Image Forgery Detection through Demosaicing Analysis: Unconcealment of a Signature

Quentin Bammey
Under the supervision of Rafael Grompone von Gioi and Jean-Michel Morel

Centre Borelli @ École Normale Supérieure Paris-Saclay

Fake images are everywhere

Fake images are everywhere

The 36.47 g/mol in HCL
@_Hydrochloric

This science fiction movie was first realised in 1963. 58 years later, here comes the real disease caused by the "Omicron' variant of covid 19. coincidence ??

Fake images are everywhere

https://factcheck.afp.com/http%253A%252F%252Fdoc.afp.com%252F9TU3HE-1

Fake images are everywhere

https://factuel.afp.com/non-cette-image-ne-montre-pas-des-femmes-voilees-devant-la-caf-de-rosny-sous-bois

Fake images are everywhere

https://factuel.afp.com/non-cette-image-ne-montre-pas-des-femmes-voilees-devant-la-caf-de-rosny-sous-bois

How are forgeries created?

How are forgeries created?

How are forgeries created?

How are forgeries created?

How are forgeries created?

Images can be taken out of context

DO NOT COMPLY.

Ginny "Let's Go Brandon!" Robin...
@realginnyrobins

Follow

Protest in Vienna against COVID restrictions, including mandatory vaccine passports.

People are waking up and we will all do this if we have to.

1:12 PM - 21 Nov 2021

Images can be taken out of context

RUSSIA UNREST 1991

Overview Get price

Hundreds of thousands of protesters pack Moscow's Manezh Square next to the Kremlin, Sunday, March 10, 1991, demanding the Soviet President Mikhail Gorbachev and his fellow Communists give up power. The crowd, estimated at 500,000 was the biggest anti-government demonstration in the 73 years since the Communists took power, and came a week before the nationwide referendum on Gorbachev's union treaty. (AP Photo/Dominique Mollard)

Use Information This content is intended for editorial use only. For other uses, additional clearances may be required.

ID:	9103100617
Creation Date:	March 10, 1991 10:45:23 AM

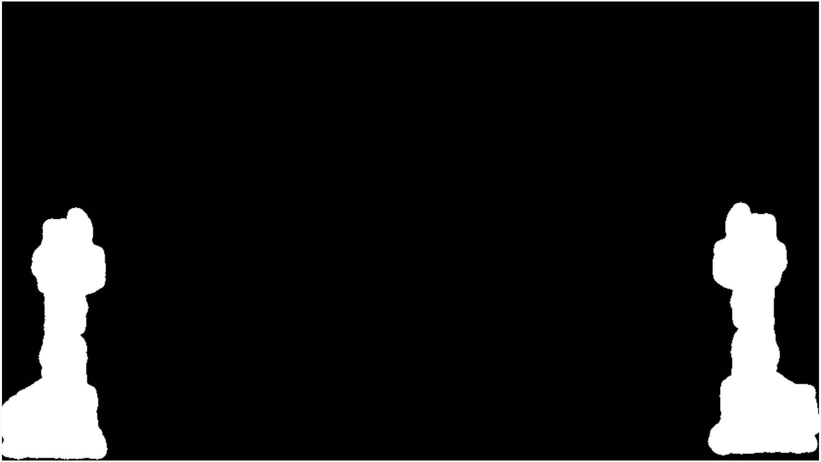
https://factcheck.afp.com/http%253A%252F%252Fdoc.afp.com%252F9T79CJ-1

Internal copy-moves

P. Korus and J. Huang, “Evaluation of random field models in multi-modal unsupervised tampering localization,” in *Proc. of IEEE Int. Workshop on Inf. Forensics and Security*, 2016, P. Korus and J. Huang, “Multi-scale analysis strategies in prnu-based tampering localization,” *IEEE Trans. on Information Forensics and Security*, 2017

Internal copy-moves

Internal copy-moves



D. Cozzolino, G. Poggi, and L. Verdoliva, "Efficient dense-field copy-move forgery detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 11, pp. 2284–2297, 2015

T. Ehret, "Automatic Detection of Internal Copy-Move Forgeries in Images," *Image Processing On Line*, vol. 8, pp. 167–191, 2018, `https://doi.org/10.5201/ipol.2018.213`

Semantic detections?

P. Korus and J. Huang, "Evaluation of random field models in multi-modal unsupervised tampering localization," in *Proc. of IEEE Int. Workshop on Inf. Forensics and Security*, 2016, P. Korus and J. Huang, "Multi-scale analysis strategies in prnu-based tampering localization," *IEEE Trans. on Information Forensics and Security*, 2017

Semantic detections?

"Day before yesterday I saw a rabbit[1], and yesterday a deer[2]."[3]

[1] Image credits https://www.pickpik.com/snowshoe-hare-rabbit-hare-wildlife-nature-outdoors-155309

[2] Image credits https://www.needpix.com/photo/download/1481788/deer-snow-whitetail-doe-outdoors-winter-cold-wildlife-nature

[3] R. F. Young, "The dandelion girl," *The Saturday Evening Post*, Apr. 1961

Semantic detections?

"Day before yesterday I saw a rabbit[1], and yesterday a blue deer[2]."[3]

[1] Image credits https://www.pickpik.com/snowshoe-hare-rabbit-hare-wildlife-nature-outdoors-155309

[2] Image credits https://www.needpix.com/photo/download/1481788/deer-snow-whitetail-doe-outdoors-winter-cold-wildlife-nature

[3] R. F. Young, "The dandelion girl," *The Saturday Evening Post*, Apr. 1961

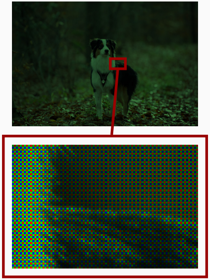
For a trace-based analysis of images

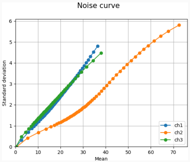
Galileo: How would it be if your Highness were now to observe these impossible as well as unnecessary stars through this telescope?
The Mathematician: One might be tempted to reply that your telescope, showing something which cannot exist, may not be a very reliable telescope, eh?

