

STAGE DATA SCIENTIST

Début : Dès que possible, Temps plein, Durée : 6 mois

Alteia

Alteia est une plateforme de gestion des actifs industriels basée sur l'intelligence artificielle. L'entreprise capitalise sur une décennie d'expérience dans la capture et l'analyse d'images. Son offre d'Intelligence Visuelle combine ainsi le meilleur de la vision par ordinateur et des technologies d'IA. La plateforme ingère et structure de très grandes quantités de données terrain en une source unique de vérité, qui permet d'élaborer des modèles prédictifs pour les infrastructures industrielles. Elle optimise l'ensemble du cycle de vie d'une opération, simplifie la gestion des risques et fournit des informations en temps réel sur ces infrastructures.

Vos missions

Au sein de nos bureaux Toulousains, vous rejoindrez une équipe de data scientists spécialisée en deep learning, traitement d'images et géomatique, en charge de la R&D sur les outils d'analyse de données déployés sur la plateforme delair.ai. L'expertise de l'équipe porte sur l'analyse de données de sources très variées : d'images de drones, nuages de points LIDAR, modèle CAO, ...

3 thématiques sont envisagées selon votre profil :

- **Deep learning appliqué à des données 3D**

Votre tâche principale sera de concevoir des algorithmes pour la segmentation sémantique des nuages des points.

Les nuages des points issus de capteurs LIDAR et/ou d'un traitement photogrammétrique ont généralement des tailles importantes et ils nécessitent des algorithmes robustes et performants pour être analysés.

Dans un premier temps, vous prendrez en main un modèle de segmentation sémantique développé en interne pour vous familiariser avec la problématique.

Dans un deuxième temps, vous devrez implémenter et comparer différentes architectures de réseau de neurones adapté à ce type de donnée : deep nets, 3D CNN, point-net, graph CNN, ...

Compétences appréciées :

- Machine Learning/ Deep learning concepts
- Python
- Expérience avec une librairie de machine learning (PyTorch, Tensorflow, sklearn)

Liens :

- PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space : <https://arxiv.org/abs/1706.02413>
- KPConv: Flexible and Deformable Convolution for Point Clouds : <https://arxiv.org/abs/1904.08889>
- PyTorch Points 3D : <https://github.com/nicolas-chaulet/torch-points3d>

- **Deep learning appliqué à la segmentation sémantique d'image**

Votre tâche principale sera de concevoir des algorithmes pour la segmentation sémantique sur des images orthorectifiées issues d'un relevé drone d'infrastructure industrielle (agriculture / mine et carrière).

Dans un premier temps, vous prendrez en main un modèle de segmentation sémantique développé en interne pour vous familiariser avec la problématique.

Dans un deuxième temps, vous devrez implémenter et tester différentes architectures de réseau de neurones qui utilisent des approches semi-supervisées ou faiblement supervisées.

Compétences appréciées :

- Machine Learning/ Deep learning concepts
- Python
- Expérience avec une librairie de machine learning (PyTorch, Tensorflow, sklearn)

Liens :

- Semi-supervised semantic segmentation needs strong, varied perturbations: <https://arxiv.org/pdf/1906.01916v4.pdf>
- Hierarchical multi-scale attention for semantic segmentation : <https://arxiv.org/pdf/2005.10821v1.pdf>
- Invariant Information Clustering for Unsupervised Image Classification and Segmentation : <https://arxiv.org/pdf/1807.06653.pdf>

- **Deep learning appliqué à la détection d'objet sur des images**

Votre tâche principale sera de concevoir des algorithmes pour la segmentation sémantique sur des images issue d'un relevé drone d'infrastructure industrielle (lignes électriques).

Dans un premier temps, vous prendrez en main les modèles de détection d'objet développés en interne pour vous familiariser avec la problématique.

Dans un deuxième temps, vous devrez implémenter et tester différentes architectures de réseau de neurones qui visent à améliorer les performances grâce une approche dite multi-view qui utilise efficacement différents points de vue sur un même objet.

Compétences appréciées :

- Machine Learning/ Deep learning concepts
- Python
- Expérience avec une librairie de machine learning (PyTorch, Tensorflow, sklearn)

Liens :

- Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks <https://arxiv.org/abs/1506.01497>
- D2-Net: A Trainable CNN for Joint Detection and Description of Local Features <https://arxiv.org/abs/1905.03561>
- End-to-end object detection with Transformers <https://ai.facebook.com/blog/end-to-end-object-detection-with-transformers/>

Votre profil

Etudiant(e) en dernière année d'un Bac+5 type Master 2 ou Ecole d'Ingénieur avec une spécialisation en traitement d'images ou Data science.