ARE CLASSIC FORENSIC TOOLS EFFECTIVE ON SATELLITE IMAGERY?

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

ARE CLASSIC FORENSIC TOOLS EFFECTIVE ON SATELLITE IMAGERY?

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

[1]Université Paris Saclay, ENS Paris Saclay, CNRS, Centre Borelli, France
[2]Kayrros SAS, France

ABSTRACT

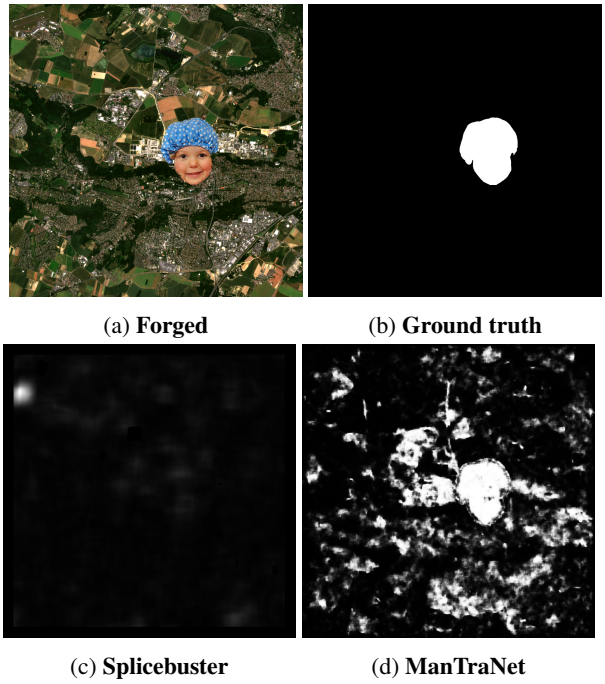
Satellite images are becoming an increasingly important part of our world. Such images are used to forecast the weather, track green house gas emissions, monitor agricultural crop health, and many other applications. Such advances are possible thanks to the free availability of a large number of satellite images. Satellite imagery now plays a key role in many areas, including external security. In this context, it is necessary to question the reliability of this data. Can the authenticity of a satellite image be guaranteed? How can one protect oneself against an entity wishing to hide illegal military material or, conversely, to incite action against another entity by falsely suggesting that it possess such material? If the forensic analysis of photographs has attracted a great deal of academic interest in recent years, this is not yet the case for satellite imagery. In this paper, we propose a methodology to create a very simple but interesting dataset to test the performance of state-of-the-art forensic methods on pristine and manipulated satellite images. Despite the strong performance of such algorithms, satellite images require special attention due to the nature of the images themselves.

***Index Terms*—** forgery detection, satellite imagery, image forensics, splicing dataset, splicing localization

1. INTRODUCTION

More widely available than ever, satellite images are now used for various applications. The sharp rise in the number of Earth observation satellites, thanks to the reduction of their development costs, enables anyone to access satellite images. This accessibility has paved the way to many improvements in satellite imagery. Unfortunately, their widespread use now makes satellite images subject to the risk of forgery as photographs. Satellite images risk tampering in a variety of ways, from splicing of external objects into the image [1] to generative methods [2, 3] that modify regions of an image or even synthesise them entirely.

The authenticity of traditional photographs has been the focus of massive academic research for several years, mainly due to the proliferation of fake news and their impact. There is a large variety of complementary methods to detect image forgeries. Methods analyse traces left by specific parts of the image processing pipeline [7, 8], check more globally the image noise fingerprints for inconsistencies [9, 5, 10] or use neural networks directly trained on forged images to detect them [6, 11]. Are these methods efficient on satellite images? There are claims [1] that satellite forgeries can only be detected with specific methods. While intuitive, these claims remain unproven.



(a) **Forged** (b) **Ground truth**

(c) **Splicebuster** (d) **ManTraNet**

Fig. 1: (a) A Sentinel-2 image forged with salient object from DUT dataset [4]; (b) its ground truth mask; (c) Splicebuster [5] detection; (d) ManTraNet detection [6]. This shows that classic algorithms struggle even on straightforward satellite forgeries. Splicebuster does not detect any forgery and ManTraNet finds it but at a cost of numerous false alarms.

