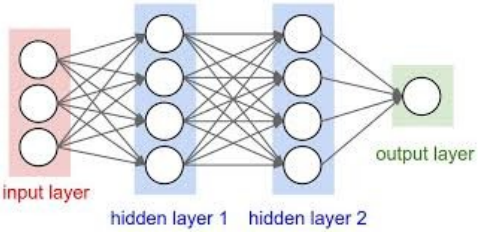

Stage : "Etude d'un générateur d'images BiGAN, pour une utilisation en stéganalyse"

Marc CHAUMONT, Mehdi YEDROUDJ, Frederic COMBY
LIRMM (Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier)
Equipe ICAR, 161 rue Ada, 34392 Montpellier cedex 5 - France
Tel : +33 4.67.14.97.59, Marc.Chaumont@lirmm.fr

Mots clefs : Traitement d'images, Stéganographie, Stéganalyse, Machine Learning, Deep Learning, GAN.

Figure 1 : **x**, des images "réelles", et *G*(*E*(**x**)), des images "générées" par un réseau BiGAN [Donahue2017 - BiGAN] (ayant appris sur ImageNet). Figure extraite de [Donahue2017 - BiGAN].

La stéganographie / stéganalyse peut être expliquée comme un jeu à trois participants. Les stéganographes classiquement appelés Alice et Bob, souhaitent envoyer un message, dont l'existence même n'est connu que d'eux seul. La stéganalyste, généralement appelé Eve, observe les échanges d'images qui ont lieu entre Alice et Bob et cherche à déterminer si Alice et Bob communiquent [Simmons83]. La stéganographie est donc l'art de dissimuler un message dans un support pour le transmettre de manière secrète, et la stéganalyse est l'art de déceler la présence de ce message. Cette discipline dans sa version moderne, c'est-à-dire numérique, a débuté au début des années 2000.

Il y a peu, l'état de l'art en stéganalyse d'image consistait à extraire d'une image un grand vecteur composé de valeurs réelles caractérisant l'image [Fridrich2012 - SRM_SRMQ1, Holub2013 - PSRM, ...], puis de donner ce vecteur à un classifieur qui décidait si l'image contenait ou non un message. Le classifieur le plus utilisé de l'état-de-l'art etait l'ensemble classifieur [Kodovsky2012 - EC].

En 2017, cette méthodologie "traditionnelle" est devenue moins performante que la méthodologie utilisant des réseaux de neurones convolutifs (deep learning - apprentissage profond) [Ye2017 - Ye-Net]. Ceci dit, en fonction du scénario, l'utilisation d'un réseau de neurones convolutifs n'est pas toujours la meilleure solution. L'apprentissage d'un réseau de neurones oblige en effet à fournir de nombreux exemples, ce qui n'est pas toujours possible dans un contexte "réel" de stéganalyse. La voie que nous souhaitons explorer dans ce stage consiste à générer de nouveaux exemples "qui ressembleraient" à des exemples "réel". L'objectif est donc d'évaluer le gain que l'on peut obtenir en augmentant artificiellement la base d'apprentissage.

Le réseau que nous utiliseront pour la stéganalyse sera basé sur un réseau inspiré de celui décrit dans [Yedroudj17 - Yed-Net]. Le stagiaire devra prendre en main ce réseau, les bases de données d'image "cover" et "stego" [Bas2011 - BOSS], [Holub2014 - S-UNIWARD], et la méthodologie GAN [Goodfellow2014 - GAN], [Donahue2017 - BiGAN].

Pour mener à bien ce sujet, il est préférable d'avoir certaines connaissances : en traitement des images, et/ou en classification/fouille de données, et/ou en architecture des machines/installation d'OS. Il est également intéressant d'avoir de bonnes bases en programmation et en math.

Profil recherché : Master (M2) ou Ecole d'Ingénieur (3ème année) ayant une bonne maitrise de la programmation (C++, Python…), des connaissances en fouille de données / indexation / classification, traitement des images, sécurité.

Encadrement : Marc CHAUMONT (Enseignant Chercheur), Mehdi YEDROUDJ (Doctorant), Frederic COMBY (Enseignant Chercheur).

Modalité de candidature : Envoyez un CV, une lettre de motivation ainsi que votre relevé de notes de M1 le plus tôt possible. Après pré-sélection des candidatures, des entretiens téléphoniques ou en personne seront planifiés.

Contacts : Marc Chaumont (*marc.chaumont@lirmm.fr*)

Lieu du stage : LIRMM, équipe ICAR.

Période du stage : 1er semestre 2018 (5-6 mois).

Gratification de stage : environ 550€ mois.

Bibliographie:

[Bas2011 - BOSS] P. Bas, T. Filler, and T. Pevny, "'Break Our Steganographic System': ´ The Ins and Outs of Organizing BOSS," in *Proceedings of the 13th International Conference on Information Hiding, IH'2011*, ser. Lecture Notes in Computer Science, vol. 6958. Prague, Czech Republic: Springer, May 2011, pp. 59–70.

[Fridrich2012 - SRM_SRMQ1] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images", IEEE TIFS.

[Holub2013 - PSRM] Random Projections of Residuals as an Alternative to Co-occurrences in Steganalysis, with V. Holub and T. Denemark, Proc. SPIE, Electronic Imaging, Vol. 8665, San Francisco, CA, February 3–7, 2013.

[Holub2014 - S-UNIWARD] V. Holub, J. Fridrich, and T. Denemark, "Universal Distortion Function for Steganography in an Arbitrary Domain," *EURASIP Journal on Information Security, JIS*, vol. 2014, no. 1, 2014.

[Kodovsky2012 - EC] J. Kodovský, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," IEEE Transactions on Information Forensics and Security, vol. 7, no. 2, pp. 432–444, 2012.

[Simmons83] G. J. Simmons, "The prisoners problem and the subliminal channel," in Advances in Cryptography, CRYPTO, Aug. 1983, pp. 51–67.

[Ye2017 - Ye-Net] YeNet J. Ye, J. Ni, and Y. Yi, “Deep Learning Hierarchical Representations for Image Steganalysis,” *IEEE Transactions on Information Forensics and Security, TIFS*, p. 13, 2017.

[Yedroudj17 - Yed-Net] Mehdi Yedroudj, Marc Chaumont, Frédéric Comby, " Yedroudj-Net: un réseaux de neurones efficace pour la stéganalyse spatiale ", CORESA’2017, COmpression et REprésentation des Signaux Audiovisuels, Caen, France, 20-21 septembre, 2017, 6 pages, sciencesconf.org:coresa2017:161315.

[Goodfellow2014 - GAN], Ian J.; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua (2014). "Generative Adversarial Networks". arXiv*:*1406.2661.

[Donahue2017 - BiGAN] Jeff Donahue, Philipp Krähenbühl, Trevor Darrell. "Adversarial Feature Learning", 2017, Published as a conference paper at ICLR 2017. arXiv:1605.09782.