Can printable unclonable codes be copied? Evaluating the performance of attacks and highlighting the role of the detector

Iuliia Tkachenko

GDR SI - June 26, 2023

LIRIS, Université Lumière Lyon 2, CNRS, Lyon, France

Motivation

Counterfeiting risks:

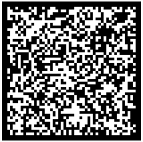
- Health and safety risks
- Loss of consumer trust
- Loss of market share and liability risks
- Damage to the brand reputation

Printable unclonable codes:

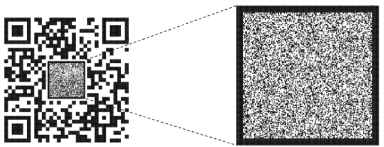
- Integrates easily to existing printing processes
- Leverages information loss during the printing process
- Robust to naïve attacks and sensitive to estimation attack using neural networks

Copy Detection Pattern

CDP (Copy Detection Pattern) is a small random or pseudo-random digital image which is printed at the native printer resolution and designed to maximize information loss during printing or copying.



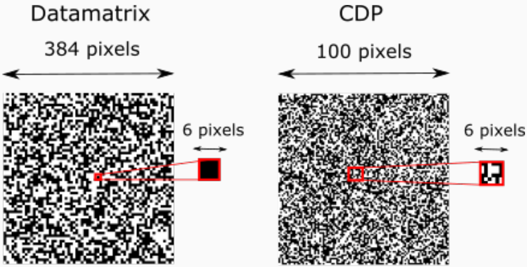
[Picard 2004]



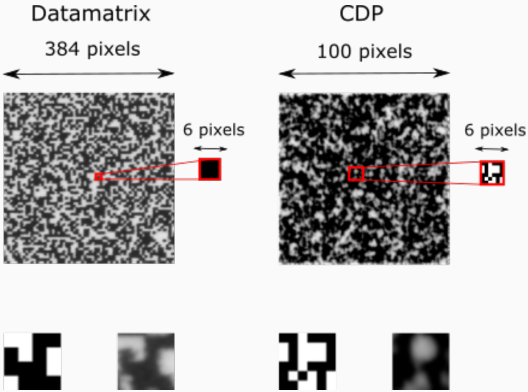
[Picard et al. 2021]

J. Picard, “Digital authentication with copy-detection patterns,” Electronic Imaging 2004, pp. 176–183.
J. Picard, P. Landry and M.Bolay, "Counterfeit detection with QR codes”,” Proceedings of the 21st ACM Symposium on Document Engineering.

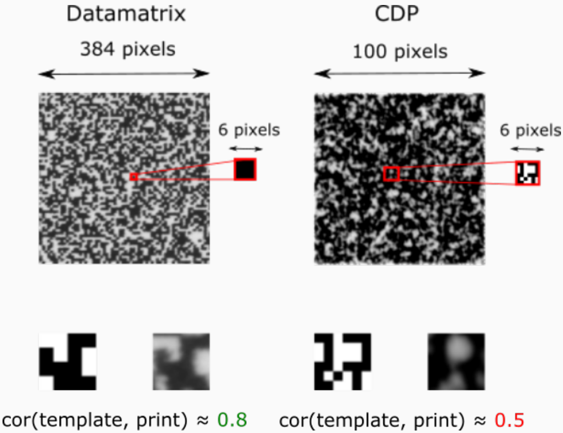
Datamatrix vs CDP: digital



Datamatrix vs CDP: printed



Datamatrix vs CDP: information loss



Authentication process

Original template	I
Original hardcopy	$\Pi(I)$
Original after P&S	$\Sigma(\Pi(I))$
Estimated softcopy	$\hat{I}$
Fake after P&S	$\Sigma(\Pi'(\hat{I}))$

Authentication process

Original template	I
Original hardcopy	$\Pi(I)$
Original after P&S	$\Sigma(\Pi(I))$
Estimated softcopy	$\hat{I}$
Fake after P&S	$\Sigma(\Pi'(\hat{I}))$

<u>Authentication test</u>

$$\mathcal{H}_0 : \tilde{I} \sim \Sigma(\Pi(I)),$$

$$\mathcal{H}_1 : \tilde{I} \nsim \Sigma(\Pi(I)),$$

where $\tilde{I}$ is a grayscale image of CDP to authenticate.
Comparison metrics: distance or correlation coefficient between digital template and P&S CDP.