Indeed, satellite image generation has little in common with photographs. The pipelines even differ from one constellation to the other, due to many differences in the sensors and acquisition modes. Furthermore, satellite images are stored in specific formats such as GeoTIFF or JPEG2000. JPEG compression artefacts, a major focus of photographic forensics, are absent in satellite images. Nevertheless, not all forensic tools work on specific traces. The question of their efficiency on satellite images thus remains valid.

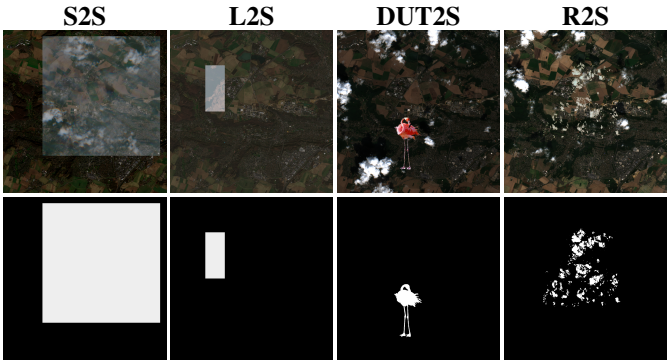


Fig. 2: Examples from the proposed S2S, L2S, DUT2S and R2S datasets (first row) and their corresponding ground truth (second row).

The contribution of this paper is two-fold. We create a database of very simple satellite forgeries, that are easily detectable by the naked eye. The full dataset is made publicly available here: `https://zenodo.org/record/7984723`. This database is then used to test publicly available SOTA classic forensic tools on forged satellite images. We show that while these methods can detect some easily detectable forgeries, they struggle on most of the images, as seen in Fig. 1.

2. RELATED WORK

Forensics methods of photographic images can be broadly divided into two categories. Historically, forensics tools worked on specific traces in images left by the image processing pipeline. These traces are altered when the image is modified, or different traces can even be imported from other images during splicing. Analysis of these traces can thus lead to inconsistent areas in the forged regions. Many SOTA methods still apply this paradigm with more advanced ideas, such as Bammey [7] which analyses demosaicing traces, or ZERO [8] which detects JPEG block dephasing. Some methods also detect more generic traces, such as Noisesniffer [9] that analyses the noise level of an image, or Splicebuster [5] and Noiseprint [10] that extract and analyze a fingerprint of the image. A second, more recent category of methods are those trained directly to detect forgeries, with the most well-known being ManTraNet [6, 11] and its variants. They consist of end-to-end neural networks, trained on forged images to learn to directly detect forgeries.

Only recently has attention been given to satellite forensics. In [1], an auto-encoder is trained on pristine satellite images, then a support vector machine is used on the encoded features of an image to distinguish pristine and forged parts. Similarly, conditional generative adversarial networks [12], deep belief networks [13] or transformers [14] can be used to encode images prior to check their authenticity.

The efficiency of these methods, however, can hardly be tested. Despite the increasing number of satellite data available, there are no publicly available datasets of spliced satellite images. In most cases, authors of a satellite image forgery detection tools create their own non-public datasets by copy-pasting salient objects from other images onto satellite images. The objects are usually sourced from photographs, they are thus easier to detect.

3. METHODOLOGY

3.1. Data creation

We selected 100 Sentinel-2 L2A images to obtain 1000×1000-pixels images with the red, green and blue bands at a resolution of 10 metres per pixel. Using these pristine images, we created four basic forgery datasets. The aim here is not to create realistic images; most of the forgeries are actually evident to the naked eye, as seen in Fig. 2. Quite the opposite, these images are created to show that even roughly done forgeries are challenging for SOTA forensic methods.

Sentinel to Sentinel (S2S) In this dataset, a square patch of random size and position is selected from a Sentinel-2 image and pasted to another Sentinel-2 image. This dataset is the closest (from the traces viewpoint) to what might be done by malicious entities, and is probably the most challenging of the four, as forged regions may present similar characteristic than pristine ones, except at the border.

Landsat to Sentinel (L2S) The creation of this dataset is similar to the S2S dataset discussed above, but the source region is taken from a Landsat-8 image instead of a Sentinel-2 image. As the resolution of the RGB bands of Landsat-8 is 30 metres per pixel, the forged regions present different characteristics than the pristine image and might thus be more easily detected.

