Positional Learning to reverse-engineer the camera processing pipeline and detect image forgeries

Quentin Bammey, Rafael Grompone von Gioi, Jean-Michel Morel

Positional Learning to reverse-engineer the camera processing pipeline and detect image forgeries

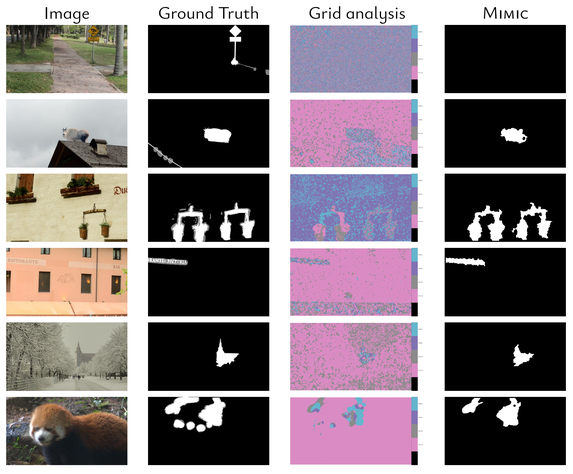
Quentin Bammey[1], Rafael Grompone von Gioi[1], Jean-Michel Morel[2]
https://bammey.com/

1. : ENS Paris-Saclay, Université Paris-Saclay, Centre Borelli, CNRS, France
2. Department of Mathematics, City University of Hong Kong, Kowloon Tong, Hong Kong

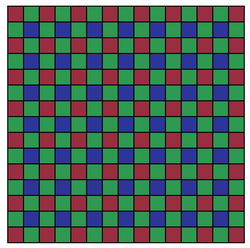
Abstract

As conventional tools and generative-AI-based methods alike can alter images in visually convincing ways, image editing is no longer reserved to experts. However, this ease of manipulation has given rise to malicious manipulation of images, resulting in the creation and dissemination of realistic but fake content to spread **disinformation** online, wrongfully incriminate someone, or commit fraud. The **detection** of such **forgeries** is paramount in exposing those deceitful acts. One approach involves **reverse-engineering** the image signal processing pipeline, to detect and localize inconsistencies. In this context, **positional learning** has emerged as a promising and explainable approach to **reveal underlying traces** of the signal processing pipeline. We show how it can be used to detect forgeries from inconsistencies in the image mosaic or compression history.

Results

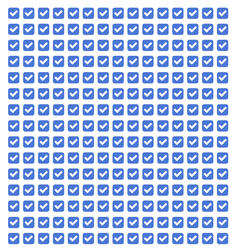


Train a CNN to detect the modulo-$(2, 2)$ position of each pixel

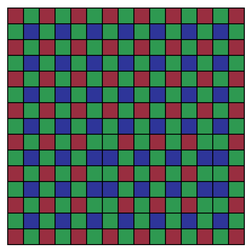


Sampling mosaic of an authentic image

Trained/expected output on an authentic image

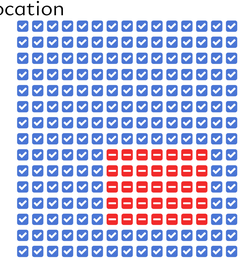


All pixels are estimated at their correct location



On this forged image, the sampling mosaic is altered

The CNN relies on the mosaic to make its estimation, so the positional information is dephased as well



The forgery is detected as a significant error in the CNN output.

- Train the CNN on authentic images only, to estimate the modulo-$(2, 2)$ position of each pixel.
- The network will need to learn to rely on **demosaicing artefacts** to extract the **positional information**.
- At **inference**: image forgeries lead to inconsistencies in the mosaic, which lead to errors in the output.
- Detect errors in the output to know the image is **forged**!

Automatic *a contrario* detection to control the tolerated rate of false positives

- Globally: ratio p_0 of pixels that have their position incorrectly detected
- In a given rectangle: k out of n pixels wrongly detected
- **Is there a forgery in this rectangle?**
- Threshold on the expected **number of false alarms** to statistically control the tolerated rate of false positives:

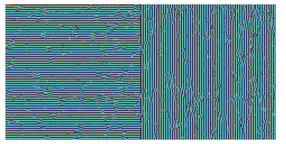
$$NFA = n_{tests} \cdot \mathbb{P}(k_{wrong} \geq k) = 2(X \cdot Y)^2 \sum_{i=k}^{n} \binom{n}{k} p_0^i (1 - p_0)^{n-i}$$

- Threshold at 10^{-3} $\Rightarrow$ expect one false detection every 1000 images

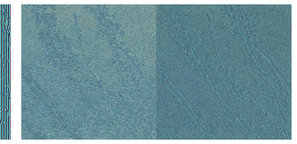
Positional Learning also helps detect generated images

(a) Stable diffusion image

(b) Real image from the Raise dataset



(c) Detected position of each pixel modulo 8, horizontally and vertically, on a Stable Diffusion image



(d) Detected position of each pixel modulo 8, horizontally and vertically, on a natural image from the Raise dataset

References

- Quentin Bammey, Rafael Grompone von Gioi, and Jean-Michel Morel. "An Adaptive Neural Network for Unsupervised Mosaic Consistency Analysis in Image Forensics". In: *CVPR*. 2020
- Quentin Bammey, Rafael Grompone von Gioi, and Jean-Michel Morel. "Forgery Detection by Internal Positional Learning of Demosaicing Traces". In: *WACV*. 2022
- Quentin Bammey. "A Contrario Mosaic Analysis for Image Forensics". In: *ACIVS*. Springer Nature Switzerland, 2023

Acknowledgements

This work has received funding by the European Union under the Horizon Europe vera.ai project, grant agreement number 101070093, and by the ANR under the APATE project, grant number ANR-22-CE39-0016.
Centre Borelli is also a member of Université Paris Cité, SSA and INSERM.

https://www.veraai.eu/ https://improved-anr.github.io/