Image Processing Pipeline

Raw

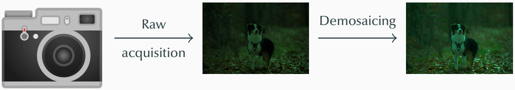
acquisition

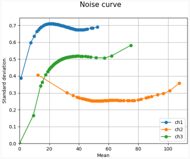




The image processing pipeline *leaves traces* on the image.
We *analyse* those traces to *reverse-engineer the pipeline*.

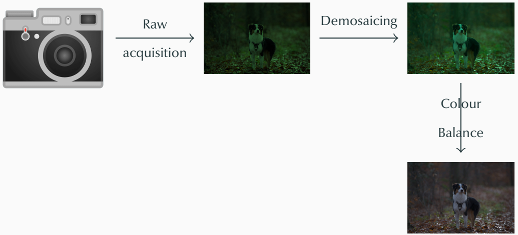
Image Processing Pipeline

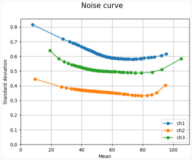




The image processing pipeline *leaves traces* on the image.
We *analyse* those traces to *reverse-engineer the pipeline*.

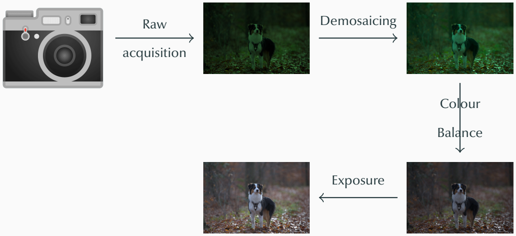
Image Processing Pipeline

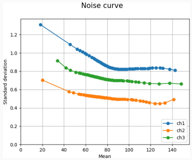




The image processing pipeline *leaves traces* on the image.
We *analyse* those traces to *reverse-engineer the pipeline*.

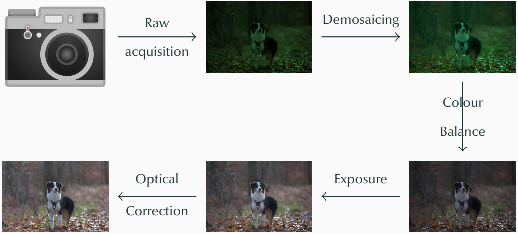
Image Processing Pipeline

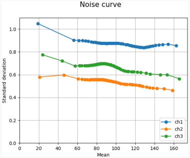




The image processing pipeline *leaves traces* on the image.
We *analyse* those traces to *reverse-engineer the pipeline*.

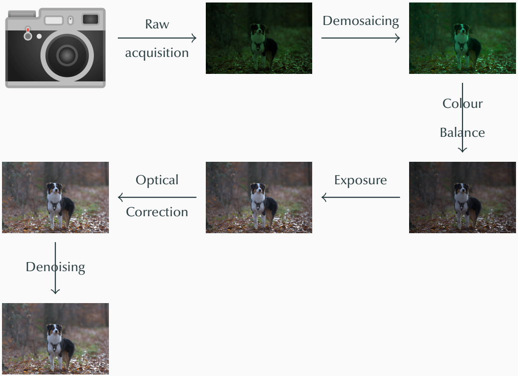
Image Processing Pipeline

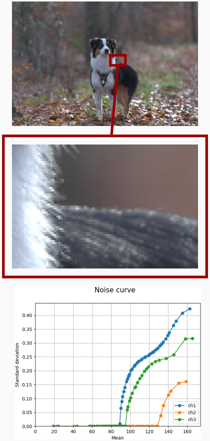




The image processing pipeline *leaves traces* on the image.
We *analyse* those traces to *reverse-engineer the pipeline*.

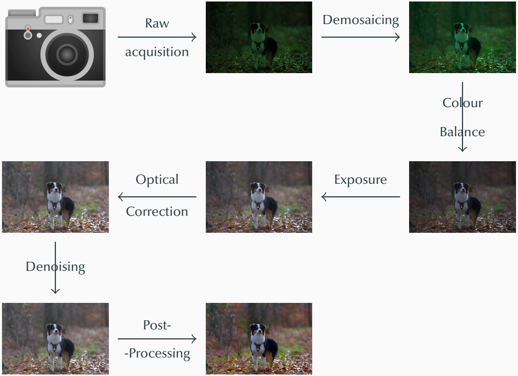
Image Processing Pipeline

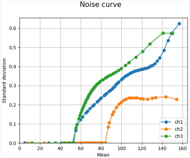




The image processing pipeline *leaves traces* on the image.
We *analyse* those traces to *reverse-engineer the pipeline*.

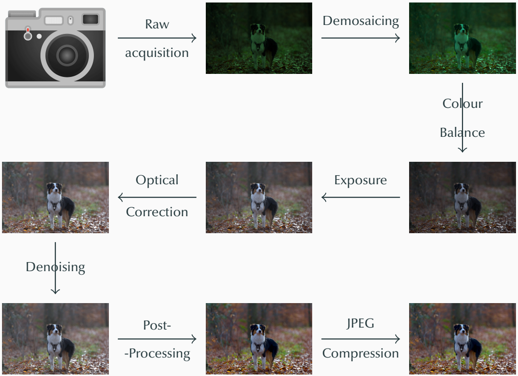
Image Processing Pipeline

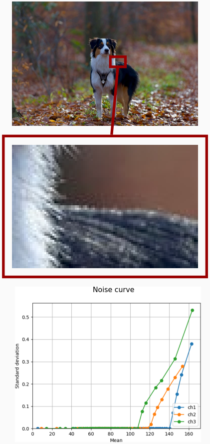




The image processing pipeline *leaves traces* on the image.
We *analyse* those traces to *reverse-engineer the pipeline*.