DUT to Sentinel (DUT2S) In this dataset, the forged regions are no longer copied from satellite images. Instead, we use salient objects from the DUT [4] dataset, consisting of 8-bit JPEG images. To match the pixel range value of JPEG files, the 16-bit Sentinel-2 images were requantized to 8-bit prior to forgery.

Real to Sentinel (R2S) The semantic contents of DUT2S images is markedly incoherent, including giant people or flying cars. Our last dataset is an attempt to generate slightly more realistic images. For this, sample clouds from [15] and planes from [16] were segmented and pasted on Sentinel images following the same protocol as for DUT2S.

3.2. Evaluation

We test the performance of publicly available SOTA methods Splicebuster [5], Noiseprint [10], Bammey [7], ZERO [8], Noisesniffer [9], and ManTraNet [6] using the implementation in [11]. We evaluate the results of these methods with the Matthews Correlation Coefficient (MCC) [17], a metric that ranges from -1, opposite of the ground truth, to 1 for a perfect detection. Input-independent methods, such as always

	S2S	L2S	DUT2S	R2S
Splicebuster [5]	0.04	0.07	0.13	0.09
Noiseprint [10]	0.12	0.05	0.14	0.09
Bammey [7]	0.02	0.01	0.01	0.00
ZERO [8]	0.00	0.00	0.45	0.00
Noisesniffer [9]	0.00	0.00	0.01	0.00
ManTraNet [6, 11]	0.11	0.09	0.38	0.38

Table 1: Results of forensic tools presented in Sec. 3.2 on our four datasets using the MCC. While some methods detect forgeries in the DUT2S and R2S datasets, where the forgeries are taken from real photographs, detections are poor in the true satellite forgeries of S2S and L2S, proving those methods are not suited for satellite images.

Splicebuster	Noiseprint	Bammey	ZERO	Noisesniffer	ManTraNet
39	98	91	0	4	41

Table 2: For each method, number of authentic images where there is at least least one false alarm, among a total of 100 pristine images of size 1000×1000. A false alarm is defined here as a connected region of 1024 pixels over which all pixels are detected with a confidence over 0.8.

returning a white mask, have an expected score of 0. The MCC is computed for each image, and then averaged over each dataset. Our experiments focus on the performance of photographic forensic tools on satellite images. Methods specific to satellite forgeries such as [1, 12, 13, 14] were not used in the comparison, as their code was not publicly available at the time of writing.

4. RESULTS

Results are presented in Table 1. With few exceptions, SOTA photographic forensic tools yield globally bad results, failing to detect most forgeries. The lack of detections of Noisesniffer [9] are not surprising, as it estimates the noise level of an image using knowledge about camera models, which have nothing in common with satellite images. Likewise, ZERO and Bammey predictably fail to yield any detection on the S2S, L2S and R2S datasets, as satellite images lack the demosaicing (for Bammey) and JPEG compression (for ZERO) steps these methods use to find forgeries. On DUT2S, ZERO nevertheless detects some forgeries, as the method picks up the compression traces of the forged image. Likewise, ManTraNet [6] detect many forgeries in the DUT2S and R2S datasets, whose forgeries come from real images. In the other two datasets, while it can behave differently in the forged regions, as seen in Fig. 3, it fails to detect it. Furthermore, given its large number of false positive seen in Tab. 2, detections cannot always be distinguished from its false positives. The bad results of Noiseprint [10] are expected, due to its being trained on photographs. However, those of Splicebuster [5] are more surprising. Indeed, as Splicebuster extract high-frequency information about the image without assumptions of any generation model, it should theoretically be able to detect forgeries when some traces are different for those of pristine area, which is the case in all but the S2S dataset.

It follows that photographic forensic methods are not designed to, and globally unable to detect forged satellite images. The methods that make detections only do with the simplest forgeries, at the cost of a high number of false alarms even on pristine images. Despite these results, some of these methods could probably be adapted to satellite images, be it by retraining the learning-based methods on suitable data, or by using the ideas behind camera-traces-based methods and applying them to satellite processing pipelines.

