Image forensic tools

Marina Gardella

Free Software and Reproducible Research Course
November 23rd, 2022

CENTRE BORELLI
école normale supérieure paris—saclay

Meet the DEFALS team at the Centre Borelli

Yanhao Li
PhD student

Thibaud Ehret
Post-doctoral researcher

Quentin Bammey
Post-doctoral researcher

Tina Nikoukhah
Post-doctoral researcher

Marina Gardella
PhD student

Pablo Musé
Senior researcher

Jean-Michel Morel
Senior researcher

Rafael Grompone
Senior researcher

Miguel Colom
Senior researcher

Collaborations and projects

- 2015–2019: Partnership with **Surys** (Authentication)
- 2017–: Partnership with the **Police Scientifique**
- 2017–2021: **ANR/DGA Defals Challenge**: image forgery detection
- 2018–: Partnership with **Agence France Presse** (Journalists and fact-checkers)
- 2021: **Envisu4** project for an online fact-checking platform
- 2022–: **Vera.ai** project for verification assisted by AI.
- 2022–: **Apate** project, A Prototype deepfake Assessment Toolbox for forensic Experts

Is seeing believing?

Photomontages throughout history

Is seeing believing?

Photomontages throughout history

Is seeing believing?

Photomontages throughout history

Is seeing believing?

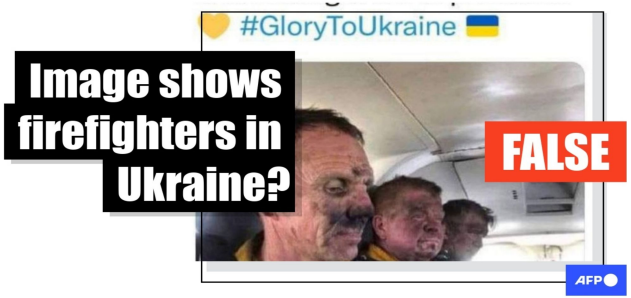
Photomontages throughout history

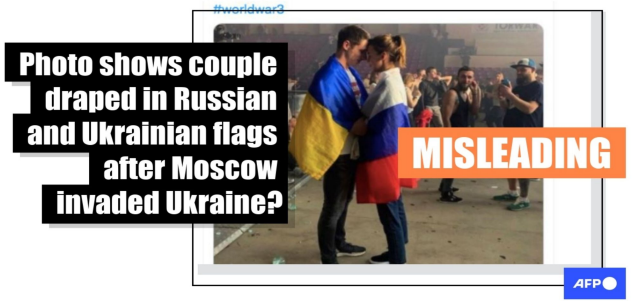
Is seeing believing?

Photomontages throughout history

Is seeing believing?

Fake images and fake news





How to verify an image?

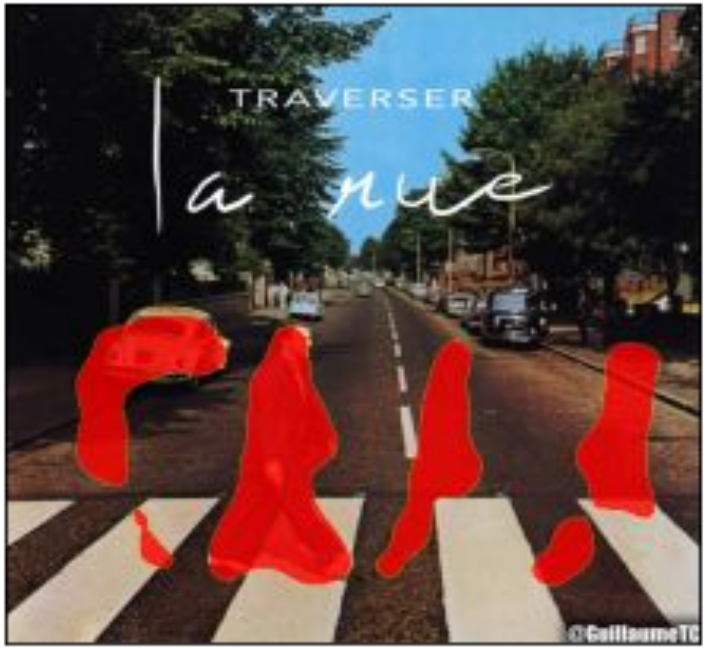
How to verify an image?

