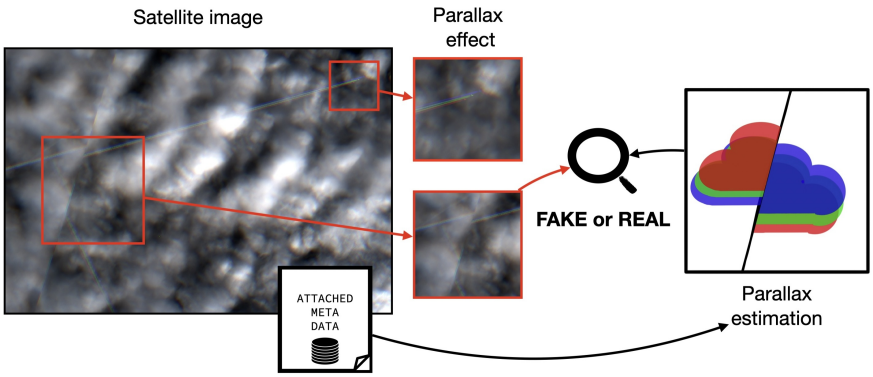

Fig. 1: Illustration of the parallax effect visible in the cloud and on a flying plane. The parallax order is reversed at each joint between sensors. This effect is particularly visible on plane vapour trails and clouds. We use this alternating parallax to detect fake clouds, by comparing the estimated parallax properties with the theoretical ones derived from the tile metadata.

Abstract. The availability and significance of satellite imagery in our world is continuously growing. Satellite images now play a crucial role in various applications such as weather forecasting, greenhouse gas monitoring, agricultural crop health assessment, and external security. However, this also exposes them to malicious attacks aimed at hiding or manipulating information. Forensic analysis is, therefore, a necessary shield against disinformation and disruption attempts in these areas. While forensic analysis of photographs has received considerable academic attention in recent years, the same cannot be said for satellite imagery. In this study, we present two methodologies to create realistic datasets of images forged with added clouds, that may inconspicuously hide information. We show results produced by state-of-the-art forensic methods are unable to detect the forged satellite images. To overcome this problem, we propose a parallax-based method to detect inconsistent satellite images.

1 Introduction

Satellite images have gained widespread use and accessibility for various applications due to the increased availability and reduced costs of Earth observation satellites. However, as any images, satellite images are also vulnerable to forgery. Tampering risks range from inserting external objects into the image [27] to using generative methods that synthetically alter regions [29].

Extensive research exists for detecting image forgeries in traditional photographs. These techniques include analysing traces left by the image processing pipeline [6, 7, 11, 13, 31], examining global inconsistencies in image noise [17, 18, 22, 24], or training neural networks on forged images [34]. However, it is unclear whether these methods are equally effective for satellite images. Satellite image generation feature unique characteristics, such as largely different pipelines across satellite constellations, and specialized formats such as GeoTIFF or JPEG2000 rather than the ubiquitous JPEG compression of natural images. These specificities likely pose challenges to existing forensic tools [27], whose research is focused on traditional images. Despite their intuitive appeal, however, these claims are yet to be substantiated. Unfortunately, satellite image forgery datasets are scarce [15, 23, 25–27] and unrealistic.

Our contribution is three-fold:

- We explore a new protocol to produce realistic forged satellite images by adding fake clouds to them, a realistic scenario which may be used to hide data.
- We show that existing forensic tools, both generic and specific to satellite images, are unable to detect such forgeries,
- We propose a new method to detect satellite image forgeries by detecting inconsistencies between the estimated and metadata-derived parallax effect changes introduced by push broom scanners.

Our forgery method relies on cloud addition in an image, either using fully-generated clouds with Perlin noise or splicing them from another image. Indeed, clouds can be used to hide structures beneath them without arousing suspicion, and fake clouds are thus prime candidates to evaluate satellite image forgery detection tools. Also the ground itself can be used as a splicing area to hide something. In contrast, a realistic splicing of objects onto satellite images would not be as useful, as most objects that could be spliced would only be a few pixels large at the resolution of public satellites. Indeed, we focus on tampering Sentinel-2 products. Sentinel-2 sensors resolutions are too small to realistically observe real objects such as a plane, which would only be a few pixels large.

Another possible approach to create realistic fake satellite images would be to retrain and use recent generative models to generate or modify content. However, conditioning and retraining a model on satellite images would require extensive resources. Most importantly, generated images can be detected regardless of whether images used for training the detection model come from a generative model trained on different images using spectral analysis [8, 30]. Such an analysis might even be easier on satellite images, free from the modifications introduced

by common photo processing on social media [28]. Thus, there is no actual need to focus specifically on satellite images when it comes to generated image detection. Therefore, we choose not to focus on this orthogonal protocol.

1.1 Sentinel-2 mission

We use Sentinel-2 images as the foundation for our forgery dataset. The Sentinel-2 mission includes two identical satellites, Sentinel-2A and Sentinel-2B, orbiting 783 kilometers above Earth in a sun-synchronous orbit. Each satellite is equipped with a MultiSpectral Instrument (MSI) that collects sunlight reflected from Earth using a push-broom scanning method, where the satellite's motion captures new data rows.

