

Sujet de Stage de Master 2 / PFE,  
suivi d'une thèse de doctorat en contrat CIFRE

## Apprentissage profond pour la détection de défauts sur des objets en contexte industriel

### Contexte

L'équipe Imagine du LIRIS (Laboratoire en Image et Système d'Information) a notamment pour objectif de développer de nouvelles méthodes de vision par ordinateur pour analyser automatiquement le contenu d'images ou de vidéos. Ces méthodes sont appliquées dans divers contextes, notamment dans le contexte industriel. Ainsi, elle collabore depuis 4 ans avec le groupe DEMS, bureau d'étude et de design français spécialisé dans le développement de véhicules (nautisme, automobiles, véhicules sanitaires/professionnels, véhicules industriels lourds, nouvelles mobilités), de produits et de machines. La performance du groupe DEMS se base sur une approche projet innovante : le design, les usages, les facteurs humains, les nouvelles technologies et l'ingénierie travaillent ensemble afin de développer une idée jusqu'au produit final. La recherche et le développement sont au cœur des activités de l'entreprise.

Une première collaboration a eu lieu entre le LIRIS et DEMS dans le cadre de la thèse de Julia Cohen [1, 2, 3, 4] qui s'est concentrée sur la détection d'objets industriels à l'aide de modèles 3D dans des images égocentriques. Nous comptons poursuivre cette collaboration dans le contexte d'un stage puis d'une nouvelle thèse CIFRE dont le sujet est décrit ci-dessous.

### Sujet

Dans l'industrie, l'assemblage de produits peut de nos jours bénéficier d'outils basés sur la réalité augmentée. Ainsi les opérateurs munis de lunettes de réalité augmentée ou de tablettes peuvent être guidés dans le montage ou la manipulation d'objets. En effet, des modèles 3D des objets et des produits finis sont en général disponibles. Basés sur ces données, il est possible de développer des méthodes de vision par ordinateur permettant la détection des objets dans les vidéos issues des capteurs égocentrés pour communiquer des informations aux opérateurs.

Plus précisément, dans le cadre de cette thèse, nous nous concentrons sur la phase finale de montage d'un objet afin de vérifier par inspection vidéo que l'objet est conforme à ce qui est attendu. Il s'agira d'une part de détecter les éventuelles pièces manquantes et d'autre part, de vérifier l'emplacement des pièces mises en place. Les travaux de cette thèse devraient également amener à une étude de la typologie des objets reconnaissables par un tel procédé (taille nécessaire des objets apparents dans les images, impact de leur couleur/texture notamment).

Les méthodes développées se baseront sur l'apprentissage profond et prendront en entrée une représentation 3D de la scène, acquise par un opérateur à partir d'un point de vue égocentrique de la scène. A noter que l'opérateur pourra être guidé dans l'acquisition de la vidéo. L'apprentissage de tels modèles sera fait en se basant sur des données multimodales (représentation 3D de l'objet et de ses pièces constitutives, et éventuellement des informations textuelles, telles que des instructions d'assemblage).

Dans un premier temps, un état de l'art sera fait sur les méthodes de la littérature autour des de l'analyse automatique de l'assemblage d'objets (par exemple, [11, 12, 13]), de la métrologie virtuelle (par exemple [5, 6]), ainsi que de l'apprentissage profond en lien avec la détection d'objets à partir de modèles 2D et 3D (par exemple [7,8]) et l'adaptation de domaine (par exemple [9,10]).

Dans un second temps, des méthodes adaptées à la problématique précise de la thèse seront mise en place prenant en compte l'apprentissage par adaptation de domaine, l'analyse automatique de l'assemblage d'objets par inspection vidéo, et le traitement embarqué et temps réel, indispensable dans ce contexte.

Pour tester ces méthodes et leurs limites, des jeux de données seront proposés en coordination avec l'entreprise DEMS. Ces jeux de données seront mis à disposition de la communauté scientifiques afin que les chercheurs du domaine puissent tester leurs algorithmes.

## Modalités du stage et de la thèse

Le stage et la thèse se dérouleront en alternance au laboratoire LIRIS (Bron, 69) et dans l'entreprise DEMS (Saint Bonnet de Mure, 69).

Indemnité de stage : indemnités légales + indemnités transport

Salaire brut mensuel pour la thèse : environ 2300 €.

Le stage se déroulera de février 2023 à juillet 2023 (dates à adapter en fonction de la formation).

La thèse débutera en septembre 2023, pour une durée de 36 mois.