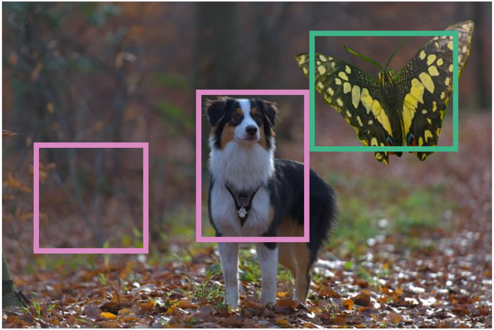
Image Processing Pipeline

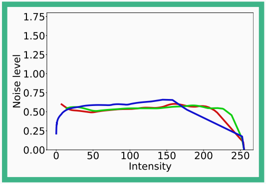




The image processing pipeline *leaves traces* on the image.
We *analyse* those traces to *reverse-engineer the pipeline*.

For a trace-based analysis of images





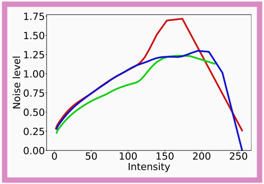


Image credit: Tina Nikoukhah

The 3 paradigms of forgery detection

Trace-specific analysis

- Noise level
- JPEG compression
- Demosaicing
- ...

Noise residual analysis

1. Extract the noise residual
2. Is it consistent across the image?

Forged/Authentic classification

1. Train a CNN on forgery datasets
2. Is the given image authentic or forged?

Datasets and evaluation

Let's make a new dataset to benchmark splicing detection!

1. Take an object from one image
2. Copy it to another image
3. Add some blur to the border

Nothing could go wrong with that...

Datasets and evaluation

Let's make a new dataset to benchmark splicing detection!

1. Take an object from one image
2. Copy it to another image
3. Add some blur to the border

Nothing could go wrong with that...Could it?

Training on this dataset, my classification network reaches 99% accuracy on the testing set!

The questions to ask

Trace-specific analysis

- Where can we evaluate those methods?
- How do methods of the same category compare?

Noise residual analysis

1. To what kind of traces are they sensitive?
2. What are their synergies?

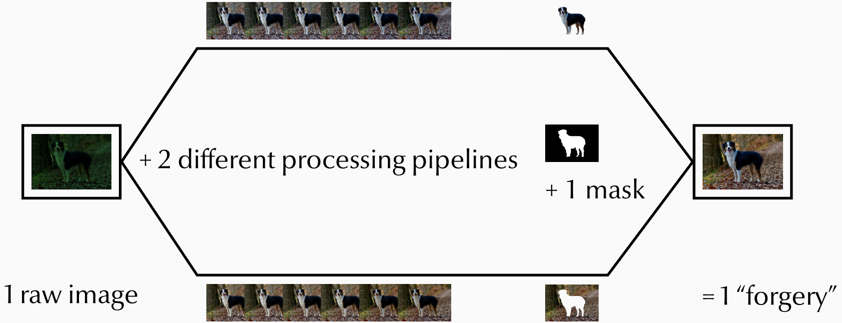
Forged/Authentic classification

1. Can they make detections on grounds other than semantics?
2. How to make sure they did not overfit to a specific method?

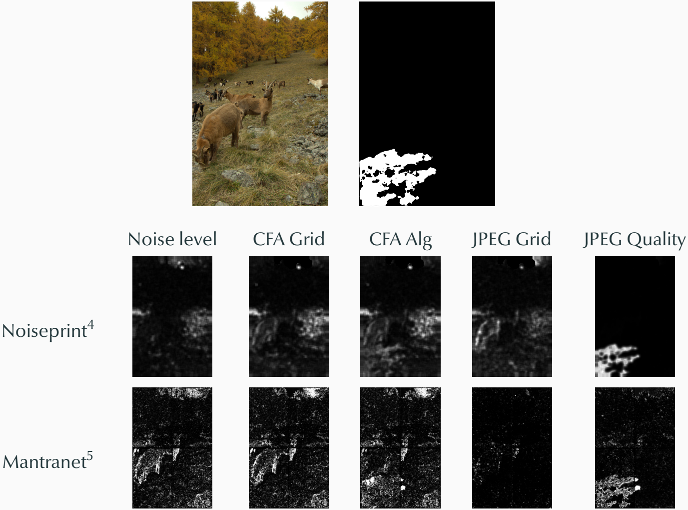
How to answer those questions?

Trace: forgeries without forgeries

Trace database: Forgeries without forgeries



Evaluation results



[4] D. Cozzolino and L. Verdoliva, "Noiseprint: A cnn-based camera model fingerprint," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 144–159, 2020. DOI: `10.1109/TIFS.2019.2916364`

[5] Y. Wu, W. AbdAlmageed, and P. Natarajan, "Mantra-net: Manipulation tracing network for detection and localization of image forgeries with anomalous features," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019

Demosaicing

Most camera can only sample one colour per pixel.

Figure 1: The camera does not see this...

Demosaicing

Most camera can only sample one colour per pixel.

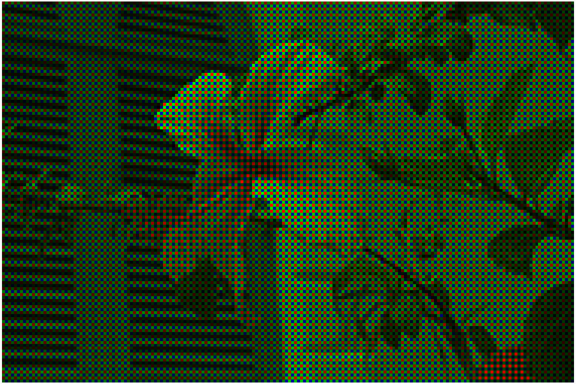


Figure 1: ...but this!