5. CONCLUSION

In this paper, we have proposed four simple satellite imagery datasets. Our experiments show that even though these forgeries are voluntarily not sophisticated and lack semantic coherence, classic SOTA forensic methods struggle to detect them. This confirms the intuition that forensic tools specific to satellite imagery are needed, due to the different nature of satellite imagery compared to photography. Future work should thus focus on trying to adapt existing methods, port their ideas to satellite imagery pipelines, or design entirely new, specific methods. Such work can be made difficult by the variety of processing pipelines for different satellite constellations, with even fewer standards and common points than between cameras. However, this is needed to prevent a total vulnerability to satellite image forgeries.

6. ACKNOWLEDGEMENT

This work has received funding by the European Union under the Horizon Europe vera.ai project, Grant Agreement number 101070093, and from ANR under APATE project, grant ANR-22-CE39-0016. Centre Borelli is also a member of Université Paris Cité, SSA and INSERM. Supported by PHD Grant DGA/AID (No 01D22020572)

7. REFERENCES

[1] S. Kalyan Yarlagadda, D. Güera, P. Bestagini, FM Zhu, S. Tubaro, and EJ Delp, "Satellite image forgery detection and localization using gan and one-class classifier," *arXiv e-prints*, 2018.

[2] Y. Zhan, D. Hu, Y. Wang, and X. Yu, "Semisupervised hyperspectral image classification based on generative adversarial networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 212–216, 2018.

[3] L. Abady, M. Barni, A. Garzelli, and B. Tondi, "Gan generation of synthetic multispectral satellite images," in *Image and Signal Processing for Remote Sensing XXVI*. SPIE, 2020.