In this paper, we address all the reflectance bands captured by the MSI. The bands, along with their time delay relative to band 2, central wavelength, bandwidth, and spatial resolution, are detailed in Table 1. The push-broom technology causes a parallax effect, especially noticeable in clouds, between bands. Our method leverages this effect by analyzing cloud parallax to determine its consistency with the MSI. We use Level-1C product images, which have already been geometrically and radiometrically corrected, and compressed using JPEG2000.

The Sentinel-2 satellite's smallest precision is 10 meters per pixel, making it challenging to splice small objects like planes or drones realistically, as they would only span a few pixels, as shown in Figure 1. In contrast, clouds are much larger and more likely to be captured by Sentinel-2. If someone wanted to conceal something in a satellite image, clouds would serve as natural occluders. This work focuses on splicing clouds into satellite images. Since cloud detection is crucial for many applications, where algorithms rely on cloudless images, distinguishing real clouds from fake ones becomes essential.

2 Related Works

2.1 Classical forgery detection

Forensic tools can focus on specific traces within images that were left by the image processing pipeline [14]. Local inconsistencies of these traces constitute evidence of forgery. For example, dephasing inconsistencies can be detected from the image mosaic [9,10,12] as well as from JPEG [31] and JPEG2000 [7] compression traces. Many SOTA methods still follow this approach while incorporating more advanced ideas, such as positional learning [6, 11, 13]. Noise level [22] and fingerprints [17, 18, 24] can also lead to inconsistency detection. These methods contrast with models trained to detect forgeries directly [34].

2.2 Satellite forgery detection

Satellite image forensics has only recently gained attention. Kalyan et al. [27] train a model to distinguish genuine and manipulated parts of a satellite image. SeeTheSeams [23] proposes to detect seam carving in satellite images. Jade

Band #	Time order	Delay with band 2 (s)	Central Wavelength (nm)	Bandwidth (nm)	Spatial Resolution (m)
1	12	2.314	442.2	21	60
2	1	0.000	492.1	66	10
3	3	0.527	559.0	36	10
4	5	1.005	664.9	31	10
5	6	1.269	703.8	16	20
6	8	1.525	739.1	15	20
7	9	1.790	779.7	20	20
8	2	0.264	832.9	106	10
8A	10	2.055	864.0	22	20
9	13	2.586	943.2	21	60
10	4	0.851	1376.9	30	60
11	7	1.468	1610.4	94	20
12	11	2.085	2185.7	185	20

Table 1: Bands number and the order of appearance in the acquisition process accompanied with the time delay compared to the band number 2, the blue band, their respective central wavelength, bandwidth and spatial resolution. The bands highlighted in green are those used in the proposed method.

Owl [7] looks for inconsistencies in the traces left by JPEG2000 compression, an often used compression format in satellite images. Cannas et al. [16] adapts deep splicing detection tools to satellite images by retrieving satellite attribution markers. Their findings indicate the importance of meticulous attention and thorough evaluation of the data life cycle, especially when dealing with data types that diverge from conventional digital images. Other approaches using conditional generative adversarial networks [15], deep belief networks [25], or vision transformers [26] have also been used to encode images before verifying their authenticity.

2.3 Cloud detection

Cloud detection is the sinews of war in this paper. In remote sensing, cloud detection has been studied for decades. Recently the UNetMobV2 method [4] was proposed based on the neural network MobilNetV2 [33] as a backbone. In that work, great attention has been dedicated to data annotation resulting in a highly precise output cloud map. The network takes as input all 13 reflectance bands and output a cloud map containing 4 labels: background, thick clouds, thin clouds and shadows. Some clouds maps output are displayed on Figures 5 and 8.

In this paper, cloud detection will be used in two scenarios. Firstly, it is used to detect real clouds and then splice them onto a cloudless image. Secondly, we attempt to mislead the network into detecting synthetic generated clouds.

3 Proposed Dataset

Some literature [15, 23, 25–27] currently exists on the topic of falsified satellite image datasets, although none of the datasets are openly accessible yet. The manipulations present in these datasets are primarily obtained through splicing of non-satellite images onto satellite images. As the spliced objects originate from non-satellite sources, their characteristics markedly differ from those of satellite images, rendering them conspicuous and easy to detect. Furthermore, these spliced objects are generally larger and visually different than if said object had actually been captured by a satellite, thereby reducing the realistic aspect of the forgery.

Another approach is to insert clouds into satellite images. Such a manipulation is plausible in real-life scenarii as clouds are a common natural occluder on satellite images. To a malicious mind, they would be ideal to inconspicuously conceal relevant elements within an image. To test and challenge the robustness of existing methods to such attacks, we propose two ways of inserting clouds, by either splicing existing clouds from donor satellite images, or by synthesizing Perlin noise to create new clouds.

To create our dataset, we use open-access level-1C (resp. level-1T) satellite images from the Sentinel-2 twin satellites constellation. Note that the proposed methodology can be applied similarly to other constellations using a push-broom technology, such as Landsat 8-9 or WorldView-2.

3.1 Splicing dataset

To introduce clouds into an image, one method is to extract cloud formations from another satellite picture. A cloud detection algorithm [4] is used to identify cloudy pixels and create a mask. Alpha matting refines the mask to align better with cloud boundaries. These clouds are then transferred to another image. In the interest of maintaining the dataset's realism, shadows are introduced by darkening the ground in accordance with the cloud mask, with a slight random offset from the cloud itself.

