

# Détecteurs de manipulation d'images par réseaux de neurones profonds : utilisation de l'adaptation de domaine pour faire face à la diversité des images

Patrick Bas, Jérémie Boulanger, Vincent Itier

## — Objectif et contexte :

L'objectif de ce projet sera de construire un classifieur fonctionnant sur une base d'images différente de celle sur laquelle l'entraînement a été réalisé.

Le but de ce classifieur sera de détecter des manipulations d'images (manipulation par photoshop, traitements basiques, création de deepfake, ...). Les méthodes actuelles utilisent pour cela des réseaux de neurones profonds spécifiques [1] où la première couche est conçue de manière à permettre l'extraction de signaux faibles relatifs à l'image.

Une base d'images se caractérise par la chaîne de développement utilisée pour le passage de l'image brute (en sortie des capteurs de l'appareil photo) jusqu'au format JPEG.

## — Problème :

Etant donné que la création d'images manipulées est extrêmement couteuse et qu'il est très difficile de construire des images manipulées à souhait pour toutes les bases d'images possibles, on souhaite utiliser un classifieur entraîné pour une base donnée sur une autre base. Malheureusement, la transposition directe du classifieur donnera des mauvais résultats car la chaîne de développement comprend une succession d'opérations propre à chaque base qui laissent une trace statistique spécifique et provoque ce manque de généralisation du classifieur.

## — Solution proposée :

L'idée est alors d'utiliser des techniques de transfert de domaine pour les réseaux profonds [2]. Après avoir réalisé une étude bibliographique, l'étudiant aura pour tâche d'implémenter la structure proposée dans [1, 3], puis ensuite de mettre en oeuvre des méthodes de transfert. Des tests seront effectués sur des bases d'images manipulées artificiellement. L'implémentation s'effectuera sous TensorFlow ou Pytorch et un serveur muni d'un GPU sera mis à disposition.

## Contacts :

patrick.bas@centralelille.fr (Chercheur à CRIStAL, Bâtiment ESPRIT)

jeremie.boulanger@univ-lille.fr(Enseignant/Chercheur à CRIStAL, Bâtiment ESPRIT)

vincent.itier@imt-lille-douai.fr(Enseignant/Chercheur à CRIStAL, Bâtiment ESPRIT)

## Références

- [1] Ivan Castillo Camacho and Kai Wang. A simple and effective initialization of cnn for forensics of image processing operations. 2019. (document)
- [2] Mingsheng Long, Yue Cao, Jianmin Wang, and Michael I Jordan. Learning transferable features with deep adaptation networks. *arXiv preprint arXiv :1502.02791*, 2015. (document)
- [3] Jindong Wang. A pytorch implementation of 'learning transferable features with deep adaptation networks'. <https://github.com/jindongwang/transferlearning/tree/master/code/deep/DAN>. (document)