

Sujet : De l'image au texte : l'apprentissage profond et l'indexation en Imagerie par Rayons X (stage M2/Ingénieur ou thèse CIFRE)

Sujet de recherche proposé par l'entreprise **Smiths Detection** en partenariat avec le laboratoire d'Informatique **LIPADE**¹ de l'**Université Paris-Descartes**.

→ Ce sujet peut être traité sous la forme d'un stage de fin d'étude (niveau M2 / Ingénieur) suivi d'une thèse CIFRE ou être directement abordé dans le cadre d'une thèse CIFRE.

Mots clefs : imagerie par rayons X, « deep learning », indexation d'images, « image2text », « visual learning », recherche de matières illicites

Contexte : Smiths Detection est l'une des cinq divisions opérationnelles de Smiths Group plc, un groupe d'ingénierie d'envergure mondiale coté à la bourse de Londres. Née en 1997, l'entreprise s'est développée en intégrant les compétences technologiques d'acteurs clés sur de nombreux marchés afin de créer un leader mondial du secteur. En août 2003, l'entreprise est devenue une division à part entière du groupe Smiths. Elle développe et commercialise des systèmes d'imagerie utilisant des rayons X pour le contrôle des bagages dans les aéroports et sites « sensibles » mais aussi pour **contrôler les cargaisons au niveau des ports et des frontières terrestres**. Elle offre des solutions avancées en matière de sécurité à des fins de détection et d'identification des explosifs, des agents chimiques et biologiques, des armes et des drogues.

Passer aux rayons X des containers ou des voitures exige un niveau de radiation bien supérieur à ce qui est par exemple nécessaire pour les applications médicales ou le contrôle des bagages dans les aéroports. Pour obtenir ces rayons X très énergétiques, on utilise un accélérateur d'électrons, qui peut être linéaire ou circulaire (plus compact). Les particules γ sont accélérées jusqu'à atteindre plusieurs millions d'électron-volts (MeV). Elles se dirigent ensuite vers une cible de freinage (aussi appelée Bremstarget), qui les transforme en rayons X à spectre continu. L'énergie des rayons X correspond alors à celle des électrons accélérés.

Les rayons X sont ensuite soigneusement collimatés pour obtenir des faisceaux en éventail dans un plan vertical. Les rayons traversent l'objet et viennent frapper des détecteurs scintillants, adaptés au niveau d'énergie des rayons X. Des photodiodes transforment le signal en impulsions électriques. Après amplification et numérisation, ces dernières sont alors envoyées vers le module de traitement des données. Le système se déplace le long du chargement en prenant des vues en coupe successives, qu'il rassemble ensuite pour former l'image entière. À la différence de la radioscopie médicale, qui affiche les résultats sous forme d'un négatif, ici les objets qui absorbent le plus de rayonnement sont aussi les plus sombres sur l'image (voir Figure 1).

Enfin, les produits et technologies de scanning aux rayons X de Smiths Detection aident les agents de sécurité ou de douane à prendre des décisions délicates. Notre priorité est de fournir la meilleure qualité d'image possible pour faciliter la prise de décision. La possibilité de distinguer différents types de matériaux constitue un atout indéniable dans cette optique.

Objectifs du sujet de recherche : Les équipes CIS (Cargo Inspection System) comptent des ingénieurs et des scientifiques qui couvrent un spectre complet en matière de recherche, depuis les développements de base, jusqu'à l'intégration de systèmes complexes.

¹ <http://lipade.mi.parisdescartes.fr/>

Pour concevoir et réaliser des systèmes d'un haut niveau de technologie, Smiths Detection mobilise depuis ses origines d'importants moyens en recherche et développement en traitement d'images et « machine learning ». Ainsi, afin d'aider les agents au contrôle et à l'inspection des chargements, nous désirons mener des recherches sur différentes questions relatives à l'analyse (semi-)automatique d'images par rayons X. En particulier, le sujet de stage + thèse CIFRE proposé s'intéresse aux domaines de l'indexation d'images par « deep learning » (*apprentissage profond*), en particulier sur des images de véhicules et de conteneurs. Indexer une image consiste en la vectorisation de cette dernière via l'extraction de caractéristiques visuelles (textures, couleur, indices de formes) pour permettre un stockage dans une base de données et une recherche ultérieure d'informations.