3.2 Perlin noise clouds dataset

While clouds can be sourced from existing satellite imagery for subsequent inclusion in cloudless images, the availability of such cloudy images can be inconsistent. Hence, an alternative method involves the creation of synthetic clouds, which once generated are subsequently spliced into a separate image.

We propose to use Enomoto's method [20] and Czerkawski's [19] in which synthetic clouds are produced using the Perlin noise algorithm, a technique known for generating natural-looking textures like clouds. Perlin noise [32] works by connecting isolated points generated by a noise function via an interpolation function. The look of the clouds is controlled by parameters such as the number of octaves and the resolution. As in the previous dataset, shadows are added on the ground, and some noise is added on the cloud to simulate a texture.

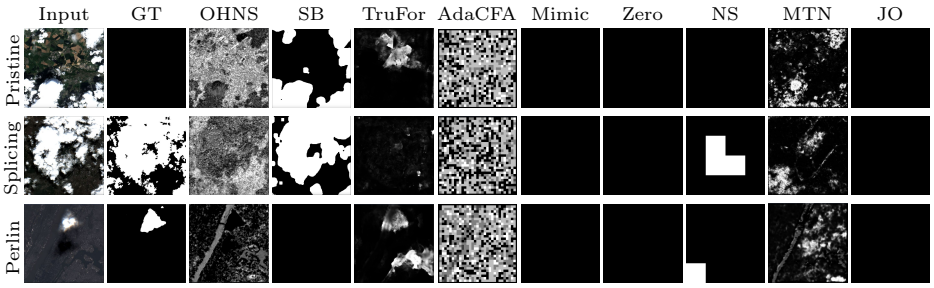

Fig. 2: Results of forensic tools on a pristine image (first row) and on forged images from the two datasets. The second row shows the results on the splicing dataset and the third row the Perlin cloud generation dataset results. Classical forensic tools are unable to detect the forgeries but still make errors on many images.

3.3 Dataset presentation

Our dataset is composed of 6 complete cloudless tiles downloaded using the Copernicus data browser [1]. Those tiles are forged by adding Perlin-noise-generated clouds. For each tile, we generate ten different cloud maps. Generated clouds are added to the cloudless image either with or without parallax, resulting in 100 forged tiles of size 10980×10980. We also use 100 cropped Sentinel-2 images to build spliced and generated clouds forgeries resulting in 100 pristine images, 100 spliced-clouds forgeries and 100 Perlin-noise-generated ones.

4 Proposed detection procedure

The Sentinel-2 constellation is composed of two sun-synchronous satellites. Those satellites acquire image with the MSI. As mention earlier, the architecture of the MSI leads to some parallax in the L1C product. This parallax is strongly evident in clouds due to their altitude. The core of our method lies in the evaluation and prediction of this parallax effect in the clouds.

The first step of our method is to estimate the orientation and magnitude of the parallax effect based on information such as the satellite's orbital angle and position at the time of acquisition, available in the L1C product. The parallax effect is then computed independently for each pair of bands. Subsequently, we compare the theoretical assumptions with our measurements to determine whether the cloud is real or not. To explain the entire methodology, we need to take a closer look at the MSI first.

4.1 MSI description and geometrical assumptions

The MSI is made of 12 successive sensors each themselves containing 10 bands that capture light reflected by the Earth. On Figure 3 taken from [2] is showed

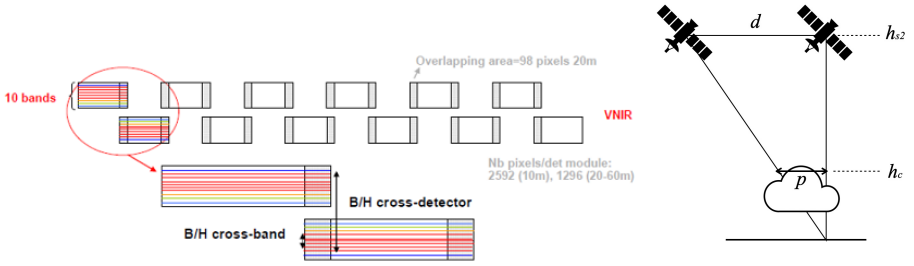

Fig. 3: On the left the Multi-Spectral Instrument arrangement taken from [2]. Bands are stacked other each others producing a parallax effect in the final product. Also, sensors overlap and the order of the bands is reversed for each neighbouring sensor. The order of the visible parallax effect is reversed. On the right a representation of the acquisition between two bands. On one satellite, each band is acquired with a time delay. During this time, the satellite travels a distance d. For elevated objects such as clouds or planes, the pixels of the same object are not at the same location on the output image.

the staggered arrangement of the 12 individual sensors. Each sensor has an overlapping zone with his neighbours. Also all upper sensors have the same bands arrangement, lowers ones also have a similar arrangement but in the reverse order compare to the upper ones. The parallax effect arises from this specific band arrangement on each sensor. At the time of acquisition, the first band to receive light is the lowest band, specifically band 2, the blue band. The order of acquisition is describe in Table 1. Note this table is reversed for the lower sensors. Table 1 also provides information about the time delay between the acquisition of band 2 and each of the other bands. From this information, we can make our first assumption regarding the magnitude of the parallax effect for a cloud.