First step: reverse image search

Imágenes visualmente similares

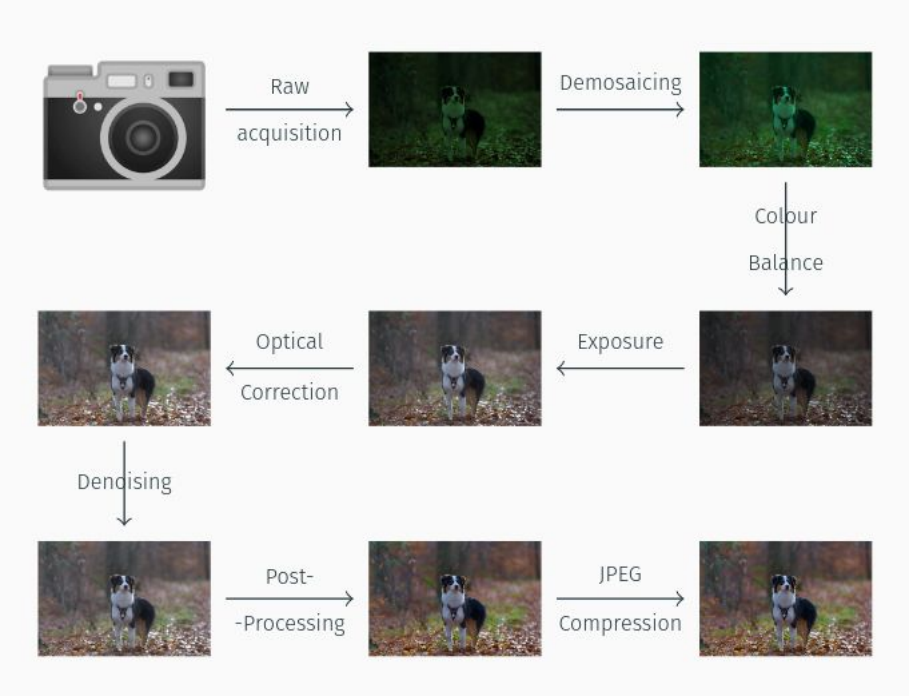
How to verify an image?

Forgery detection methods



Forgery detection methods: how do they work?

From its acquisition to its digital storage, an image undergoes a certain number of **operations** that will give it its final form. Each of these operations embeds in the image a **specific trace** that we can analyze.

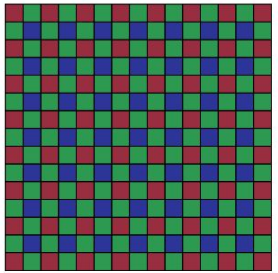


Methods based in the mosaicing pattern

(a) In a raw image, each pixel is only sampled in one color channel.

(b) Demosaicing interpolates the missing colors.

The raw image is not a 3-channel color image, but each pixel is sampled in one color, according to a color filter array (CFA).

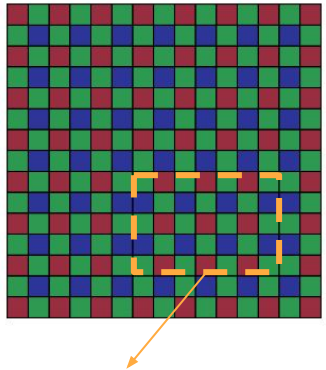


Bayer CFA

Methods based in the mosaicing pattern

Forgery detection -> look for shifts in the CFA pattern

In case of copy-paste, there is a chance of 3 4 that the CFA pattern of the pasted region is not aligned with that of the original image.



Methods based in the JPEG artefacts

The JPEG algorithm is the most common method for compression of digital photography.

