

PROPOSITION DE SUJET DE THESE

Intitulé : Détection et classification de cibles de petite taille par deep learning

Référence : **PHY-DOTA-2020-38**
(à rappeler dans toute correspondance)

Début de la thèse : Dès que possible

Date limite de candidature : Août 2020

Mots clés : deep learning, détection, classification

Profil et compétences recherchées : Master 2 ou Ecole d'ingénieur

Mathématiques appliquées, data science, machine learning, statistiques, optimisation

Présentation du projet doctoral, contexte et objectif

Le développement de capteurs qui permettent d'acquérir de façon simultanée des images dans plusieurs bandes spectrales du spectre visible à l'infrarouge lointain et leur exploitation automatique, rendue possible grâce à l'essor des méthodes de machine learning (SVM, arbres de décision, réseaux de neurones...), ont permis de réaliser ces 20 dernières années de grands progrès en détection, reconnaissance et identification de cibles (DRI). Cependant, il demeure difficile de détecter des cibles de petite taille (entre 1 et 50 pixels sur la surface de la cible), notamment quand elles présentent un faible contraste par rapport au fond ou quand on cherche à les détecter sur un fond texturé comme un ciel nuageux.

Les méthodes d'apprentissage profond (ou deep learning) ont connu un très fort développement ces dernières années, car elles ont démontré qu'elles pouvaient, en extrayant automatiquement des caractéristiques non linéaires bien adaptées à la tâche à accomplir, surpasser les techniques conventionnelles, et même les humains, en termes de performances dans de nombreux domaines : classification automatique, détection et reconnaissance d'objets dans des images, traduction de texte...). Toutefois, en imagerie, ces performances en détection et classification sont obtenues pour des objets de grande taille dans des images RGB. De plus, l'apprentissage des paramètres des réseaux profonds nécessite d'avoir accès à des grandes bases de données annotées (plusieurs centaines de milliers ou plusieurs millions d'exemples) pour éviter le phénomène de sur-apprentissage. Il est difficile d'accéder à ce volume de données pour des instruments fonctionnant en infrarouge.

L'objectif de la thèse sera donc de proposer des méthodes de détection et classification à base de deep learning adaptées au cas de cibles de petites tailles, potentiellement de faible contraste par rapport au fond, pour lesquelles la signature est mesurée ou apprise dans une bande spectrale ou plusieurs bandes spectrales dans différents domaines en visible et infrarouge (les images associées seront appelées dans la suite images multispectrales).

Deux grands types d'approches peuvent être envisagés : la première méthode consiste à utiliser une fenêtre glissante permettant d'extraire des imagerie 2D ou 3D (spatial + spectral) que l'on peut ensuite entrer successivement dans une architecture de type autoencoder, réseau dont l'objectif est d'apprendre une représentation de dimension réduite d'un ensemble de données qui permette de reconstruire au mieux les entrées, ou CNN (Convolutional Neural Network) afin de mener à bien la détection puis la classification. La seconde méthode consiste à s'appuyer sur une méthode de segmentation sémantique. On obtient alors une carte de segmentation où chaque pixel porte une information sémantique sur la classe à laquelle il appartient.

Le doctorant sera ainsi amené à s'approprier ces deux approches. Afin de pallier le manque de données dans le domaine infrarouge, on s'intéressera aux méthodes d'apprentissage par transfert (transfer learning), qui permettent de compenser l'absence de données étiquetées et de tirer profit de l'information associée à un domaine donné (images multimédia par exemple) pour résoudre un problème dans un autre domaine (images infrarouge dans notre cas).

Le doctorant s'intéressera, dans un premier temps, aux méthodes de détection d'anomalies et aux algorithmes non supervisés qui permettent de pallier le manque de données étiquetées. La détection d'anomalies est couramment utilisée pour rechercher les cibles potentielles d'intérêt sur une image. Le principe consiste à analyser le signal lu sur les pixels de l'image et à distinguer deux classes : une classe

majoritaire associée au fond, et une classe de pixels qui se démarquent du signal de fond, les anomalies. Celles-ci peuvent ensuite être analysées de façon plus détaillée par des méthodes de classification. Plusieurs méthodes de détection d'anomalies par deep learning prometteuses ont été proposées récemment et leur adaptation aux images mono et multispectrales infrarouge sera envisagée : des méthodes à base de réseaux antagonistes génératifs (GAN) [Schlegl2019], ou transposant le concept de one-class SVM, très utilisé en détection d'anomalies dans des images hyperspectrales, acquises dans une centaine de fines bandes spectrales contiguës, aux réseaux de neurones [Ruff2018], et enfin les méthodes les plus courantes, à base d'autoencoders. Le point clé se situe dans le choix du critère pour déterminer les anomalies. Pour les autoencoders, par exemple, le critère le plus évident est l'erreur de reconstruction, mais le choix du seuil est délicat. Une solution pour s'en affranchir pourrait être d'estimer le fond à l'aide d'un auto-encodeur, puis de réaliser une détection a contrario sur une image différence par rapport à ce modèle.