We know the time delay between the blue band and the others. Assume, for now, a static cloud. Regarding Figure 3 we can make the following statement: for a given cloud altitude h_c, the satellite altitude h_{s2} and his orbital speed v_{s2}, the distance d travelled by the the satellite between t_2 (acquisition of the band 2) and t_n (acquisition of band n with $n \in \{1, 3, 4, 5, 6, 7, 8, 8A, 9, 10, 11, 12\}$), the parallax effect p and by applying Thales' theorem we have:

$$\frac{p}{d} = \frac{h_c}{h_{s2}} \tag{1}$$

then we can compute the parallax magnitude in meters:

$$p = \frac{h_c(t_n - t_2)v_{s2}}{h_{s2}} \tag{2}$$

where all parameters are known except for the cloud altitude. We can also assume that $h_{s2} >> h_c$, since clouds typically lie between around 500 meters and 12 kilometers at most, while $h_{s2} = 786$ km. This indicates that the magnitude

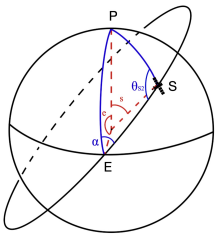

Fig. 4: Representation of the satellite orbital course. We known the side angle a between the equator and the North Pole. The satellite tilt and its latitude. We obtained θ_{S2} with the law of sines.

of the parallax in the clouds depends only on the cloud altitude. Overall, the parallax magnitude varies between 1 and 300 meters.

We have seen that we can estimate the parallax magnitude depending on the cloud altitude. The next step is to assess the orientation of the parallax. Sentinel-2 satellites have a specific orbital angle. Their polar orbit is slightly shifted, descending with an inclination angle of 98.62° at the equator on the day side. Due to the Earth's curvature and given the latitude λ_{S2}, the orientation of the satellite track θ_{S2} is:

$$\theta_{S2} = \arcsin\left(\frac{\sin s \sin \alpha}{\sin(90 - \lambda_{S2})}\right) \tag{3}$$

where s is the angle between the North pole to the equator thus $s = 90°$ and $\alpha = 98.62°$ the inclination angle.

This θ_{S2} provides a good approximation, but we can be more precise. In a satellite image, the y axis does not point exactly North. It points to a direction called Northing. Northing is slightly different from true North and depends on the UTM zone from which the image is acquired due to the curvature of the Earth. It also depends on the position of the satellite within a given UTM zone. The angle will vary if the satellite is at the center of the UTM zone compared to the edges. The Northing angle must then be taken into account. The UTM zone information is stored in the metadata of the L1C product, so it is accessible. From this, we can compute the Δ angle between Northing and θ_{S2}. The trick is to take two points $p_A(x_A, y_A)$ and $p_B(x_B, y_B)$ on the same column but not on the same row, such that $x_A = x_B$ and $y_A \neq y_B$. These coordinates are then converted using the transform function so that p_A and p_B are projected onto the Earth. Finally, Δ is added to θ_{S2} for a better estimation.

4.2 Parallax estimation in the image

We have shown that the parallax effect can be theoretically estimated from the metadata provided with an L1C product image. The parallax effect is particularly visible in clouds. The higher the clouds, the stronger the parallax. One way

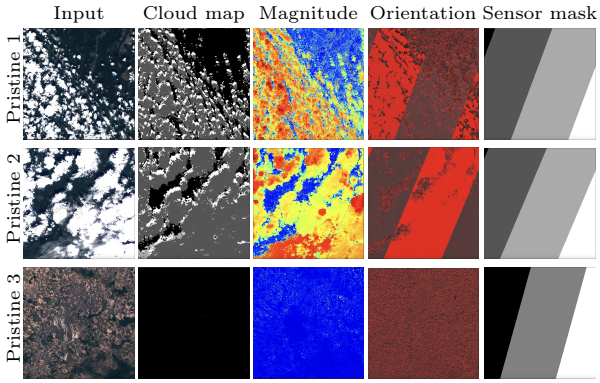

Fig. 5: Some examples of the optical flow on satellite images between band 2 (blue) and band 4 (red). The first two rows show pristine cloudy images, and the last one shows a pristine cloudless image. From left to right, the columns display the input image (13 bands) represented as RGB, the output cloud map obtained by [4], where black pixels represent the background, dark gray represents thick clouds, light gray represents thin clouds, and white represents shadows. Next, the magnitude of the optical flow, which ranges from blue to red, the stronger the magnitude, the more red it gets. Then, the orientation of the optical flow, where red heads North-East and dark gray to the South-West. Finally, the sensor mask indicates the border between each band of acquisition.

to estimate this parallax in images is to compute the optical flow between two bands.

Optical flow is the motion pattern of objects or surfaces in a scene, detected by tracking pixel movement between consecutive video frames, here two bands of the same L1C product. To compute the optical flow between two images we use the implementation of the Gunnar Farneback's algorithm [21] provide by OpenCV.