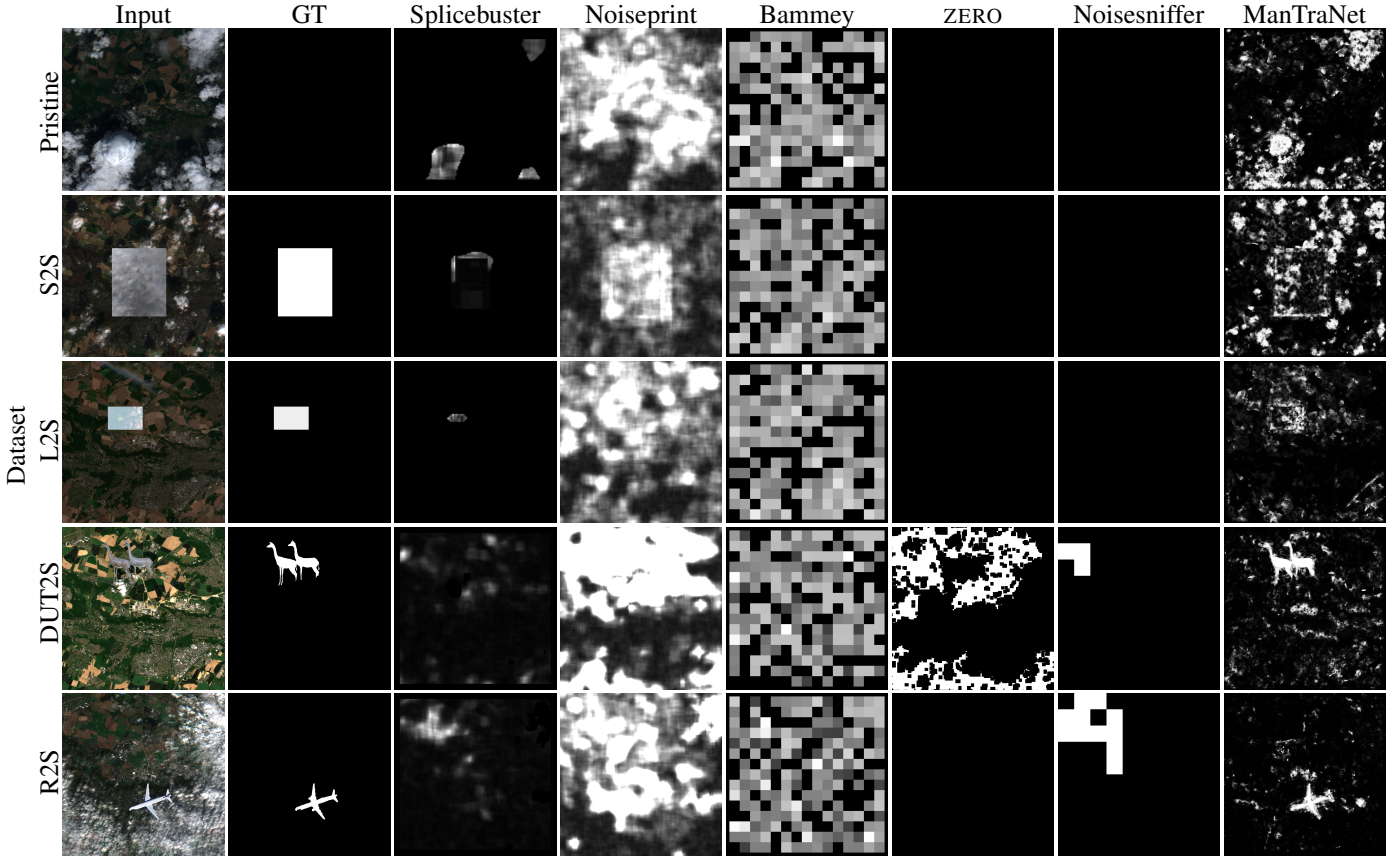


Fig. 3: Presentation of the results of classic forensic tools for one image for each dataset. Some methods, like Noiseprint [10] and ManTraNet [6] detect some forgeries, at the cost of many false alarms. Detections are otherwise poor, especially on the pure satellite splicing forgeries from S2S and L2S.

[4] L. Wang, H. Lu, Y. Wang, M. Feng, D. Wang, B. Yin, and X. Ruan, "Learning to detect salient objects with image-level supervision," in *IEEE/CVF CVPR*, 2017.

[5] D. Cozzolino, G. Poggi, and L. Verdoliva, "Splicebuster: A new blind image splicing detector," in *2015 IEEE WIFS*, 2015.

[6] Y. Wu, W. AbdAlmageed, and P. Natarajan, "Mantra-net: Manipulation tracing network for detection and localization of image forgeries with anomalous features," in *IEEE/CVF CVPR*, 2019.

[7] Q. Bammey, R. Grompone von Gioi, and J.-M. Morel, "An adaptive neural network for unsupervised mosaic consistency analysis in image forensics," *CVPR*, 2020.

[8] T. Nikoukhah, J. Anger, M. Colom, J.-M. Morel, and R. Grompone von Gioi, "ZERO: a Local JPEG Grid Origin Detector Based on the Number of DCT Zeros and its Applications in Image Forensics," *IPOL*, vol. 11, pp. 396–433, 2021.

[9] M. Gardella, P. Musé, J.-M. Morel, and M. Colom, "Noisesniffer: a fully automatic image forgery detector based on noise analysis," in *2021 IEEE IWBF*, 2021.

[10] D. Cozzolino and L. Verdoliva, "Noiseprint: A cnn-based camera model fingerprint," *IEEE TIFS*, vol. 15, pp. 144–159, 2019.

[11] Q. Bammey, "Analysis and Experimentation on the ManTraNet Image Forgery Detector," *IPOL*, vol. 12, pp. 457–468, 2022.

[12] ER Bartusiak, SK Yarlagadda, D. Güera, P. Bestagini, S. Tubaro, FM Zhu, and EJ Delp, "Splicing detection and localization in satellite imagery using conditional gans," in *2019 IEEE MIPR*, 2019.

[13] J. Horváth, DM Montserrat, H. Hao, and EJ Delp, "Manipulation detection in satellite images using deep belief networks," in *IEEE/CVF CVPR Workshops*, 2020.

[14] J. Horváth, S. Baireddy, H. Hao, DM Montserrat, and EJ Delp, "Manipulation detection in satellite images using vision transformer," in *IEEE/CVF CVPR*, 2021.

[15] S. Mohajerani, T. A. Krammer, and P. Saeedi, "A cloud detection algorithm for remote sensing images using fully convolutional neural networks," in *2018 IEEE 20th International Workshop on MMSP*, 2018.

[16] D. Saini, "Satellite images for computer vision tasks," 2022, Kaggle.

[17] D. Chicco and G. Jurman, "The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation," *BMC genomics*, 2020.