Detector improvements

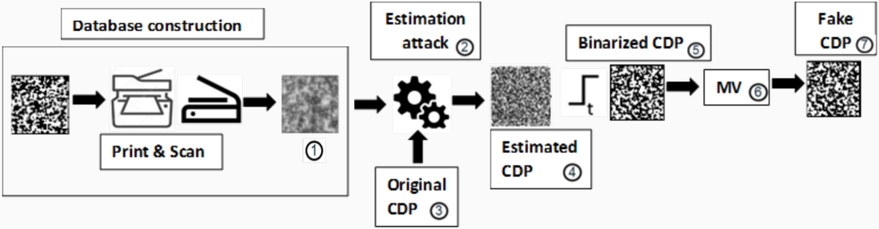
Better separability during authentication test can be ensured by:

- **template resizing**: scaling the template by an integer factor allows a more precise sub-pixel matching.
- **template matching**: allows to take into account sub-pixel geometric distortion following the P&S process.

  It consists in matching the slightly cropped template with the grayscale scan by maximizing the correlation score.
- **high pass filtering** (such as unsharp masking) before the correlation score calculation reduces the low frequencies which are less discriminative.

E. Khermaza, I. Tkachenko, J. Picard, "Can Copy Detection Patterns be copied? Evaluating the performance of attacks and highlighting the role of the detector", IEEE WIFS 2021, December 2021, Montpellier, France

Estimation attack



Possible estimations:

- Image processing attack
- Attack based on estimation using neural networks
- Averaging attack for batch CDP

I. Tkachenko, C. Destruel, “Exploitation of redundancy for pattern estimation of copy-sensitive two level QR code”, IEEE WIFS 2018, December 2018, Hong Kong, China.

O. Taran, S. Bonev, and S. Voloshynovskiy, "Clonability of anti- counterfeiting printable graphical codes: a machine learning approach", IEEE ICASSP, May 2019, Brighton, United Kingdom.

R. Yadav, I. Tkachenko, A. Trémeau, T. Fournel, “Copy Sensitive Graphical Code Estimation: Physical vs Numerical Resolution”, IEEE WIFS 2019, December 2019, Delft, Netherlands.

Possible estimations

- Image processing attacks [Tkachenko et al.]:
  - Otsu thresholding
  - pre-preprocessing operation using unsharp mask

I. Tkachenko, C. Destruel, "Exploitation of redundancy for pattern estimation of copy-sensitive two level QR code", IEEE WIFS 2018, December 2018, Hong Kong, China.

I. Tkachenko, C. Destruel, O. Strauss, W. Puech, "Sensitivity of different correlation measures to print-and-scan process", Electronic Imaging 2017, February 2017, Burlingame, USA.

Possible estimations

- Image processing attack [Tkachenko et al.]:
    - Otsu thresholding
    - pre-preprocessing operation using unsharp mask
- Attack based on estimation using neural networks:
    - bottleneck DNN with 2 fully connected hidden layers [Taran et al. 2019]
    - Selectional Auto-Encoder [Yadav et al. 2019b1]
    - Super Resolution Generative Adversarial Networks [Yadav et al. 2019b2]

O. Taran, S. Bonev, and S. Voloshynovskiy, "Clonability of anti-counterfeiting printable graphical codes: a machine learning approach", IEEE ICASSP, May 2019, Brighton, United Kingdom.

R. Yadav, I. Tkachenko, A. Trémeau, T. Fournel, "Estimation of copy-sensitive codes using a neuronal approach", IH&MMSec 2019, July 2019, Paris, France.

R. Yadav, I. Tkachenko, A. Trémeau, T. Fournel, "Copy Sensitive Graphical Code Estimation: Physical vs Numerical Resolution", IEEE WIFS 2019, December 2019, Delft, Netherlands.

Possible estimations

- Image processing attack
- Attack based on estimation using neural networks
- Averaging attack for batch CDP [Tkachenko et al 2018]:
    1. CDP C is printed and scanned m times that gives us $P_j, j = 1, \cdots, m$ samples for estimation attack.
    2. Samples $P_j, j = 1, \cdots, m$ are binarized using either image processing or neural networks.
    3. Averaging step consists in counting the number of black and white pixels for each position: if the majority of binarized batch samples have a white pixel in this position, the pixel on the estimated code will also be white, otherwise it will be black.