In Figure 5, we can clearly observe the movement of the clouds from one image to another. Note that on the cloud-free image, this movement is not visible. In Figure 5, we can also observe the separation between each band. This effect is due to the arrangement of the bands from one sensor to another (cf. section MSI). This implies that the orientation of the parallax is phased at 180° to each other. Figure 6 shows the distribution of the optical flow computed only in the clouds. The optical flow is indeed phased at 180° for each upper and lower sensor.

It is then possible to compare theoretical assumptions with experimental evaluations to determine whether a cloud is real or not.

4.3 Forgery detection methodology

In Figure 6, the theoretical parallax is shown as a red bar. This value is close to but different from the orientation of the optical flow due to the wind, a factor previously unconsidered. Our earlier assumptions were based on static clouds,

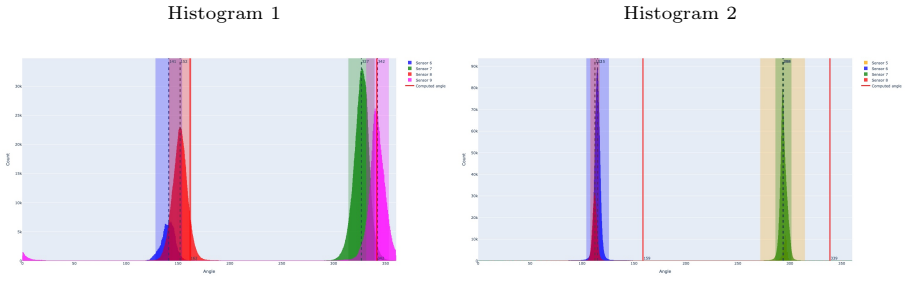

Fig. 6: Histograms of the distribution of the optical flow orientation in the detected clouds. Histogram 1 (resp. 2) is linked to pristine image 1 (resp. 2) in Figure 5. Each color corresponds to a specific sensor. Red bars indicate the computed theoretical angle found with equations 3. First, we can see that odd (resp. even) sensors are grouped together and that they are 180° phased. In histogram 2, the theoretical angle is shifted from the other distributions because of the wind, which changes the orientation of the optical flow.

but clouds move with the wind. According to [3], wind speeds vary between $10, m/s$ and $75, m/s$, depending on altitude. We estimated that the MSI's parallax is around a dozen meters, depending on cloud altitude, with the smallest band resolution being $10, m/pixel$. Thus, both wind and the MSI can produce a similar parallax effect. However, wind orientation is not correlated with the MSI, explaining the difference between theoretical and experimental values in Figure 6. This discrepancy could be resolved with meteorological data, but such data are unavailable with the L1C product.

In a nutshell, based on our assumptions, we can compare the observed and theoretical parallax magnitude and orientation. Essentially, if both are similar, we can assess that no evidence of forgery is found. Here are several practical cases.

There is no parallax observed in the cloud: This event may be very rare, but it can be observed on pristine images. Indeed, it is possible that the sum of the parallax of the MSI and the wind cancel each other out. In this particular case, we can't say it is a fake cloud.

The cloud overlaps two sensors: By computing the optical flow, we should observe a border on the cloud indicating a phase shift of 180° in the parallax orientation. Thus, if there is no border in the clouds, they are indeed fake. Examples are provided in the first two row of Figure 5 on the fourth column.

The cloud is captured by one sensor only: This is the most difficult case because the optical flow here depends on the cloud altitude and wind speed, which are not provided with the L1C product. However, if a cloud is real, there is one property that will always remain: the consistency of the parallax between each band.

We may not know the cloud's altitude or the wind speed, but during one tile acquisition, we can reasonably assume that these factors are fixed. Therefore, the

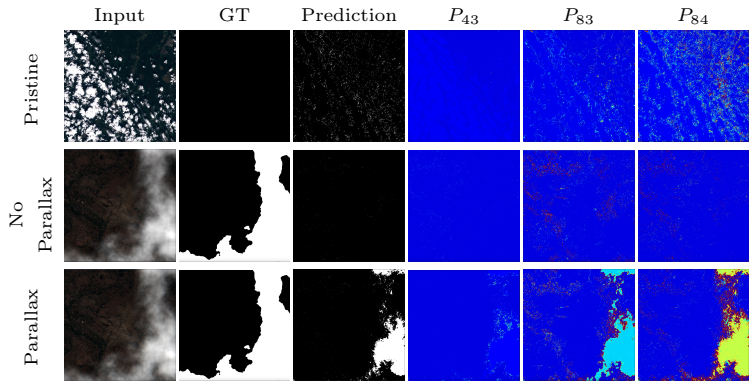

Fig. 7: Some examples of the parallax ratio between pairs of bands and their associated final predictions. From left to right: the input image, the ground truth, the forgery detection map, the ratio between the parallax from band 4 and band 2 with band 3 and band 2, then band 8 and band 2 with band 4 and band 2, and finally band 8 and band 2 with band 3 and band 2. We subtracted the theoretical value of the expected ratio from the three ratio heat maps to display the error distance.

observable parallax should be consistent from one band to another, respecting the time delay between the acquisition of two bands as reported in Table 1 and the order of the bands, whether it is an upper or a lower sensor. This is the main idea of our method.

