

IMPACT DAD :

Utilisation de réseaux de neurones profonds pour la détection de manipulations d'images

Patrick Bas – CNRS/Ecole Centrale Lille
Patrick.Bas@ec-lille.fr
John Klein – Université de Lille
John.Klein@univ-lille1.fr

La détection de manipulations d'images¹ repose sur l'extraction de caractéristiques pertinentes (c'est à dire permettant la discrimination entre une image naturelle et une image manipulée). Ce problème peut être vu comme un problème d'extraction supervisée de caractéristiques (l'algorithme utilisé pouvant apprendre par l'exemple les caractéristiques utiles à la tâche). De par la grande dimensionnalité des images, ce problème était jusqu'à présent intraitable mais l'arrivée des réseaux de neurones convolutifs profonds [4, 3] et des implémentations GPU associées [2] a permis de construire des réseaux de neurones et d'extraire des caractéristiques ayant de très bonnes performances en détection de manipulations simples [1].

Durant cet impact, l'élève devra se familiariser dans un premier temps avec le principe des systèmes d'apprentissage par réseaux de neurones profonds (et plus particulièrement des réseaux convolutifs), puis ensuite d'utiliser les bibliothèques existantes telles que Caffe [2] afin de procéder à l'entraînement et au test du réseau pour reproduire les résultats publiés dans [1]. Ces travaux sont prometteurs puisque la première couche du réseau est guidée pour une tâche de forensics, ce qui permet d'obtenir de très bonnes performances (de l'ordre de 99% de taux de bonne classification). Dans un second temps, l'élève devra évaluer les limites du système construit.

L'élève aura à sa disposition l'accès à une carte GPU puissante (sur un serveur de l'école centrale ou au centre de calcul de Lille-1) ainsi que des bases d'apprentissage et de tests (200 000 images de taille 256x256 de chaque classe).

Connaissances et thèmes abordés : apprentissage statistique - réseaux de neurones profonds - traitement du signal

Capacités en programmation : Python - Caffe (<http://caffe.berkeleyvision.org>)

Références

- [1] Belhassen Bayar and Matthew C Stamm. A deep learning approach to universal image manipulation detection using a new convolutional layer. In *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*, pages 5–10. ACM, 2016.
- [2] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe : Convolutional architecture for fast feature embedding. *arXiv preprint arXiv :1408.5093*, 2014.
- [3] Wikipedia. https://en.wikipedia.org/wiki/convolutional_neural_network.
- [4] Wikipedia. https://en.wikipedia.org/wiki/deep_learning.

1. aussi appelé « image forensics ».