## Profil recherché

Le candidat doit être inscrit dans un Master 2 en informatique ou dans une école d'ingénieurs en Informatique et avoir de bonnes connaissances en analyse d'images ou de vidéos et en deep learning.

La complexité du sujet et des algorithmes à manipuler au cours de ces travaux, et le contexte industriel de cette thèse (bourse Cifre), demande à ce que le candidat soit extrêmement motivé, rigoureux, et exigeant vis-à-vis de lui-même.

L'étudiant évoluera dans un environnement innovant au sein d'une société jeune et en lien avec un laboratoire de renommée internationale de plus de 350 chercheurs.

### Compétences attendues :

- Langage Python
- Bibliothèque OpenCV

### Compétences souhaitées :

- Logiciel de gestion de version (GIT)
- Framework PyTorch.

## Contacts :

Laure Tougne Rodet : [laure.tougne@univ-lyon2.fr](mailto:laure.tougne@univ-lyon2.fr)

Carlos Crispim-Junior : [carlos.crispimjunior@univ-lyon2.fr](mailto:carlos.crispimjunior@univ-lyon2.fr)

## Bibliographie

- [1] Julia Cohen, Carlos Crispim-Junior, Jean-Marc Chiappa et Laure Tougne: Training an embedded object detector for industrial settings without real images. In 2021 IEEE International Conference on Image Processing (ICIP), pages 714-718. IEEE, 2021a.
- [2] Julia Cohen, Carlos Crispim-Junior, Céline Grange-Faivre et Laure Tougne: CAD-based learning for egocentric object detection in industrial context. In 15th International Conference on Computer Vision Theory and Applications (VISAPP), volume 5, pages 644-651. SCITEPRESS - Science and Technology Publications, 2020.
- [3] Julia Cohen, Carlos F Crispim-Junior, Jean-Marc Chiappa et Laure Tougne : MobileNet SSD : étude d'un détecteur d'objets embarquable entraîné sans images réelles. In ORASIS 2021, Saint Ferréol, France, septembre 2021b. Centre National de la Recherche Scientifique [CNRS]. URL <https://hal.archives-ouvertes.fr/hal-03531390>.
- [4] Julia Cohen, Carlos F Crispim-Junior, Jean-Marc Chiappa et Laure Tougne : Industrial object detection with multi-modal SSD: closing the gap between synthetic and real images. Machine Vision and Applications, in revision.
- [5] Dreyfus Paul-Arthur, Psarommatos Foivos, May Gokan, Kiritsis Dimitris Virtual : Metrology as an approach for product quality estimation in Industry 4.0: a systematic review and integrative conceptual framework. In International Journal of Production Research, 742-765, 60 : 2, 2022.
- [6] T. C. Tin, S. C. Tan and C. K. Lee, "Virtual Metrology in Semiconductor Fabrication Foundry Using Deep Learning Neural Networks," in *IEEE Access*, vol. 10, pp. 81960-81973, 2022, doi: 10.1109/ACCESS.2022.3193783.
- [7] R. Qi Charles, Hao Su, Mo Kaichun et Leonidas J. Guibas : PointNet : Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, jul 2017.
- [8] Armen Avetisyan, Manuel Dahnert, Angela Dai, Manolis Savva, Angel X. Chang et Matthias Nießner : Scan2cad : Learning CAD model alignment in RGB-d scans. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, jun 2019.
- [9] Stefan Hinterstoisser, Vincent Lepetit, Paul Wohlhart et Kurt Konolige : On pre-trained image features and synthetic images for deep learning. In Proceedings of the IEEE European Conference on Computer Vision (ECCV), pages 682-697, 2018.
- [10] Fernando Camaro Nogues, Andrew Huie et Sakyasingha Dasgupta : Object detection using domain randomization and generative adversarial refinement of synthetic images. arXiv preprint arXiv :1805.11778, 2018.
- [11] Y. Li, K. Mo, L. Shao, M. Sung, et L. Guibas : Learning 3D Part Assembly from a Single Image, 2020, arXiv <https://doi.org/10.48550/arXiv.2003.09754>
- [12] Y. Li, S. Kan, J. Yuan, W. Cao and Z. He : Spatial Assembly Networks for Image Representation Learning. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 13871-13880, 2021, doi: 10.1109/CVPR46437.2021.01366.
- [13] K. Zakka, A. Zeng, J. Lee, S. Song : Form2Fit: Learning Shape Priors for Generalizable Assembly from Disassembly. ArXiv, 2019, <https://doi.org/10.48550/arXiv.1910.13675>