Demosaicing

Most camera can only sample one colour per pixel.

Figure 1: Demosaicing algorithms interpolate missing values.

Bayer Matrix

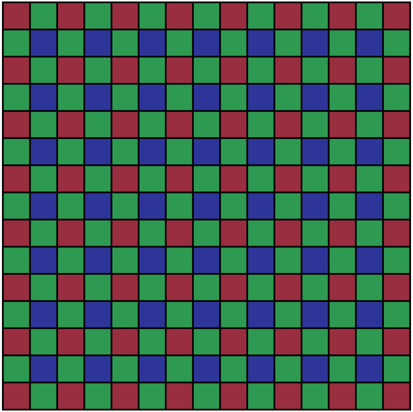


Figure 2: Bayer Matrix: in which colour each pixel is sampled

How to use demosaicing analysis to detect forgeries?

Check for the local absence of demosaicing traces!

- Popescu and Farid[6]: Probability map that each pixel has been sampled in green: should have a 2-periodic component.
- Le and Retraint[7]: Improve the estimation by using several filter and focusing on smoother areas.
- Ferrara, Bianchi, De Rosa, *et al.*[8]: Difference of variance after reinterpolation.

---

[6] A. C. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3948–3959, Oct. 2005, ISSN: 1053-587X. DOI: 10.1109/TSP.2005.855406

[7] N. Le and F. Retraint, "An improved algorithm for digital image authentication and forgery localization using demosaicing artifacts," *IEEE Access*, vol. 7, pp. 125 038–125 053, 2019. DOI: 10.1109/ACCESS.2019.2938467

[8] P. Ferrara, T. Bianchi, A. De Rosa, *et al.*, "Image forgery localization via fine-grained analysis of cfa artifacts," *IEEE TIFS*, vol. 7, no. 5, pp. 1566–1577, 2012

Local shift of the demosaicing pattern?

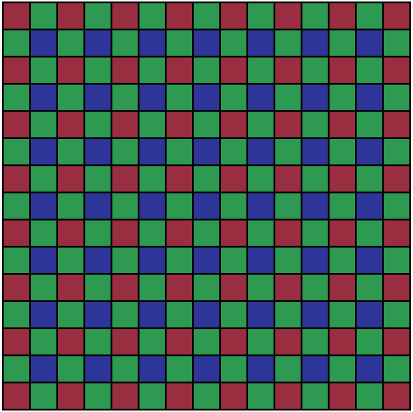


Figure 3: Bayer Matrix: in which colour each pixel is sampled

Local shift of the demosaicing pattern?

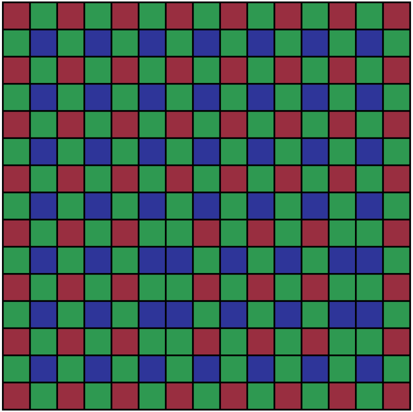
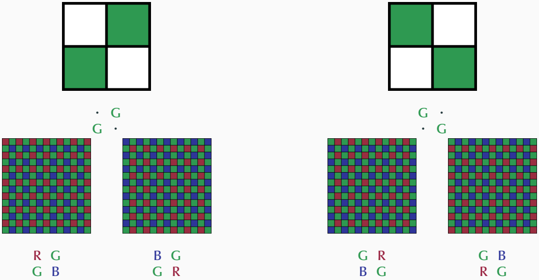


Figure 3: In case of a copy-paste, $\frac{3}{4}$ chances that the sampling colour of pixels won't be correctly arranged

Demosaicing patterns



Can we detect the offset?

Several methods exist to detect the demosaicing pattern:

- Kirchner[9]: Re-demosaick an image with the bilinear algorithm in the four possible configurations: Which one has the lowest residual?
- Choi et al.[10]: Originally sampled pixels are more likely to take extremal value: Count them.

[9] M. Kirchner, "Efficient estimation of CFA pattern configuration in digital camera images," in *Media Forensics and Security*, 2010

[10] C.-H. Choi, J.-H. Choi, and H.-K. Lee, "CFA pattern identification of digital cameras using intermediate value counting," in *MM&Sec*, ser. MM&Sec '11, Buffalo, New York, USA: ACM, 2011, pp. 21–26, ISBN: 978-1-4503-0806-9. DOI: 10.1145/2037252.2037258

Can we detect the offset?

139	240	154	16	94	56	72	20
92	131	168	76	72	94	24	43
85	24	100	48	102	224	130	72
60	107	160	68	64	122	200	153
92	184	125	0	50	0	133	108
52	155	156	76	136	117	224	127
146	228	111	12	110	108	107	44
56	114	48	90	184	141	52	90

Can we detect the offset?

Several methods exist to detect the demosaicing pattern:

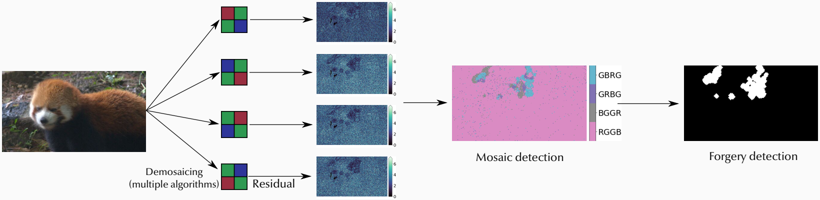
- Kirchner[9]: Re-demosaick an image with the bilinear algorithm in the four possible configurations: Which one has the lowest residual?
- Choi et al.[10]: Originally sampled pixels are more likely to take extremal value: Count them.