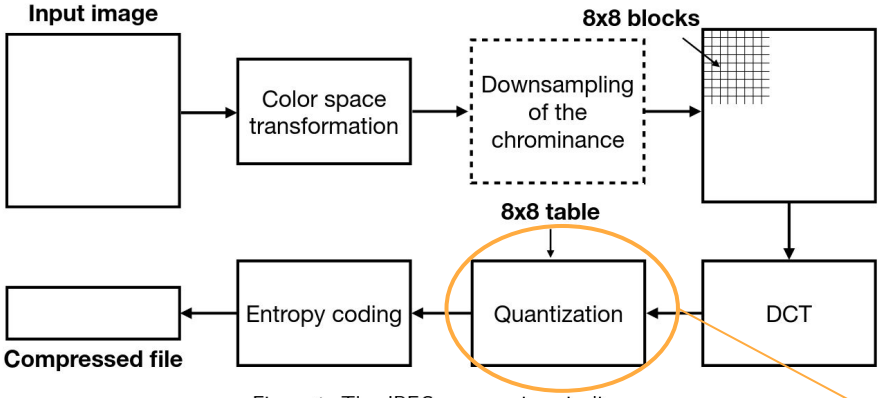
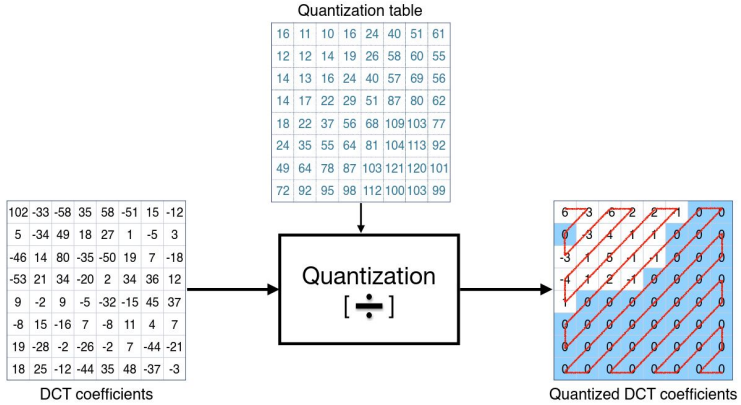
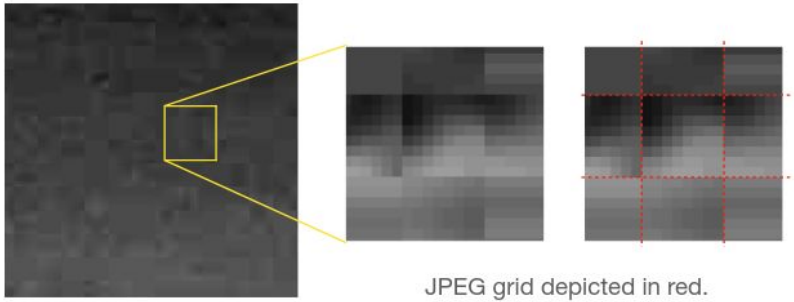


Figure 1: The JPEG compression pipeline.



Methods based in the JPEG artefacts

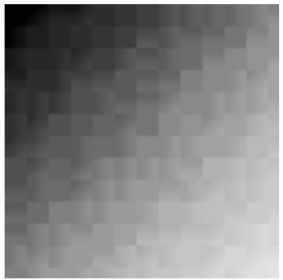
In JPEG encoding, the 8×8 blocking and quantization steps lead to the appearance of discontinuities at the edge of the blocks of the decompressed image.



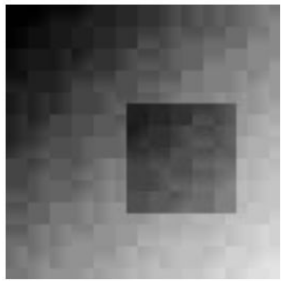
JPEG grid depicted in red.

Forgery detection -> look for shifts in the JPEG grid pattern

In case of copy-paste, there is a chance of 63/64 that the JPEG pattern of the pasted region is not aligned with that of the original image.