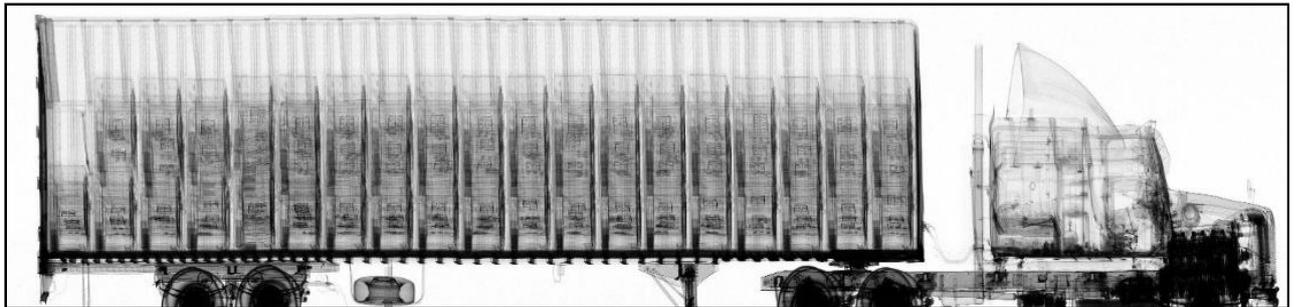


Figure 1 : image à rayons X d'un camion

Cette problématique intervient dans notre contexte dans deux cas d'usage :

- Aider un opérateur débutant à profiter de l'expertise de ses collègues experts pour interpréter le contenu d'une image en lui proposant, automatiquement, les verdicts de photo-interprétation de ses derniers, via une recherche d'images par le contenu [Kurtz2014] ;
- Générer automatiquement des rapports textuels décrivant le contenu du chargement à partir de l'image : « visual learning » ou « image2text » [Karpathy2017].

Dans les deux cas, l'objectif commun est de considérer une représentation haut-niveau du contenu des images permettant une indexation efficace, rapide et « évolutive » (robuste envers l'ajout de nouvelles images dans la collection).

Actuellement, la plupart des moteurs d'indexation par le contenu se basent sur des représentations dites intermédiaires (l'avant dernière couche d'un réseau de neurones, par exemple). Il a été montré que les représentations sémantiques des images, dans laquelle le vecteur caractéristique est formé de probabilités associées à des concepts sémantiques tels « pistolet », « cigarette », « chaise », « conserves » etc. constituent une bonne alternative. Le but de ce sujet de recherche est ici de coupler ces deux approches pour permettre de prédire automatiquement la description sémantique d'une image (de la notion de mots-clés à la notion de phrases) à partir de modèles d'apprentissage profond [Krishna2017, Li2018] entraînés sur des bases de données existantes.

Les verrous scientifiques sont liés principalement aux détournements de modèles entraînés sur des bases d'images naturelles (ImageNet, etc.) à l'analyse d'images à rayons X de chargements.

Profil du candidat :

- Ingénieur ou Master 2 avec des fortes compétences en Programmation/Algorithmique
- Connaissances en apprentissage profond (deep learning)
- Connaissances en Vision par ordinateur

et un gout prononcé pour la lecture/rédaction (en anglais) de documents (brevets, publications) scientifiques.

Aspects techniques :

Le développement logiciel sera réalisé en Python et C++. Le doctorant sera amené à utiliser les bibliothèques de traitement d'images développées par l'entreprise, et à y intégrer le développement de modules d'indexation réalisée pendant la thèse.

Candidatures :

Envoyer un CV étendu et une lettre de motivation à :

- Najib GADI <najib.gadi@smithsdetection.com>
- Camille KURTZ <camille.kurtz@parisdescartes.fr>

Références :

[Krishna2017] Ranjay Krishna, Yuke Zhu, Oliver Groth, Justin Johnson, Kenji Hata, Joshua Kravitz, Stephanie Chen, Yannis Kalantidis, Li-Jia Li, David A. Shamma, Michael S. Bernstein, Li Fei-Fei: Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations. International Journal of Computer Vision 123(1): 32-73 (2017)

[Li2018] Christy Y. Li, Xiaodan Liang, Zhiting Hu, Eric P. Xing: Hybrid Retrieval-Generation Reinforced Agent for Medical Image Report Generation. Proceedings of NIPS (2018)

[Karpathy2017] Andrej Karpathy, Li Fei-Fei: Deep Visual-Semantic Alignments for Generating Image Descriptions. IEEE Trans. Pattern Anal. Mach. Intell. 39(4): 664-676 (2017)

[Kurtz2014] Camille Kurtz, Adrien Depeursinge, Sandy Napel, Christopher F. Beaulieu, Daniel L. Rubin: On combining image-based and ontological semantic dissimilarities for medical image retrieval applications. Medical Image Analysis 18(7): 1082-1100 (2014)