

# Détection de données multimodales hors distribution

## Stage M2 - 2025

**Mots-clés :** Incertitude, OOD, données multimodales

## 1 Contexte

Les modèles de *machine learning* (ML) sont souvent conçus pour des données de test similaires aux données d'entraînement, mais en pratique, des échantillons hors distribution (*Out of Distribution* – OOD) peuvent survenir, compromettant leur robustesse. La détection d'OOD est donc cruciale dans des domaines sensibles comme la conduite autonome et la santé. De nombreuses méthodes ont été proposées en vision par ordinateur pour la détection de données OOD reposant sur des méthodes de classification [2], de distance [3], d'estimation de densité [6] ou de reconstruction [5].

La recherche actuelle se concentre principalement sur des données unimodales qui sont très souvent des images naturelles. Certains travaux récents étudient les modèles combinant vision et langage pour améliorer les performances OOD mais s'évaluent sur des tâches purement visuelles. Ces modèles ne sont donc pas adaptés pour tirer pleinement parti des informations complémentaires issues de diverses modalités, comme des données LiDAR et caméra en conduite autonome ; des images échographiques, scanner et IRM en imagerie médicale ou encore des images multispectrales et radar en observation de la Terre.

Très récemment une nouvelle méthode étudiant la détection d'OOD en présence de différentes modalités a été proposé [1]. Cette approche part du constat que les prédictions seront relativement homogènes entre les différentes modalités lorsque les entrées suivent bien la distribution des données d'apprentissage. À l'inverse, sur des données OOD, les prédictions sont plus chaotiques permettant ainsi de les détecter.

## 2 Objectifs

Ce stage a pour but d'explorer l'adaptation de méthodes connues et efficaces pour la détection d'échantillons OOD dans des données mono-modale au contexte multimodal. La première partie de ce stage consistera à reprendre la méthode introduite dans [1] pour en reproduire les résultats sur le benchmark MultiOOD proposé. D'autres jeux de données issus de l'observation de la Terre tels que So2Sat ou AID pourront également être explorés. La seconde partie de ce stage visera à s'inspirer de stratégies de fusion récentes de détecteur d'OOD pour des données multimodales.

## 3 Profil

Le·a candidat·e doit être en Master en informatique, mathématiques appliquées ou traitement du signal. Le·a candidat·e doit avoir un fort intérêt pour la recherche scientifique et des bases théoriques en *machine learning*. La connaissance de la programmation en Python fortement

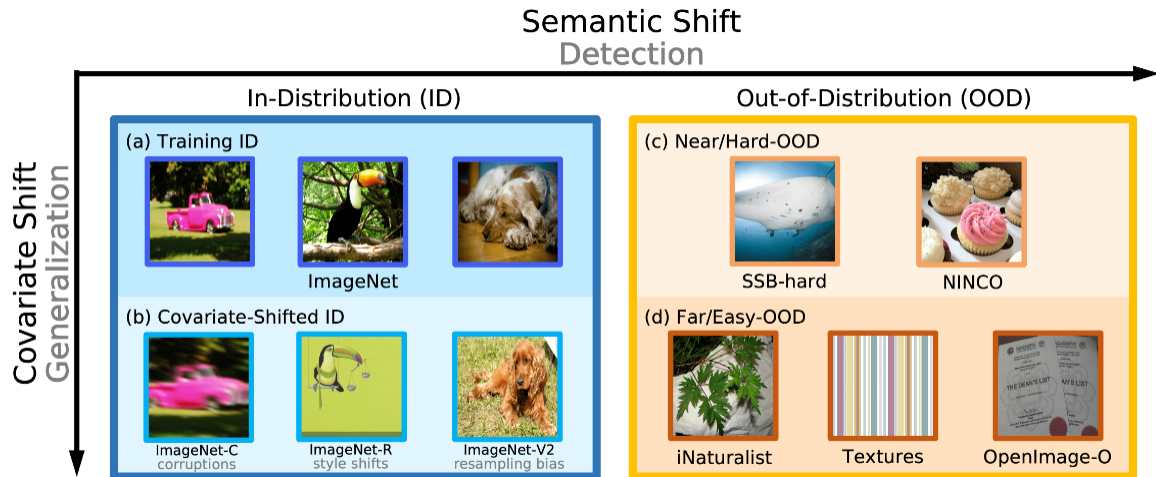


Figure 1: Illustration de la détection de données OOD issue de [4]. La détection standard se limite à détecter les décalages sémantiques en identifiant (c) + (d) à partir de (a). Dans un cadre étendu, on souhaite prendre en compte les transformations éventuelles des données et séparer (c) + (d) de (a) + (b).

conseillée ainsi qu’une première expérience avec une bibliothèque d’apprentissage profond telle que PyTorch est la bienvenue.

## 4 Organisation

Ce stage de 5 à 6 mois devrait débuter entre février et mars 2025. Le stage se déroulera à l’Institut des systèmes intelligents et robotique (ISIR) localisé au sein du campus Pierre et Marie Curie de Sorbonne Université. L’ISIR est un laboratoire de recherche pluridisciplinaire qui rassemble des chercheurs·ses et enseignant·e·s-chercheurs·ses, relevant des disciplines de la mécanique, de l’automatique, du traitement du signal et de l’informatique. Les travaux de recherche développés à l’ISIR portent sur de nombreux domaines d’application (apprentissage automatique, robotique, interface homme machine, santé, etc.). Ils sont en particulier tournés vers les applications émergentes de la robotique et des systèmes intelligents dans les domaines des sciences du vivant.

Le stage sera supervisé par Clément Rambour et Nicolas Thome de l’équipe *Machine Learning Deep Learning and Information Access* (MLIA).

## 5 Candidature

Envoyez un CV et un court message expliquant votre intérêt pour le stage à : [clement.rambour@sorbonne-universite.fr](mailto:clement.rambour@sorbonne-universite.fr)

## References

- [1] H. Dong, Y. Zhao, E. Chatzi, and O. Fink. MultiOOD: Scaling Out-of-Distribution Detection for Multiple Modalities, Oct. 2024.
- [2] D. Hendrycks and K. Gimpel. A Baseline for Detecting Misclassified and Out-of-Distribution Examples in Neural Networks. In *International Conference on Learning Representations*, Nov. 2016.

- 
- [3] K. Lee, K. Lee, H. Lee, and J. Shin. A Simple Unified Framework for Detecting Out-of-Distribution Samples and Adversarial Attacks. In *Advances in Neural Information Processing Systems*, volume 31, 2018.
  - [4] J. Yang, P. Wang, D. Zou, Z. Zhou, K. Ding, W.-H. Peng, H. Wang, G. Chen, B. Li, Y. Sun, X. Du, K. Zhou, W. Zhang, D. Hendrycks, Y. Li, and Z. Liu. Openood: Benchmarking generalized out-of-distribution detection. *ArXiv*, abs/2210.07242, 2022.
  - [5] Y. Zhou. Rethinking Reconstruction Autoencoder-Based Out-of-Distribution Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.
  - [6] E. Zisselman and A. Tamar. Deep Residual Flow for Out of Distribution Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.