---

[9] M. Kirchner, "Efficient estimation of CFA pattern configuration in digital camera images," in *Media Forensics and Security*, 2010

[10] C.-H. Choi, J.-H. Choi, and H.-K. Lee, "CFA pattern identification of digital cameras using intermediate value counting," in *MM&Sec*, ser. MM&Sec '11, Buffalo, New York, USA: ACM, 2011, pp. 21–26, ISBN: 978-1-4503-0806-9. DOI: 10.1145/2037252.2037258

Can we detect the offset?



Can we detect the offset?

Several methods exist to detect the demosaicing pattern:

- Kirchner[9]: Re-demosaick an image with the bilinear algorithm in the four possible configurations: Which one has the lowest residual?
- Choi et al.[10]: Originally sampled pixels are more likely to take extremal value: Count them.

These methods can detect the pattern...on the global image.

[9] M. Kirchner, "Efficient estimation of CFA pattern configuration in digital camera images," in *Media Forensics and Security*, 2010

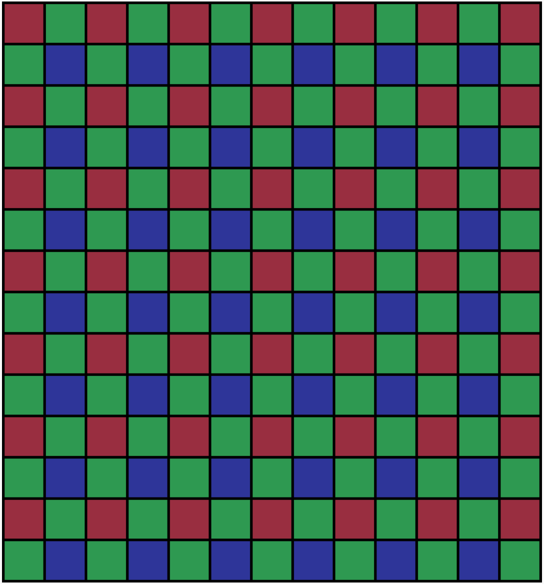
[10] C.-H. Choi, J.-H. Choi, and H.-K. Lee, "CFA pattern identification of digital cameras using intermediate value counting," in *MM&Sec*, ser. MM&Sec '11, Buffalo, New York, USA: ACM, 2011, pp. 21–26, ISBN: 978-1-4503-0806-9. DOI: 10.1145/2037252.2037258

What is wrong with current methods?

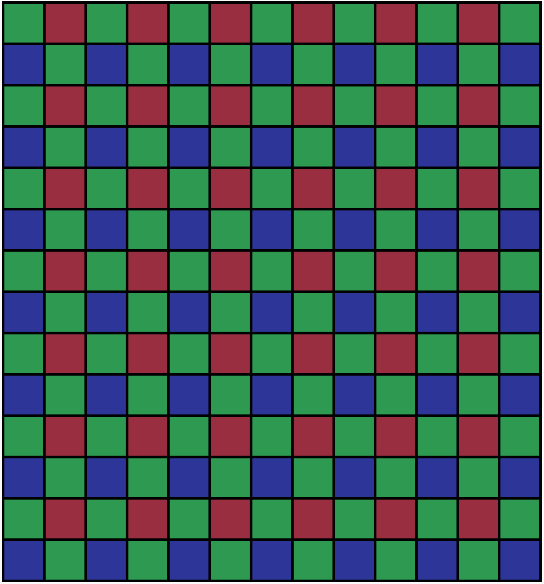
- Channel independence?
- Linear estimation?
- Too many scenarios?

We need a better method!

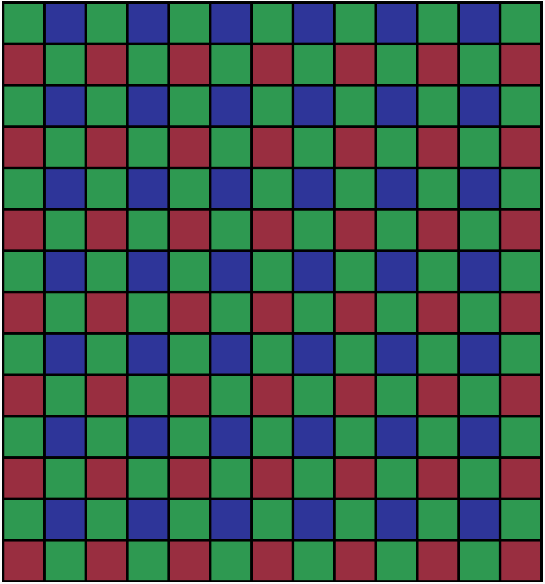
CNNs are translation-invariant!



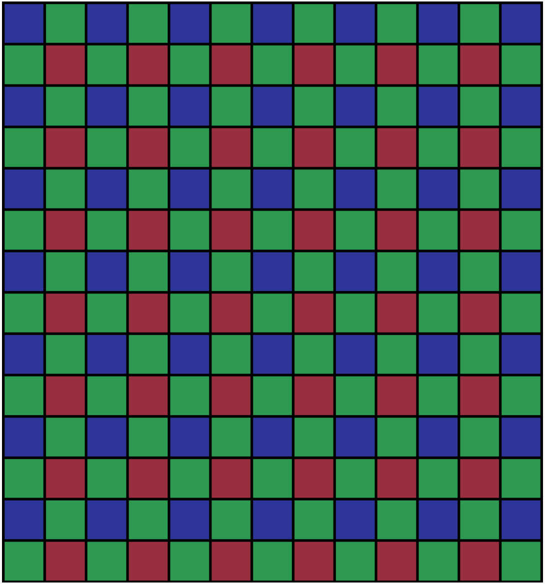
CNNs are translation-invariant!