We know how to approximate the parallax thanks to Equation 1. Based on that, we can compute the ratio between the parallaxes of two given bands and two others. For example, the ratio between p_{24} and p_{23} will always be constant:

$$\frac{p_{24}}{p_{23}} = \frac{t_4 - t_2}{t_3 - t_2} = c \tag{4}$$

Indeed, by computing the ratios of pairs of bands, we should always find the same constant in the parallax. If one of these ratios is not consistent with the theoretical one, the clouds are considered fake. This is the first criterion. The second criterion is the orientation of the parallax. As we have seen in Figure 6, the theoretical orientation and the computed orientation are relatively close, but because of the wind, it can be challenging to set a precise threshold. However, we are certain of the main direction: even if the wind and parallax can compensate for each other, one cannot completely overcome the other. Therefore, we can exclude orientations that are orthogonal (or close to orthogonal) to the theoretical angle.

For the experiment, we compute the optical flow between one band and band 2. We measure the magnitude of the optical flow in the image and the orientation where the magnitude is high enough. If these measurements align with the expected values, the cloud is considered real; otherwise, it is considered fake.

5 Experiments

We subdivide the experiment section in two sections. One dedicated to the cloud detection where the goal is to evaluate how well the fake cloud are detected by a cloud detection network [4]. The second ones aims at evaluating forgery detection methods on the detection of the fake clouds.

The Matthews Correlation Coefficient (MCC) and the F1 score will be used as evaluation metrics. The MCC ranges from -1 (opposite to the ground truth) to 1 (perfect detection). The F1 score ranges from 0 to 1, with 1 indicating perfect detection.

A baseline input-independent method (e.g., random or constant guess) is expected to have a MCC of 0. A method that always detects the entire image as forged will have an F1-score equal to the proportion of forged pixels in the ground truth. The cloud detection is performed over the dataset of the 100 forged tile of the size (10980×10980). Forgery detection will be performed on the cropped images dataset.

5.1 Clouds detection

Our goal is to mislead the UNetMobV2 network created by [4]. The network is evaluated on 50 forged tiles containing fake clouds. Every band of each tile has been forged with the same cloud at different resolutions to match the resolution of the bands.

Around 54% (F1 score) of clouds pixel are detected by the network. Additionally, Figure 8 shows some predicted cloud maps obtained by the network. The detection rate of the fake clouds are similar to one reported in their publication [4]. We successfully mislead the network in most cases, even with very small clouds. Failures generally occur when the cloud is not thick enough, allowing background pixels to be seen and resulting in the cloud not being detected.

5.2 Forgery detection

In our experiment, we incorporate datasets involving cloud generation in addition to a pristine dataset. A dataset of pristine images act as a critical control group, enabling us to accurately assess the false alarm rate, especially with cloudy images. The cropped datasets are evaluated using traditional forensic methodologies as outlined in section 2. As of the time of writing, forensic methodologies for satellite imagery as published by [15, 25–27] are not publicly accessible.

Tab. 2 and Fig. 2 illustrate the suboptimal performance of traditional forensic tools on our dataset. Approaches like Mantranet [34] and TruFor [24] manage to detect some forgeries, but they also generate a high number of false alarms as evidenced in the first row of Fig. 2. Methods that are based on clues specific to photography also understandably fail to detect inconsistencies in the proposed dataset. This is the case of Noisesniffer [22], AdaCFA [11], Mimic [6], and Zero [31].

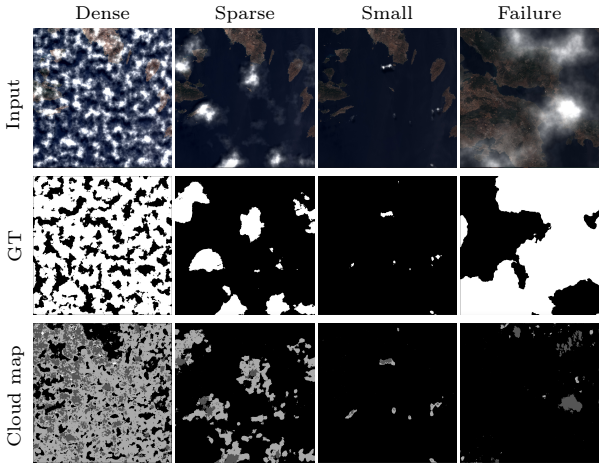

Fig. 8: Some examples of cloud map detection. The first row displays input generated clouds, the second row shows their associated ground truth, and the final row displays the output map obtained by the UNetMobV2 [4]. From left to right, there are dense, sparse, and small fake clouds. The last column is an example of a failure case.

Potential adaptation strategies for these methods could include retraining existing tools, such as Mantranet or TruFor, using Sentinel-2 or other satellites images. This has been proposed in OHNS [16] and shows potential good results, but the performances sharply decrease as soon as data moves away from the training set. Indeed, we can see in Tab. 2 that its results on our two datasets are no better than random. Another strategy could be incorporating the principles from camera-traces-based methods into the satellite image processing pipelines.