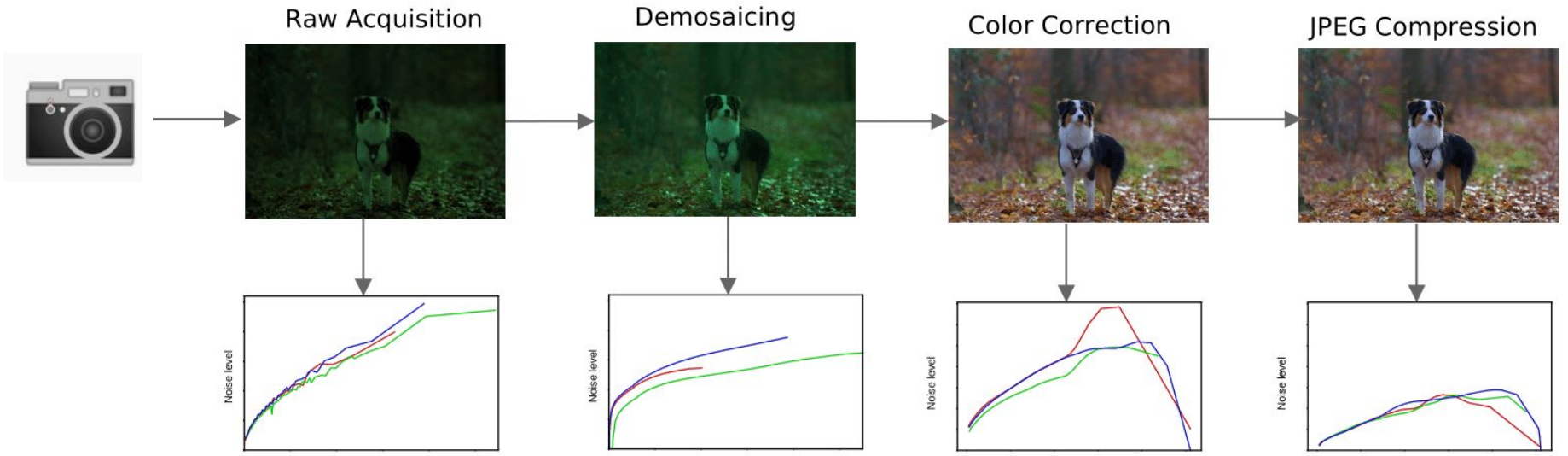
Authentic



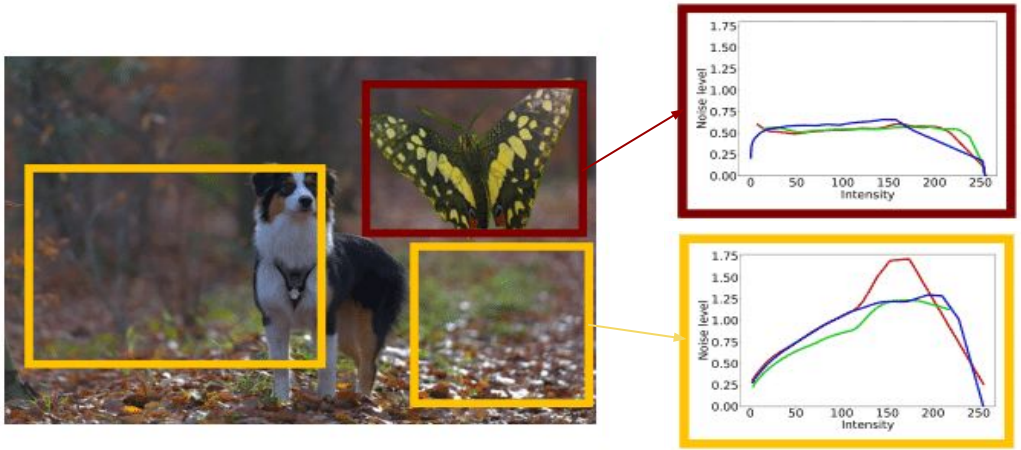
Local shifted grid.