Une autre piste intéressante, notamment pour la classification, serait de s'inspirer des travaux de la thèse de F. Maire sur les modèles déformables [Maire2017], dans lesquels les observations sont considérées comme des réalisations d'un modèle statistique d'observations fluctuant spectralement et spatialement autour de formes caractéristiques inconnues et qui ont donné des très bons résultats pour détecter et reconnaître des avions faiblement résolus sur fonds nuageux. Les travaux d'enrichissement d'un des réseaux les plus utilisés en détection d'objets, Faster R-CNN [Ren2015], menés dans le cadre de la thèse de J. Ogier du Terrail [Ogier2018] ont permis de prendre en compte des boîtes tournantes et d'obtenir de très bonnes performances pour la détection de petits véhicules en imagerie aérienne. Ils pourraient être complétés par l'utilisation de méthodes de pooling déformable [Mordan2018] afin d'adresser des petites cibles non rectangulaires. Des approches à base de mélanges d'autoencoders [Zong2018] ou de GANs [Springenberg2016] ont été publiées récemment et pourraient également servir de base à une adaptation du concept en version deep learning, afin de prendre en compte la variabilité des cibles et des fonds.

Ces développements méthodologiques seront mis en œuvre et évalués dans le cadre de différentes applications de détection et classification (visée rasante sol-sol ou sol-air sur des fonds naturels : zone rurale, zone désertique, forêt, fond nuageux, milieu maritime...), pour lesquelles nous disposons de données simulées, en grand nombre, et de données mesurées, en nombre plus restreint. Ces données seront prélevées dans des bases de données de viseurs ou caméras multivoies, comportant par exemple du visible couleur, du proche infrarouge, du thermique en bande MWIR ou LWIR. Certains viseurs disposent en outre d'une voie SWIR ou d'un capteur multispectral dans la bande proche-infrarouge.

Les résultats seront comparés à ceux obtenus à l'aide des méthodes de segmentation sémantique, du type U-net par exemple [Ronneberger2015] et à leur transposition aux images mono ou multispectrales infrarouge.

Bibliographie

[Maire2017] F. Maire, S. Lefebvre, E. Moulines, Online EM for Functional Data, Computational Statistics and Data Analysis 111, 2017.

[Mordan2018] T. Mordan, N. Thome, G. Henaff, M. Cord, End-to-End Learning of Latent Deformable Part-Based Representations for Object Detection, Int J Comput Vis, 10.1007/s11263-018-1109-z, 2018.

[Ogier2018] J. Ogier du Terrail, Réseaux de neurones convolutionnels profonds pour la détection de petits véhicules en imagerie aérienne, thèse CIFRE Safran Electronics and Defense 2018.

[Ren2015] S. Ren, K. He, R. Girshick, J. Sun, Faster R-CNN : Towards real-time object detection with region proposal networks, Advances in Neural Information Processing Systems (NIPS), 2015.

[Ronneberger2015] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI, Springer, 9351, 2015.

[Ruff2018] L. Ruff, R. A. Vandermeulen, N. Görnitz, et al, Deep One-Class Classification, ICML 2018.

[Schlegl2019] T. Schlegl, P. Seeböck, S. M. Waldstein et al, f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks, Medical Image Analysis 54, 2019.

[Springenberg2016] J. T. Springenberg, Unsupervised and semi-supervised learning with categorical generative adversarial networks, ICLR 2016.

[Zong2018] B. Zong, Q. Song, M. Renqiang Min et al, Deep autoencoding Gaussian Mixture Model for Unsupervised anomaly detection, ICLR 2018.

Collaborations envisagées : Safran E & D (thèse CIFRE)

Laboratoire d'accueil à l'ONERA

Département : Optique et Techniques Associées

Lieu (centre ONERA) : Palaiseau

Contact : Sidonie Lefebvre

Tél. : 0180386376

sidonie.lefebvre@onera.fr

Email :

Directeur de thèse

Nom : Sylvie Le Hégarat-Masclé

Laboratoire : Université de Paris Sud

Tél. :

Email : sylvie.le-hegarat@u-psud.fr

Pour plus d'informations : <https://www.onera.fr/rejoindre-onera/la-formation-par-la-recherche>