Our results are shown in Figure 7. In this figure, we show the computed ratio of pairs of images for all high-resolution bands ($10m/pixel$) to highlight the difference between the theoretical magnitude of the parallax and the one obtained by the optical flow. As expected, in the absence of parallax, no decision can be made. In the case of real clouds, the ratio of the magnitudes in the clouds is of the same order as the theoretical value, allowing us to assess the authenticity of the clouds. However, we observe false alarms on the ground, where no parallax should be observed.

The first aspect to consider is the ratio between the band 4 and band 2 maps and the band 3 and band 2 maps (fourth column). When examining Table 1, particularly focusing on the central wavelength and bandwidth, it becomes evident that in the first comparison, the wavelengths are relatively close to each other. In contrast, for the third comparison, the wavelengths are significantly different, resulting in ground pixels appearing less similar. Consequently, the optical flow is inaccurate, leading to false alarms on the ground. Similar types of errors are also observed in water. This issue appears to be the major drawback of this method. In the third row, we have the same cloud but with some manually

	With parallax		Without parallax	
	MCC	F1	MCC	F1
Ours	**0.3709**	**0.5200**	-0.0001	0.0018
OHNS [16]	0.0003	0.2260	0.0002	0.2078
Splicebuster [17]	0.0038	0.1517	0.0082	0.1865
TruFor [24]	0.0011	0.0538	0.0034	0.1018
AdaCFA [11]	0.0033	0.1744	0.0066	0.2939
Mimic [6]	0.0000	0.0000	0.0000	0.0000
zero [31]	0.0000	0.0000	0.0000	0.0000
Noisesniffer [22]	0.0041	0.0414	0.0140	0.1005
ManTraNet [5, 34]	0.1017	0.0003	0.0007	0.1410
Jade Owl [7]	0.0000	0.0000	0.0000	0.0000

Table 2: Results of forensic tools presented in Sec. 2 on the Perlin (with and without parallax) datasets using the MCC and F1. None of the tested methods are able to detect the satellite forgeries, proving those methods are not suited for satellite images. While our manage to reach better scores.

introduced parallax. Here, the difference between the expected value and the theoretical one is too high, making it clear that the cloud is fake. Additionally, we do not manage to achieve complete detection of the cloud in the transparent parts. In these areas, ground pixels are visible, so by computing the optical flow, we do not observe major differences between bands. The cloud must be thick enough to occlude the ground. These results show the proposed method is better suited to assess the authenticity of clouds rather than to detect tampered ones. Awarded entities may have the capacity to bypass this method. However, we hope that by combining other traces present in the image, such as noise or other visual clues such as shadows, we will be able to detect any kind of cloud that is not consistent with the satellite image pipeline.

6 Conclusion

In this paper, we constructed two satellite forgery datasets using spliced Perlin noise and cloud detection with alpha-matting and splicing. These two methods show promise in creating convincing forgeries that inconspicuously hide information. These forgeries remain undetected by traditional methods and satellite-specific forensic tools. However, satellite images provide meaningful information either in the metadata or in the image itself. Therefore, we can expect some residual effects, such as parallax, to be present in the final product image. Checking the consistency between theoretical assumptions and measurements helps assess the authenticity of images. Moving forward, developing more robust, global-scale satellite image forgery detection techniques will be of the utmost importance.

Acknowledgements

This work has received funding by the European Union under the Horizon Europe vera.ai project, Grant Agreement number 101070093, from ANR under APATE project, grant ANR-22-CE39-0016, and from the DGA/AID PhD grant number 01D22020572. Centre Borelli is also a member of Université Paris Cité, SSA and INSERM.

References

1. Copernicus browser, https://dataspace.copernicus.eu 6
2. Msi description web page, https://sentiwiki.copernicus.eu/web/s2-mission 6, 7
3. Aglietti, G., Redi, S., Tatnall, A., Markvart, T., Walker, S.: Aerostat for solar power generation. Solar Energy. InTech pp. 399–412 (2010) 10
4. Aybar, C., Ysuhuaylas, L., Loja, J., Gonzales, K., Herrera, F., Bautista, L., Yali, R., Flores, A., Diaz, L., Cuenca, N., et al.: Cloudsen12, a global dataset for semantic understanding of cloud and cloud shadow in sentinel-2. Scientific data **9**(1), 782 (2022) 4, 5, 9, 12, 13
5. Bammey, Q.: Analysis and Experimentation on the ManTraNet Image Forgery Detector. Image Processing On Line (2022) 14
6. Bammey, Q.: A contrario mosaic analysis for image forensics. In: ACIVS. Springer Nature Switzerland, Cham (2023) 2, 3, 12, 14
7. Bammey, Q.: Jade owl: Jpeg 2000 forensics by wavelet offset consistency analysis. In: ICIVC. IEEE (2023) 2, 3, 4, 14
8. Bammey, Q.: Synthbuster: Towards detection of diffusion model generated images. OJSP (2023) 2
9. Bammey, Q., Grompone von Gioi, R., Morel, J.M.: Reliable demosaicing detection for image forensics. In: IEEE EUSIPCO 3
10. Bammey, Q., Grompone von Gioi, R., Morel, J.M.: Automatic detection of demosaicing image artifacts and its use in tampering detection. In: IEEE MIPR (2018) 3
11. Bammey, Q., Grompone von Gioi, R., Morel, J.M.: An adaptive neural network for unsupervised mosaic consistency analysis in image forensics. CVPR (2020) 2, 3, 12, 14
12. Bammey, Q., Grompone von Gioi, R., Morel, J.M.: Image forgeries detection through mosaic analysis: the intermediate values algorithm. Image Processing On Line (2021) 3
13. Bammey, Q., Grompone von Gioi, R., Morel, J.M.: Forgery detection by internal positional learning of demosaicing traces. In: IEEE/CVF WACV (2022) 2, 3
14. Bammey, Q., Nikoukhah, T., Gardella, M., Grompone von Gioi, R., Colom, M., Morel, J.M.: Non-semantic evaluation of image forensics tools: Methodology and database. In: IEEE/CVF WACV (2022) 3
15. Bartusiak, E.R., Yarlagadda, S.K., Güera, D., Bestagini, P., Tubaro, S., Zhu, F.M., Delp, E.J.: Splicing detection and localization in satellite imagery using conditional gans. In: IEEE MIPR (2019) 2, 4, 5, 12
16. Cannas, E.D., Baireddy, S., Bestagini, P., Tubaro, S., Delp, E.J.: Enhancement strategies for copy-paste generation & localization in rgb satellite imagery. In: IEEE WIFS (2023) 4, 13, 14

