

Approches auto-supervisées pour l'analyse tomographique de roches

à partir de Février 2025

Supervision : Dominique Fourer, Désiré Sidibé
Laboratoire : IBISC (EA 4526) - Univ. Évry/Paris-Sacay
Collaboration : IFPEN
Contact : dominique.fourer@univ-evry.fr and drodesire.sidibe@univ-evry.fr

Contexte : L'apprentissage auto-supervisé (self-supervised learning) est un paradigme émergent en apprentissage automatique permettant d'entraîner des modèles reposant sur les réseaux de neurones profonds, sans nécessiter de données étiquetées en grande quantité. En générant des labels artificiels à partir des données elles-mêmes, ces approches permettent de réduire la dépendance aux annotations manuelles, souvent coûteuses. Les approches auto-supervisées se montrent particulièrement efficaces dans des domaines tels que la vision par ordinateur [1], le traitement du langage naturel [2], et l'analyse des signaux [5]. Ainsi le travail attendu de ce stage vise à renforcer une étude existante menée avec l'IFPEN qui souhaite analyser automatiquement les caractéristiques de roches à partir d'images 2D ou de volumes 3D obtenus par forage du sol (cf. Fig. 1). Plus précisément, on souhaite estimer la porosité et la perméabilité du sol à partir d'observations du sol[4].

keywords : self-supervised learning, digital rock physics, curriculum learning, contrastive learning, tomographys

Objectifs

Le stagiaire participera à un projet de recherche visant à :

1. **Revue de l'état de l'art :** Étudier les approches auto-supervisées récentes, telles que SimCLR [1], BYOL [3], et BERT [2].
2. **Implémentation :** Reproduire et tester des algorithmes existants en utilisant des PyTorch ou TensorFlow.
3. **Innovation :** Proposer et expérimenter des variantes ou extensions des méthodes existantes.
4. **Évaluation :** Comparer les performances des méthodes auto-supervisées avec des approches supervisées et semi-supervisées sur des jeux de données spécifiques.
5. **Valorisation :** Documenter les résultats, rédiger un rapport scientifique, et contribuer à la préparation d'articles pour des conférences ou journaux.

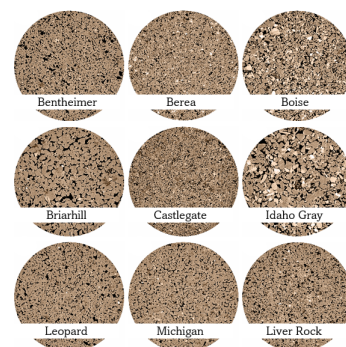


Fig. 1 : Illustration de roche 2D extraite à partir de volume tomographique 3D microCT.

Profil recherché

- Connaissances solides en apprentissage profond (ex. CNN, Transformers) et en optimisation.
- Maîtrise de Python et des bibliothèques (PyTorch, Keras, TensorFlow, etc.).
- Connaissance des concepts "self-supervised" et "contrastive learning" sont un atout
- Autonomie, curiosité scientifique, et rigueur.

Encadrement et environnement de travail

Le stage se déroulera au laboratoire IBISC site Pelvoux, sous la supervision de Dominique Fourer et Désiré Sidibé. Le stagiaire bénéficiera d'un environnement de recherche dynamique et de ressources de calcul avancées (GPU).

Références

- [1] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International Conference on Machine Learning*, pages 1597–1607. PMLR, 2020.
- [2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert : Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv :1810.04805*, 2018.
- [3] Jean-Bastien Grill, Florian Strub, Florent Altché, Corentin Tallec, Pierre H Richemond, Elena Buchatskaya, Carl Doersch, Bernardo Avila Pires, Zhaohan Daniel Guo, Mohammad Gheshlaghi Azar, et al. Bootstrap your own latent : A new approach to self-supervised learning. *Advances in Neural Information Processing Systems*, 33 :21271–21284, 2020.
- [4] Van Thao Nguyen, Dominique Fourer, Désiré Sidibé, Jean-François Lecomte, and Souhail Youssef. A comparative evaluation of self-supervised methods applied to rock images classification. In *19th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP 2023)*, 2024.
- [5] Steffen Schneider, Alexei Baevski, Ronan Collobert, Michael Auli, and Sergey Edunov. Wav2vec : Unsupervised pre-training for speech recognition. In *INTERSPEECH*, pages 3465–3469, 2019.