I. Tkachenko and C. Destruel, "Exploitation of redundancy for pattern estimation of copy-sensitive two level QR code", IEEE WIFS 2018, December 2018, Hong Kong, China.

Public datasets

- "Unrealistic" datasets - elementary unit size 5x5 pixels per module
    - DP0E, DP1E, DP1C [Taran et al. 2019]
- "Realistic" datasets - elementary unit size 1 pixel per module
    - CSGC [Yadav et al. 2019]
    - Indigo 1x1 [Chaban et al. 2021]
    - Copy Detection Pattern Dataset [Khermaza et al. 2021]

O. Taran, S. Bonev, and S. Voloshynovskiy, "Clonability of anti-counterfeiting printable graphical codes: a machine learning approach", IEEE ICASSP, May 2019, Brighton, United Kingdom.

R. Yadav, I. Tkachenko, A. Trémeau, T. Fournel, "Estimation of copy-sensitive codes using a neuronal approach", IH&MMSec 2019, July 2019, Paris, France.

R. Chaban, O. Taran, J. Tutt, T. Holotyak, S. Bonev, and S. Voloshynovskiy, "Machine learning attack on copy detection patterns: are 1x1 patterns cloneable?", IEEE WIFS 2021, December 2021, Montpellier, France.

E. Khermaza, I. Tkachenko, J. Picard, "Can Copy Detection Patterns be copied? Evaluating the performance of attacks and highlighting the role of the detector", IEEE WIFS 2021, December 2021, Montpellier, France.

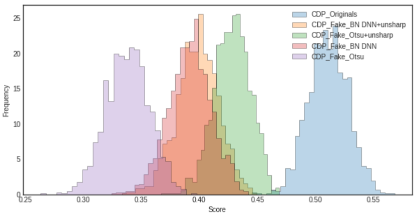
Estimation attack with realistic dataset

- 2500 / 1000 images for training and validation
- 1500 for testing
- optimal unsharp parameters: radius - 2.875, amount - 10.

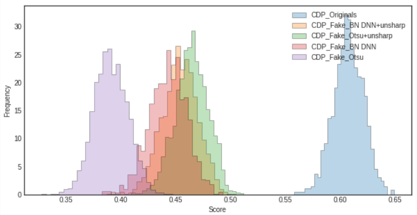
	Mean BER	Min BER	Max BER
Image processing approach			
Otsu	33.60%	29.86%	38.40%
Otsu+unsharp	**23.37%**	**19.92%**	**27.08%**
Neural network approach			
FC2	28.06%	25.68%	31.06%
FC3	26.95%	24.33%	30.26%
FC4	24.68%	21.13%	28.64%
BN DNN	**23.27%**	**20.31%**	**26.99%**

E. Khermaza, I. Tkachenko, J. Picard, "Can Copy Detection Patterns be copied? Evaluating the performance of attacks and highlighting the role of the detector", IEEE WIFS 2021, December 2021, Montpellier, France.

Correlation score with realistic dataset



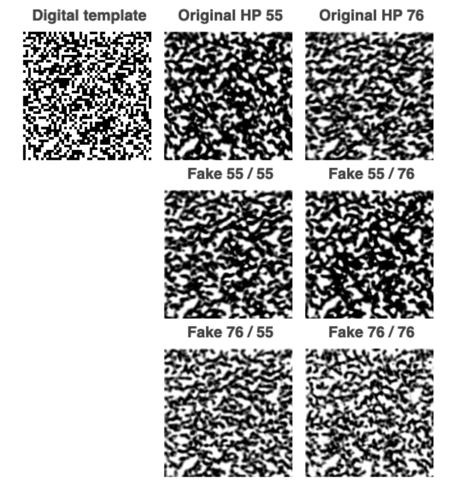
(a) Without detector improvements



(b) Template matching & unsharp masking

E. Khermaza, I. Tkachenko, J. Picard, "Can Copy Detection Patterns be copied? Evaluating the performance of attacks and highlighting the role of the detector", IEEE WIFS 2021, December 2021, Montpellier, France.