CNNs are translation-invariant!



CNNs are translation-invariant!



Positional Learning

- Train to detect pixel-position-related information
- A kind of self-supervised training: labels drawn from the image itself

Positional learning: leverage the translation invariance

0	1	0	1	0	1		0	*	0	*	0	*
1	0	1	0	1	0		*	1	*	1	*	1
0	1	0	1	0	1		0	*	0	*	0	*
1	0	1	0	1	0		*	1	*	1	*	1
0	1	0	1	0	1		0	*	0	*	0	*
1	0	1	0	1	0		*	1	*	1	*	1

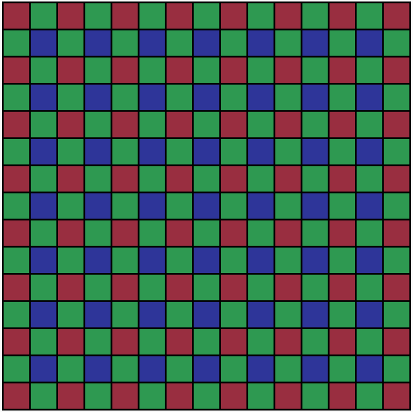
Train a CNN to output those patterns: How will it do so?

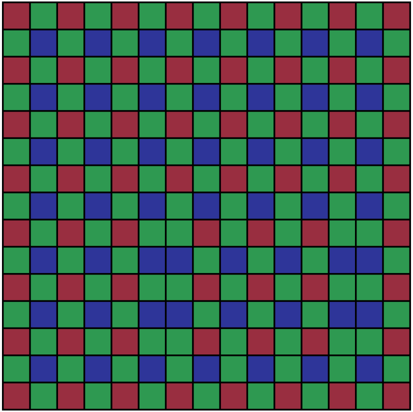
Positional learning: leverage the translation invariance

0	1	0	1	0	1
1	0	1	0	1	0
0	1	0	1	0	1
1	0	1	0	1	0
0	1	0	1	0	1
1	0	1	0	1	0

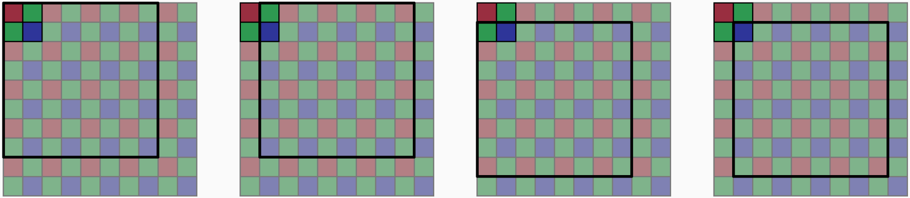
0	*	0	*	0	*
*	1	*	1	*	1
0	*	0	*	0	*
*	1	*	1	*	1
0	*	0	*	0	*
*	1	*	1	*	1

Train a CNN to output those patterns: How will it do so?**It has to rely on the underlying demosaicing traces.**

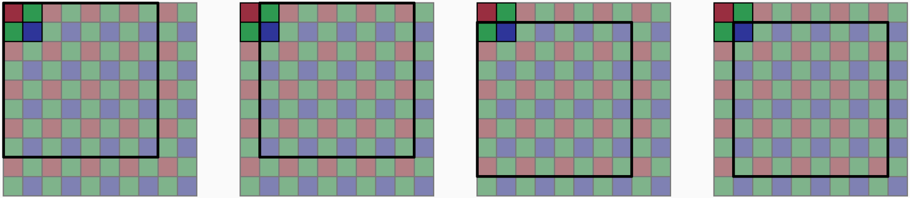




Training on JPEG-compressed images



Training on JPEG-compressed images



Helps build robustness to JPEG compression…but is not enough!

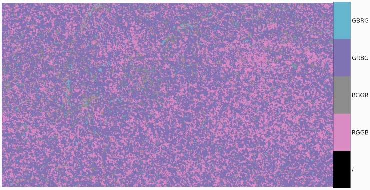
- Need even more robustness!
- What of unseen post-processing techniques?

Positional Internal Learning

- Positional learning:
  - Train to detect pixel-position-related information
  - A kind of self-supervised learning: labels are drawn from the image itself
- Internal learning:
  - Retrain a pretrained network directly on an image to analyse it
  - Possible because training is self-supervised: we can get labels from any images
  - Helps adapt the network to different scenarios, gives robustness to JPEG compression.

