Internship proposal

Internship subject

Metric learning for forgery detection

Keywords:

Similarity metric learning, image processing, image features, printable unclonable codes

Context of the study

The nowadays challenges are the fast, reliable, and cheap detection of faked packaging and documents. Due to development and broad availability of high-quality printing and scanning devices, the number of forged or counterfeited products and documents is dramatically increasing. Therefore, different security elements have been suggested to prevent this socio-economic plague. One of the most promising and cheap solutions is the use of Copy Detection Patterns (CDP) [1]. A CDP is a maximum entropy image, generated using a secret key or password, that takes full advantage of information loss principle during printing-and-scanning process. Such an unpredictable pattern is highly sensitive to distortions occurring inevitably during production (printing), verification (scanning) and reproduction (duplication) processes. Counterfeiting a printed CDP requires scanning then re-printing that increases image degradations. Initially, the detection of counterfeited CDP was devoted to evaluating the level of information loss.

However, the security of CDP based authentication system was shown to be vulnerable to neural network estimation attacks [3,4] which attempted to infer the CDP after scanning and before re-printing to fool the detector.

The project *FakeNets: morphological analysis for fake generation and similarity metrics for fake detection by neural networks* aims to explore the potential offered by deep learning methods in the context of CDP secure printing, from the attacker's point of view at first, and from the verifier's point of view, secondly. During this internship we will work on the second task: to increase the performance of counterfeit CDP detection using similarity metric learning approaches [5].

Description of the subject

We plan to study similarity metric learning to improve the fake detector accuracy. It was shown earlier that estimated CDP can be spotted using the detector based on GAN [3]. Nevertheless, these detectors need a large database of original and fake CDP samples for training. In practice, it is not realistic to construct a database with all types of fakes, therefore, in real world applications, the detector often compares the original CDPs with a printed one and decide whether it is an original or a fake. There exist several improvements for the baseline detector using image pre-processing techniques [2]. However, these pre-processing techniques are quite sensitive to the quality of fakes. Recently, Siamese neural networks were used in forensics [6] for identification if the printing source is the same and have shown its usability in real-world scenarios.

Therefore, we would like to explore the possibility to use deep similarity metric learning for CDP detector to separate between the digital CDP, printed CDP (original) and fake CDP (estimated by neural networks [3]). Ultimately, we want to verify whether contrastive and angular losses can improve the detector results against estimation attacks.

Required profile

- The candidate must currently be enrolled in a Master 2 program or in the final year of engineering school (that corresponds to Bac+5 in France) in Computer Science.
- Programming languages: Python.
- Libraries for image analysis and processing: OpenCV, scikit-image (Python).
- Machine learning frameworks: scikit-learn, Pytorch.
- Scientific knowledge: signal processing, image analysis, machine learning and deep learning.
- Knowledge in multimedia security will be considered a plus.
- Languages: French or English.

Place and allowance of internship

LIRIS (Laboratoire d'Informatique en Image et Systèmes d'information) laboratory, campus of Université Lumière Lyon 2, Bron. Internship allowance is about 550 euros/month (3.90 euros per hour).

References

[1] J. Picard. Digital authentication with copy-detection patterns. In Electronic Imaging 2004, pages 176–183. International Society for Optics and Photonics, 2004.

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

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

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

[5] S. Duffner, C. Garcia, K. Idrissi, A. Baskurt. Similarity Metric Learning. Multi- faceted Deep Learning - Models and Data, 2021.

[6] A. Ferreira, N. Purnekar and M. Barni, "Ensembling Shallow Siamese Neural Network Architectures for Printed Documents Verification in Data-Scarcity Scenarios," in IEEE Access, vol. 9, pp. 133924-133939, 2021

Contact information

E-mail: iuliia.tkachenko@liris.cnrs.fr and Carlos.CrispimJunior@liris.cnrs.fr

Please provide your CV, the motivation letter and the transcripts with your marks for the two years of Master's degree / the last two years of engineering school.