Methods based in noise analysis

Each step of the camera processing pipeline has an impact on the noise curves of the image



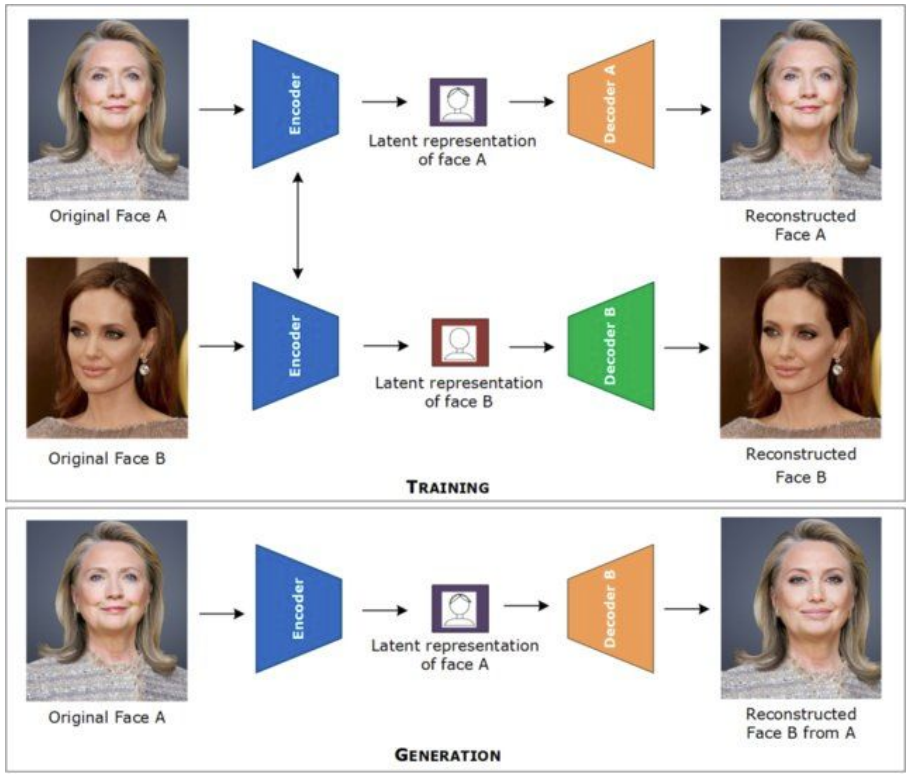
Methods based in noise analysis



Noise models differ between background donor images. The resulting tampered image presents local inconsistencies in the noise model.

Deepfake detection

Deepfakes are generated using deep learning model (GANs).
They do not undergo the camera processing pipeline.



Deepfake detection

Still, GANs leave artificial fingerprints making it possible to detect them.

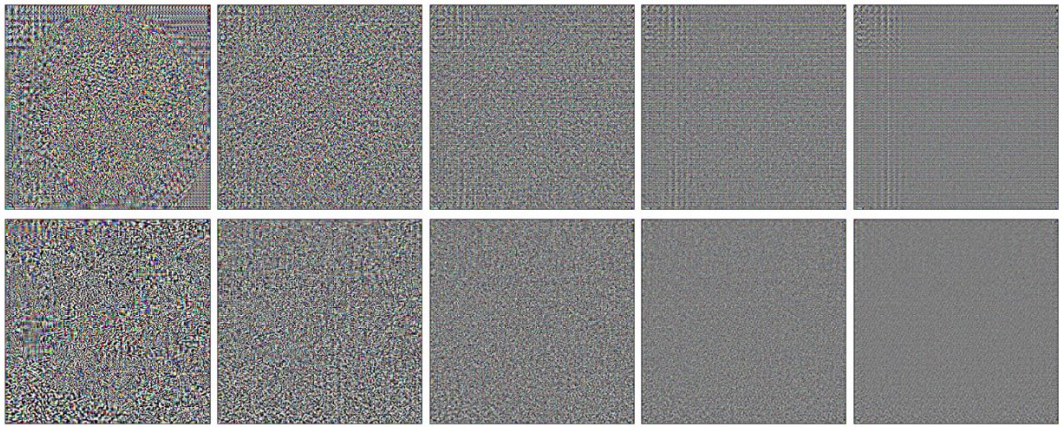


Figure 2. Cycle-GAN o2a (top) and Pro-GAN kitchen (bottom) fingerprints estimated with 2, 8, 32, 128, 512 residuals.