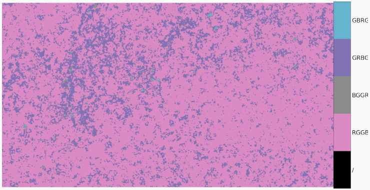
Internal learning

Let's take an authentic image and retrain the network on it



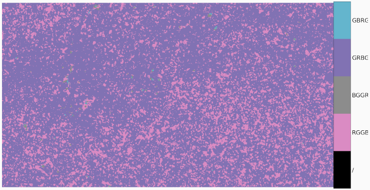
Internal learning

Let's take an authentic image and retrain the network on it



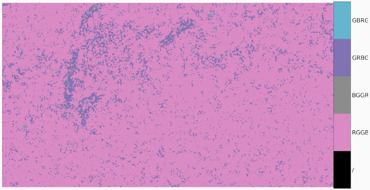
Internal learning

Let's take an authentic image and retrain the network on it



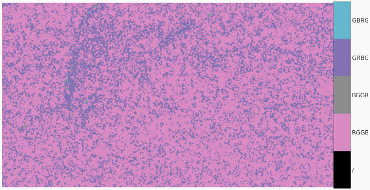
Internal learning

Let's take an authentic image and retrain the network on it



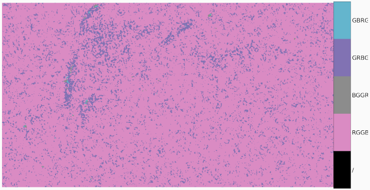
Internal learning

Let's take an authentic image and retrain the network on it



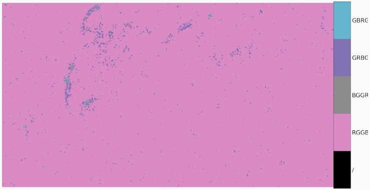
Internal learning

Let's take an authentic image and retrain the network on it



Internal learning

Let's take an authentic image and retrain the network on it



Internal learning

Let's take an authentic image and retrain the network on it

Internal learning

Let's take an authentic image and retrain the network on it

Internal learning

Let's take an authentic image and retrain the network on it

Internal learning

Let's take an authentic image and retrain the network on it

Internal learning

Let's take an authentic image and retrain the network on it

Internal learning

Let's take an authentic image and retrain the network on it

Internal learning

Let's take an authentic image and retrain the network on it

Internal learning

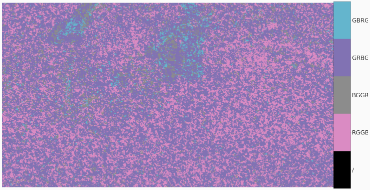
Let's take an authentic image and retrain the network on it

Internal learning

Let's take an authentic image and retrain the network on it

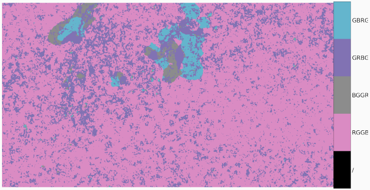
Toy example: forged image

Let's take a **forged** image and retrain the network on it



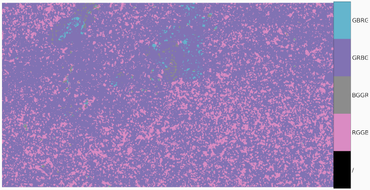
Toy example: forged image

Let's take a **forged** image and retrain the network on it



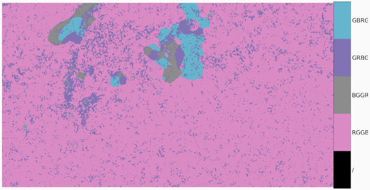
Toy example: forged image

Let's take a **forged** image and retrain the network on it



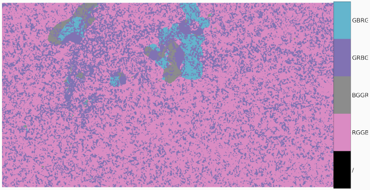
Toy example: forged image

Let's take a **forged** image and retrain the network on it



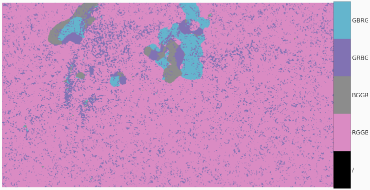
Toy example: forged image

Let's take a **forged** image and retrain the network on it



Toy example: forged image

Let's take a **forged** image and retrain the network on it



Toy example: forged image

Let's take a **forged** image and retrain the network on it

Toy example: forged image

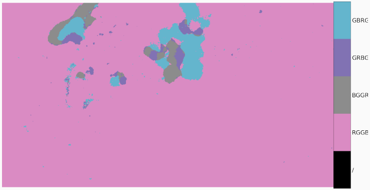
Let's take a **forged** image and retrain the network on it

Toy example: forged image

Let's take a **forged** image and retrain the network on it

Toy example: forged image

Let's take a **forged** image and retrain the network on it



Toy example: forged image

Let's take a **forged** image and retrain the network on it

Toy example: forged image

Let's take a **forged** image and retrain the network on it

Toy example: forged image

Let's take a **forged** image and retrain the network on it

Toy example: forged image

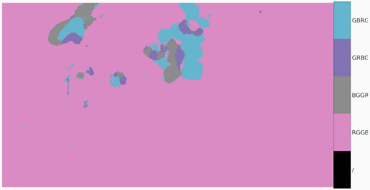
Let's take a **forged** image and retrain the network on it

Toy example: forged image

Let's take a **forged** image and retrain the network on it

Toy example: forged image

Let's take a **forged** image and retrain the network on it



Is it even possible?

Difficult

- The forged region contributes to overfitting: the network will train to detect forged pixels' positions correctly even though it should not
- JPEG compression can induce a lot of overfitting, because JPEG artefacts are also cues to pixels' positions.

But possible

- Forged regions are usually small, and their cues contradict the rest of the image: the network shouldn't learn too much on it
- The network is initialised to learn from CFA: At first it will be cheaper for it to adapt CFA interpretation than to learn JPEG artefacts from scratch.

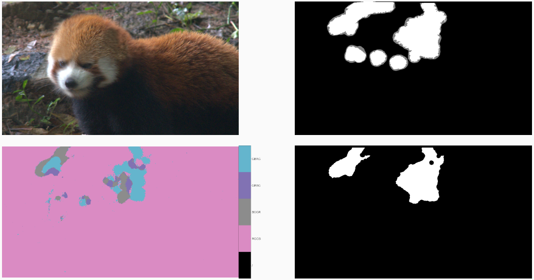
How to detect demosaicing pattern shifts?