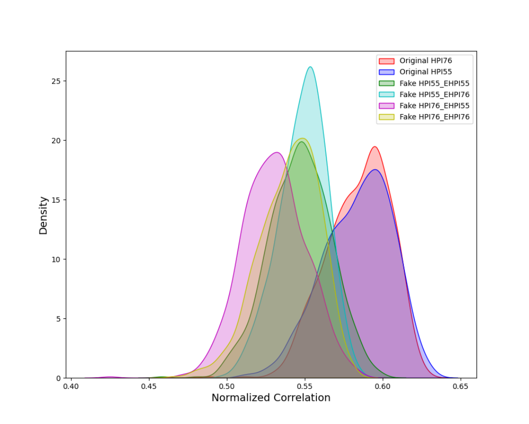
Indigo dataset



R. Chaban, O. Taran, J. Tutt, T. Holotyak, S. Bonev, and S. Voloshynovskiy, "Machine learning attack on copy detection patterns: are 1x1 patterns cloneable?", IEEE WIFS 2021, December 2021, Montpellier, France.

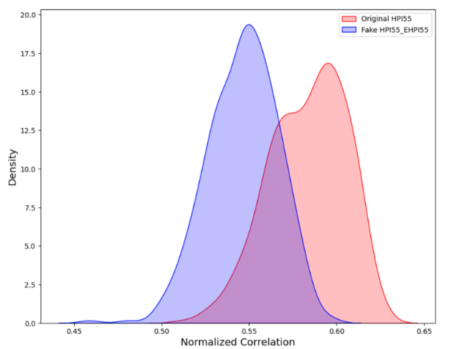
Correlation score with Indigo dataset

- 2 HP Indigo printers
- 4 fake types using estimation attack



Real world use case

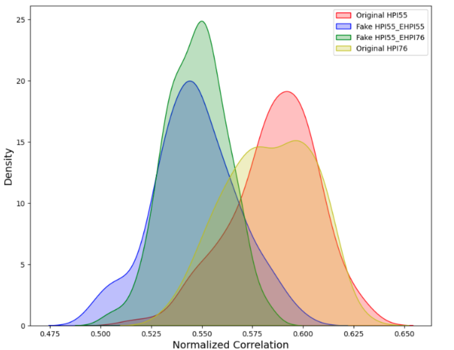
- 1 HP Indigo printer
- 1 fake type using known estimation attack



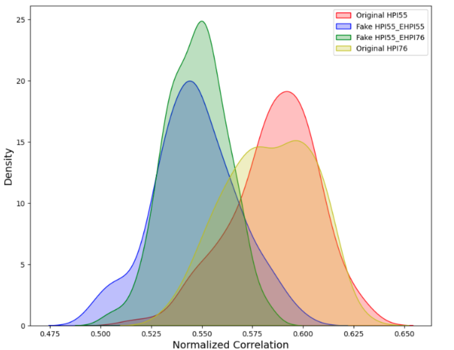
Current challenge

The CDPs come to the detector from:

- Known HP Indigo printer - Originals
- Known fakes used to train the detector
- Unknown fakes from estimation attack
- Unknown HP Indigo printer - Considered as fakes



Detector based on correlation values

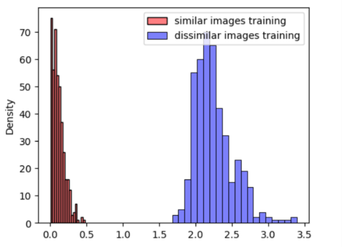


True labels	Predicted labels	
	Original	Fake
HP55	82%	18%
F 55/55	17%	83%
F 55/76	12%	88%
HP76	76%	24%

Rethink a detector?

Ongoing work

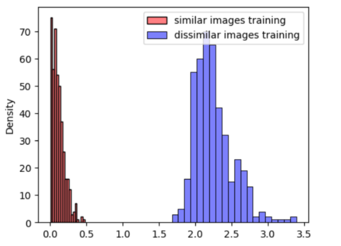
- Use of metric learning
- Try to identify the printer used



Rethink a detector?

Ongoing work

- Use of metric learning
- Try to identify the printer used



True labels	Predicted labels	
	Original	Fake
HP55	77%	23%
F 55/55	45%	55%
F 55/76	37%	63%
HP76	0%	100%

Conclusions & perspectives

Conclusions:

- Correlation based detectors are not (any more) relevant for fake CDP detection.
- The pre-processing strategies can increase the separability between the original and the fake CDPs.
  - These techniques are not sufficient to guarantee full separation while the estimation process is accurate.
- The detector that consider printer used could be a good future solution to consider.

Perspectives:

- Improve the detector accuracy by using another similarity metrics.
- Test the detectors in the real world use-cases - using smartphones

Questions ?

Contact : iuliia.tkachenko@liris.cnrs.fr