17. Cozzolino, D., Poggi, G., Verdoliva, L.: Splicebuster: A new blind image splicing detector. In: 2015 IEEE WIFS (2015) 2, 3, 14
18. Cozzolino, D., Verdoliva, L.: Noiseprint: A cnn-based camera model fingerprint. IEEE TIFS **15**, 144–159 (2019) 2, 3
19. Czerkawski, M., Atkinson, R., Michie, C., Tachtatzis, C.: Satellitecloudgenerator: controllable cloud and shadow synthesis for multi-spectral optical satellite images. MDPI Remote Sensing (2023) 5
20. Enomoto, K., Sakurada, K., Wang, W., Fukui, H., Matsuoka, M., Nakamura, R., Kawaguchi, N.: Filmy cloud removal on satellite imagery with multispectral conditional generative adversarial nets. In: IEEE/CVF CVPR Workshops (2017) 5
21. Farnebäck, G.: Two-frame motion estimation based on polynomial expansion. In: Image Analysis: 13th Scandinavian Conference, SCIA 2003 Halmstad, Sweden, June 29–July 2, 2003 Proceedings 13. pp. 363–370. Springer (2003) 9
22. Gardella, M., Musé, P., Morel, J.M., Colom, M.: Noisesniffer: a fully automatic image forgery detector based on noise analysis. In: 2021 IEEE IWBF (2021) 2, 3, 12, 14
23. Gudavalli, C., Rosten, E., Nataraj, L., Chandrasekaran, S., Manjunath, B.S.: Seetheseams: Localized detection of seam carving based image forgery in satellite imagery. In: IEEE/CVF CVPR Workshops (2022) 2, 3, 5
24. Guillaro, F., Cozzolino, D., Sud, A., Dufour, N., Verdoliva, L.: Trufor: Leveraging all-round clues for trustworthy image forgery detection and localization. In: IEEE/CVF CVPR (2023) 2, 3, 12, 14
25. Horváth, J., Montserrat, D., Hao, H., Delp, E.: Manipulation detection in satellite images using deep belief networks. In: IEEE/CVF CVPR Workshops (2020) 2, 4, 5, 12
26. Horváth, J., Baireddy, S., Hao, H., Montserrat, D.M., Delp, E.J.: Manipulation detection in satellite images using vision transformer. In: IEEE/CVF CVPR (2021) 2, 4, 5, 12
27. Kalyan Yarlagadda, S., Güera, D., Bestagini, P., Zhu, F.M., Tubaro, S., Delp, E.J.: Satellite image forgery detection and localization using gan and one-class classifier. arXiv e-prints pp. arXiv–1802 (2018) 2, 3, 5, 12
28. Karageogiou, D., Bammey, Q., Porcellini, V., Goupil, B., Teyssou, D., Papadopoulos, S.: Evolution of detection performance throughout the online lifespan of synthetic images. In: Computer Vision – ECCV 2024 Workshops. Springer Nature Switzerland, Cham (2024) 3
29. Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., Aila, T.: Analyzing and improving the image quality of stylegan (2020) 2
30. Li, Y., Bammey, Q., Gardella, M., Nikoukhah, T., Morel, J.M., Colom, M., Von Gioi, R.G.: Masksim: Detection of synthetic images by masked spectrum similarity analysis. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. pp. 3855–3865 (June 2024) 2
31. Nikoukhah, T., Anger, J., Colom, M., Morel, J.M., Grompone von Gioi, R.: ZERO: a Local JPEG Grid Origin Detector Based on the Number of DCT Zeros and its Applications in Image Forensics. IPOL (2021) 2, 3, 12, 14
32. Perlin, K.: An image synthesizer. In: SIGGRAPH (1985) 5
33. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4510–4520 (2018) 4
34. Wu, Y., AbdAlmageed, W., Natarajan, P.: Mantra-net: Manipulation tracing network for detection and localization of image forgeries with anomalous features. In: IEEE/CVF CVPR (2019) 2, 3, 12, 14