- The network only detects positional information, it does not directly see forgeries.
- However, a shift in the demosaicing pattern causes a shift in the detection.
- The demosaicing pattern can be derived from the detection's output

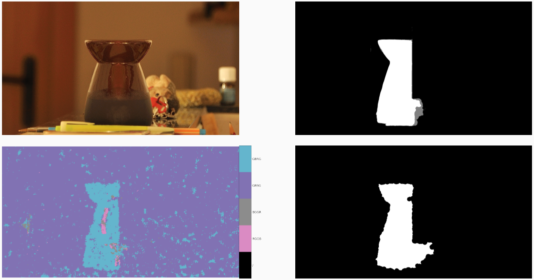
A contrario detection

- Aggregate results in 2×2 block votes. Probability p_0 that a block does not vote for the globally-correct grid.
- Is a region significantly wrong?
- If an image is authentic, probability that a given region has at least n out of k votes for a wrong grid: $\text{Binom}_{sf}(k, n, p_0)$.
- *A contrario* detection: threshold on the expected **Number of False Alarms** (NFA) that would be expected on the image: $\text{NFA}_{k,n,p_0} = n_{\text{windows}} \text{Binom}_{sf}(k, n, p_0)$.

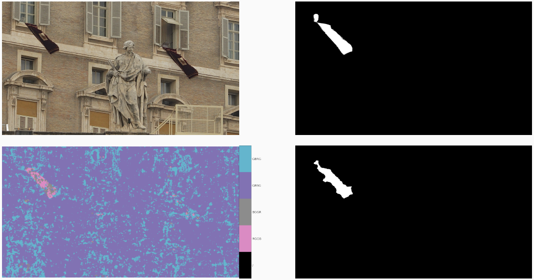
Results



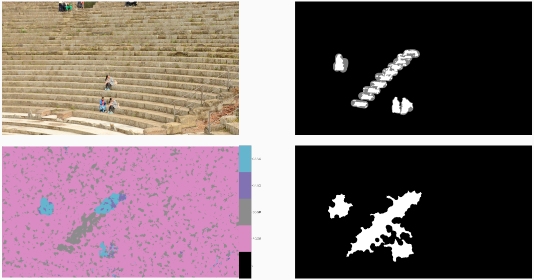
Results



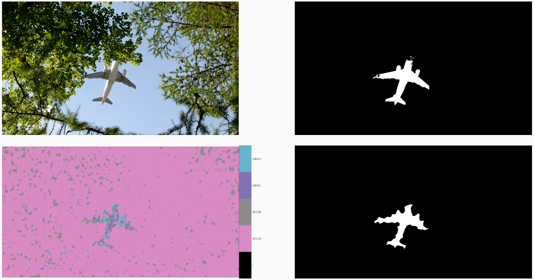
Results



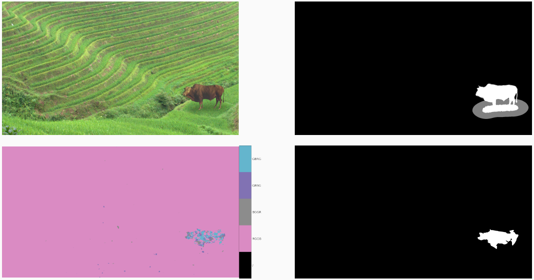
Results



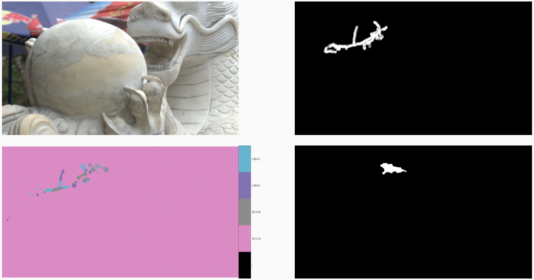
Results



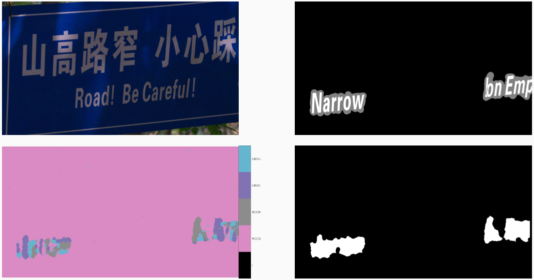
Results



Results



Results



Conclusion

- **Trace methodology and database**: A new methodology to evaluate, understand and improve forensic tools
- **Positional learning**: Leverage the translation-invariance of CNN to replicate an underlying 2-periodic component
- Coupled with **Internal learning** and ***A contrario* validation**, adapt to each image to improve detections over wider variety of scenarios.

Impact

- Integration of a demosaicing analysis method in the Envisu4 project for use by fact-checkers: `https://weverify.eu/verification-plugin/`
- Participation to the Defals (ANR/DGA) challenge for forgery detection
- Main publications:
    - Q. **Bammey**, R. G. v. Gioi, and J.-M. Morel, "An adaptive neural network for unsupervised mosaic consistency analysis in image forensics," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020
    - Q. **Bammey**, R. Grompone von Gioi, and J.-M. Morel, "Forgery detection by internal positional learning of demosaicing traces," Workshops on Applications of Computer Vision 2022 (WACV 2022), journal article in preparation
    - Q. **Bammey**, T. Nikoukhah, M. Gardella, *et al.*, "Non-semantic evaluation of image forensics tools: Methodology and database," Workshops on Applications of Computer Vision 2022 (WACV 2022)

Conclusion

- Use cases and limits of demosaicing analysis?

Conclusion

- Use cases and limits of demosaicing analysis?
- Main limits: JPEG compression (can now withstand some, but not too much), resampling

Conclusion

- Use cases and limits of demosaicing analysis?
- Main limits: JPEG compression (can now withstand some, but not too much), resampling
- Trace methodology: could it be used to design new methods?

Conclusion

- Use cases and limits of demosaicing analysis?
- Main limits: JPEG compression (can now withstand some, but not too much), resampling
- Trace methodology: could it be used to design new methods?
- Positional learning: Introduced to analyse a periodic component...what else can we do with it?