

PROPOSITION DE STAGE EN COURS D'ETUDES

Référence : **DTIS-2024-70**
(à rappeler dans toute correspondance)

Lieu : Palaiseau

Département/Dir./Serv. : DTIS/IVA

Tél. : 01 80 38 65 41

Responsable(s) du stage : Baptiste Abeloos

Email : baptiste.abeloos@onera.fr

DESCRIPTION DU STAGE

Thématique(s) : Intelligence Artificielle et Décision

Type de stage : Fin d'études bac+5 Master 2 Bac+2 à bac+4 Autres

Intitulé : **Explicabilité de l'intelligence artificielle pour la détection d'objets**

Sujet : Les modèles d'intelligence artificielle et en particulier les réseaux de neurones permettent aujourd'hui d'obtenir d'excellents résultats dans le domaine de la vision par ordinateur pour effectuer des tâches comme la classification d'images, la détection d'objets, la segmentation, etc. Cependant de nombreuses applications nécessitent une compréhension claire et détaillée des processus de décision des algorithmes. C'est par exemple le cas pour les systèmes critiques impliquant des vies humaines.

Les réseaux de neurones utilisés en vision par ordinateur n'offrent pas cette transparence, ce qui limite la confiance dans les modèles, et freine les processus de certification pour les applications industrielles ainsi que leur utilisation dans la vie courante.

En vision par ordinateur, de nombreuses méthodes telles que la méthode Grad-CAM [1] ont été développées afin d'identifier les régions de l'images qui ont eu un impact sur la décision finale du modèle, principalement pour des tâches de classification d'image. La nature de ce type d'explication, qui prend la forme d'une carte de chaleur, reste cependant limitée et ne donne pas d'indication sur le comportement intrinsèque du modèle. De plus, sans vérité terrain pour l'explication attendue, l'évaluation quantitative des explications produites et la comparaison des modèles explicatifs entre eux reste un sujet actif de recherche [2].

L'objectif de ce stage est d'aborder les défis scientifiques permettant de développer une méthode d'explicabilité pour un cas d'usage de détection d'objet [3]. Il s'agira dans un premier temps de se familiariser avec les méthodes d'explicabilité développées dans l'état de l'art, ainsi qu'avec les métriques qualitatives et quantitatives permettant d'évaluer une méthode [2]; puis dans un second temps de concevoir la méthode en proposant une définition de ce qu'est une explication du comportement du modèle pour le cas d'usage traité.

[1] R. R. Selvaraju et al. Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization. CoRR, 2016.

[2] Meike. Nauta et al. From Anecdotal Evidence to Quantitative Evaluation Methods: A Systematic Review on Evaluating Explainable AI. ACM Comput. Surv. 55, 13s, Article 295, 42 pages, 2023.

[3] Baptiste Abeloos and Stéphane Herbin. Sept 2022. Explaining object detectors: the case of transformer architectures. Workshop on Trustworthy Artificial Intelligence as a part of the ECML/PKDD, 2022.

Est-il possible d'envisager un travail en binôme ? **Non**

Méthodes à mettre en oeuvre :

- | | |
|---|---|
| <input type="checkbox"/> Recherche théorique | <input type="checkbox"/> Travail de synthèse |
| <input checked="" type="checkbox"/> Recherche appliquée | <input checked="" type="checkbox"/> Travail de documentation |
| <input checked="" type="checkbox"/> Recherche expérimentale | <input checked="" type="checkbox"/> Participation à une réalisation |

Possibilité de prolongation en thèse : **Oui**

Durée du stage : Minimum : 4

Maximum : 6

Période souhaitée : Mars-Aout 2024

PROFIL DU STAGIAIRE

Connaissances et niveau requis :

3^{ème} année d'école d'Ingénieur ou Master 2

Ecoles ou établissements souhaités :

GEN-F218-3