

Référence d'offre de stage : S12024 ([English description](#) : page 2 ; [Description Française](#) : page 1)

## Stage – Master 2 en apprentissage automatique et vision par ordinateur

### « Apprentissage transductif pour la détection d'objets sans exemple sur des images aériennes »

Ce stage s'inscrit dans le cadre du LabCom IRISER, Laboratoire Commun en « Intelligence, Reconnaissance, Surveillance Réactive » (<https://www-l2ti.univ-paris13.fr/iriser/>) financé par l'ANR. Le LabCom IRISER vise à proposer, à maîtriser de bout en bout le comportement et les performances des systèmes intelligents voire embarqués destinés à la vision artificielle pour l'analyse rapide et automatisée des images/vidéos (de très grandes tailles, multispectrales géoréférencées de hautes résolutions) capturées à bord des avions de COSE en s'appuyant sur des stratégies de traitement de l'information visuelle et d'apprentissage automatique. Nous souhaitons recruter un stagiaire de master en recherche et développement en vision par ordinateur et apprentissage machine sur une durée de 5 mois.

#### Description succincte du projet scientifique

Les avancées rapides dans les tâches d'apprentissage supervisé peuvent être attribuées à la disponibilité d'ensembles de données annotées de manière extensive et aux capacités des réseaux profonds à grande capacité. En effet, les réseaux neuronaux convolutifs profonds (CNN) affichent des performances équivalentes à celles des humains dans des tâches spécialisées impliquant la reconnaissance visuelle et la catégorisation fine. Cependant, certains domaines sont confrontés à des défis pour obtenir des annotations à grande échelle, que ce soit en raison de la nécessité d'une expertise spécialisée ou en raison de la rareté d'échantillons visuels dans le monde réel, tels que des espèces rares. L'apprentissage sans exemple (Zero-Shot Learning, ZSL) émerge comme une solution pour de tels scénarios où aucun exemple visuel des classes non-vues n'est disponible pendant l'entraînement. Il est à noter que les approches traditionnelles de ZSL sont principalement limitées au cadre de la reconnaissance (classification). La détection d'objets sans exemple (Zero-shot object detection, ZSD) représente un défi récemment introduit visant à identifier et classifier simultanément des classes d'objets non-vues. Contrairement à la tâche de reconnaissance, le ZSD est notablement plus difficile en raison de la nature intrinsèquement ambiguë et des complexités associées à la localisation de catégories entièrement inconnues. La complexité est exacerbée dans un cadre généralisé de ZSD, où l'hypothèse est que des objets vus et non vus peuvent coexister pendant l'inférence. Les approches développées (voir par exemple [2]) pour résoudre le problème de ZSD explorent principalement un cadre inductif, reposant uniquement sur des exemples étiquetés du domaine source pour l'entraînement. En réalité, un écart substantiel existe entre les domaines source (objets vus) et cible (non vus). Pour combler cet écart, un cadre transductif pour l'apprentissage sans exemple (ZSL) suppose qu'une partie des échantillons non étiquetés du domaine cible est disponible pendant l'entraînement.

Dans ce stage, il s'agira d'étudier l'impact du cadre transductif utilisé dans [1] pour résoudre les problèmes de déplacement de domaine et de biais de modèle et d'évaluer les performances du modèle de détection d'objets sur des images aériennes (DOTA, DIOR). Une importance particulière sera accordée aux petits objets.

**Mots clefs** : Object Detection, Zero-Shot Learning, Transformers, Deep learning, Aerial Images.

#### References :

- [1] Shafin Rahman, Salman Khan and Nick Barnes, Transductive Learning for Zero-Shot Object Detection, ICCV 2019
- [2] Ankan Bansal, Karan Sikka, Gaurav Sharma, Rama Chellappa, and Ajay Divakaran. Zero-shot object detection. In The European Conference on Computer Vision (ECCV), September 2018

**Profil du candidat(e)** : Vous êtes issu(e) d'une formation (Master 2) en sciences de données (traitement du signal des images, informatique avec une dominante en intelligence artificielle/apprentissage/machine learning, mathématiques appliquées) et vous êtes à l'aise en développement informatique (Python/ framework Pytorch, Tensorflow). Vous maîtrisez le traitement d'images et les techniques d'intelligence artificielle (Deep learning ...). Vous êtes à l'aise en Anglais aussi bien à l'écrit qu'à l'oral. Vous êtes motivé (e) et curieux (se) de nature, dans ce cas venez participer à la création du LabCom IRISER.