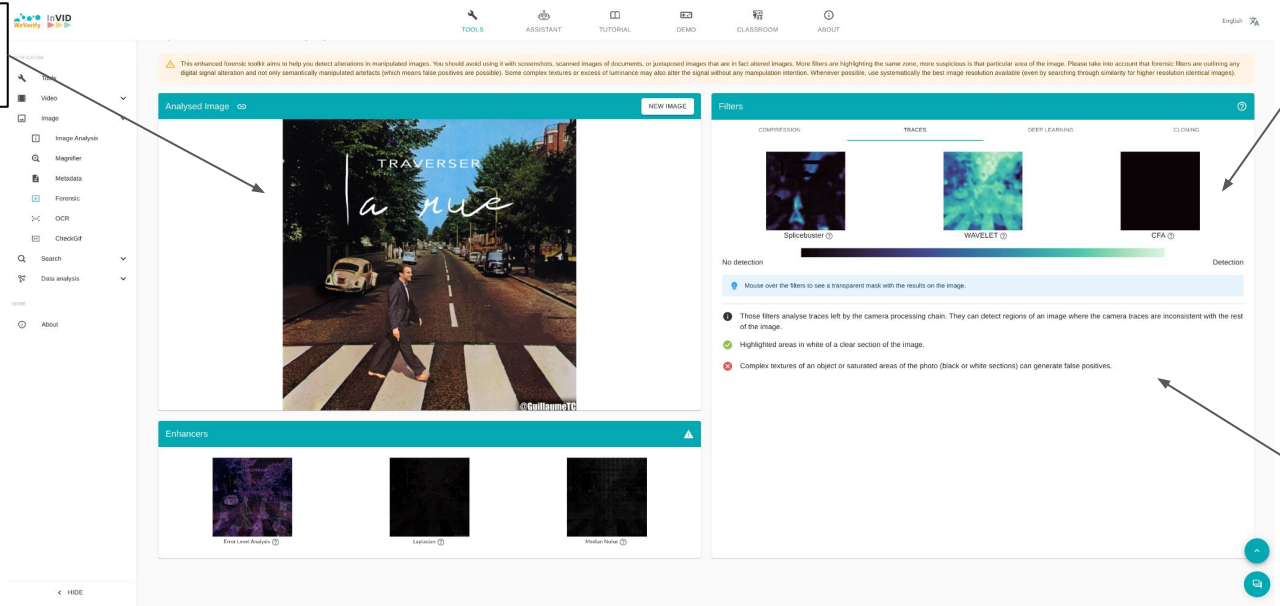
Where can I find these tools?

Public image verification platforms

- InViD-WeVerify plug-in
- REVEAL Image Verification Assistant
- FotoForensics and Forensically online tools
- Ghiro open-source project
- Authenticate software by Amped.

What do these platforms provide?

submit a suspicious image



detection results using different forensic filters

a brief explanation on how to interpret the results and the limitations

No source codes
No detailed explanation of the methods

Why do we care about reproducible, open and explainable research?

1. Because we are doing **scientific research**!

Why do we care about reproducible, open and explainable research?

2. To use the results of our methods in **courtrooms**, the following factors be considered

- Has the technique been tested in actual field conditions (and not just in a laboratory)?
- Has the technique been subject to peer review and publication?
- What is the known or potential rate of error?
- Do standards exist for the control of the technique's operation?
- Has the technique been generally accepted within the relevant scientific community?

Why do we care about reproducible, open and explainable research?

3. To **debunk fake news**:

- the justification and exact details of the methods must be available,
- methods need to be explicable to several publics so that they can understand the results,
- they need to be able to conduct their own investigations,
- limitations need to be exhaustively documented.

Conclusions

- Don't believe everything you see !

- Reproducible, open source and explainable research is needed in the field of image forensics so that the methods can be used at courtrooms and for fake news debunking. There are still a lot of efforts that need to be done in this sense.