**Durée du stage** : 5 mois

**Date de début prévue** : 15/03/2024 -- **Date limite de dépôt** : 29/02/2024

**Contact** : Anissa MOKRAOUI, Hicham TALAOUBRID

**Courriel** : [anissa.mokraoui@univ-paris13.fr](mailto:anissa.mokraoui@univ-paris13.fr); [hicham.talaoubrid1@edu.univ-paris13.fr](mailto:hicham.talaoubrid1@edu.univ-paris13.fr)

**Dossier de candidature** : CV, lettre de motivation, relevés de notes, au moins une lettre de recommandation (merci d'indiquer la référence d'offre de stage dans le sujet de votre courriel).

Référence d'offre de stage : S22024

## Internship – Master's Degree in Machine Learning and Computer Vision « Transductive Learning for Zero-Shot Object Detection in Aerial Images »

This internship is part of the LabCom IRISER, a Common Laboratory in "Intelligence, Reconnaissance, Reactive Surveillance" (<https://www-l2ti.univ-paris13.fr/iriser/>) funded by the ANR. The LabCom IRISER aims to propose end-to-end control over the behavior and performance of intelligent or embedded systems intended for artificial vision for the rapid and automated analysis of images/videos (of very large sizes, georeferenced multispectral images of high resolution) captured aboard COSE aircraft, relying on strategies for visual information processing and machine learning. We are looking to recruit a master's research and development intern in computer vision and machine learning for a duration of 5 months.

### Brief Description of the Scientific Project:

The rapid advancements in supervised learning tasks can be attributed to the availability of extensively annotated datasets and the capabilities of deep networks with high capacity. Deep Convolutional Neural Networks (CNNs), in particular, exhibit performance comparable to humans in specialized tasks involving visual recognition and fine categorization. However, some domains face challenges in obtaining large-scale annotations, either due to the need for specialized expertise or the scarcity of visual samples in the real world, such as rare species.

Zero-Shot Learning (ZSL) emerges as a solution for scenarios where no visual examples of unseen classes are available during training. Traditional approaches to ZSL are primarily limited to the realm of recognition (classification). Recently introduced, Zero-Shot Object Detection (ZSD) represents a challenge aiming to simultaneously identify and classify unseen object classes. Unlike the recognition task, ZSD is notably more challenging due to the inherently ambiguous nature and complexities associated with locating entirely unknown categories. The complexity is exacerbated in a generalized ZSD framework, where the assumption is that seen and unseen objects may coexist during inference.

Approaches developed to solve the ZSD problem primarily explore an inductive framework, relying solely on labeled examples from the source domain for training. In reality, there is a substantial gap between the source (seen objects) and target (unseen objects) domains. To bridge this gap, a transductive framework for zero-shot learning assumes that a portion of unlabeled samples from the target domain is available during training.

During this internship, the objective is to study the impact of the transductive framework used in [1] to address domain shift and model bias issues and evaluate the performance of the object detection model on aerial images (DOTA, DIOR). Special attention will be given to small objects.

**Key words** : Object Detection, Zero-Shot Learning, Transformers, Deep learning, Aerial Images.

### References :

- [1] Shafin Rahman, Salman Khan and Nick Barnes, Transductive Learning for Zero-Shot Object Detection, ICCV 2019
- [2] Ankan Bansal, Karan Sikka, Gaurav Sharma, Rama Chellappa, and Ajay Divakaran. Zero-shot object detection. In The European Conference on Computer Vision (ECCV), September 2018

### Candidate Profile:

You hold a Master's degree (Master 2) in data science (image signal processing, computer science with a focus on artificial intelligence/machine learning, applied mathematics), and you have proficiency in computer programming (Python/PyTorch framework, Tensorflow). You possess expertise in image processing and artificial intelligence techniques (Deep learning, etc.). You are comfortable with both written and spoken English. If you are motivated and naturally curious, we invite you to join the creation of LabCom IRISER.

**Internship Duration:** 5 months Scheduled

**Start Date:** 15/03/2024 -- **Application Deadline:** 29/02/2024

**Contact:** Anissa MOKRAOUI, Hicham TALAOUBRID

**Email:** [anissa.mokraoui@univ-paris13.fr](mailto:anissa.mokraoui@univ-paris13.fr) , [hicham.talaoubrid1@edu.univ-paris13.fr](mailto:hicham.talaoubrid1@edu.univ-paris13.fr)

**Application Documents:** CV, cover letter, transcripts, at least one letter of recommendation (please indicate the internship offer reference in the subject of your email).