

Laboratoire d'Informatique
Fondamentale et Appliquée
64 avenue Jean Portalis
37200 TOURS

lifat.univ-tours.fr

Équipe Reconnaissance des Formes et
Analyse d'Images

www.rfai.li.univ-tours.fr



Proposition de stage BAC+5 en IA

au LIFAT (Tours) et en collaboration avec le BRGM (Orléans)

Disciplines : informatique, géophysique

Développement d'un système d'aide au diagnostic des digues à partir de données géophysiques et géomécaniques

Employeur : Université d'Orléans (RTR DIAMS)

Grade : Stagiaire 5^{ème} année (M2 ou ingénieur)

Durée : 6 mois

Rémunération : suivant les grilles effectives au niveau national pour les stages supérieurs à 2,5 mois

Localisation : Le stagiaire sera rattaché au LIFAT

Interlocuteurs : LIFAT : Nicolas RAGOT ; BRGM : Cécile Gracianne, Kessouri Pauline, Deparis Jacques

Qualifications souhaitées : Master en cours dans le domaine de l'informatique avec des compétences en *data sciences* (analyse de données, calcul scientifique, intelligence artificielle).

Compétences requises :

- Bases scientifiques et mathématiques solides
- Programmation (Python, C++ ou Java)
- La connaissance d'au moins un environnement de développement scientifique (Pytorch, scikit-learn, Matlab, ...)
- Des connaissances en géophysique seront appréciées
- Qualités relationnelles, ouverture et curiosité afin de dialoguer et comprendre les interlocuteurs de champs disciplinaires spécifiques
- Sens de l'initiative et force de proposition
- Sens de l'organisation, autonomie
- Capacité à faire du *reporting*

Candidatures :

par courrier électronique avant le 20/01/2021 adressées à :

nicolas.ragot@univ-tours.fr ; c.gracianne@brgm.fr ; p.kessouri@brgm.fr ;
j.deparis@brgm.fr



Description du sujet

La France compte environ 9000 km de digues, de géométrie et hauteur variable bordant des canaux artificiels. Dans le cadre des articles L. 214-1 et L. 214-2 du code de l'environnement, les propriétaires de digue se doivent de surveiller l'intégrité de ces ouvrages en terre, qui peuvent être à l'origine de fuites et d'inondations. Par ailleurs, des instabilités des matériaux la constituant ou au niveau de l'encaissant géologique peuvent entraîner des ruptures totales de la digue, pouvant générer des inondations catastrophiques pour les personnes et les biens. Cette surveillance est réalisée traditionnellement par des auscultations visuelles de site et par des sondages géotechniques. Dans le premier cas, cette analyse n'est que de surface, limitant ainsi le coût de l'opération et permettant d'enclencher des alertes rapidement. Ce type d'analyse nécessite toutefois que les dommages soient apparents.

Les méthodes géophysiques sont des méthodes d'auscultation non intrusives qui permettent d'imager les variations de paramètres géophysiques dans le sous-sol, eux-mêmes liés à des paramètres mécaniques, hydrogéologiques et/ou géochimiques des formations en présence. Les méthodes géophysiques peuvent ainsi renseigner sur :

- La nature des matériaux composant la digue et l'encaissant géologique sous les digues ;
- La présence d'hétérogénéité de matériaux au sein d'un ensemble homogène dans la digue et son encaissant ;
- La variation de la teneur en eau dans les digues ;
- La présence de cavités dans le substratum rocheux sous la digue ;
- La rigidité et l'élasticité des matériaux composant la digue.

Depuis le début des années 2000, le BRGM a développé des outils haut rendement afin d'optimiser les coûts d'acquisitions et de traitements de ces méthodes. L'approche proposée par le BRGM repose sur le couplage de deux types de méthodes géophysiques à grand rendement, parfaitement adaptées à la caractérisation de grands linéaires. Il permet ainsi d'imager les paramètres physiques suivants :

- Les vitesses sismiques par analyse des ondes de surface et des premières arrivées.
- La résistivité électrique par méthodes capacitives ou électromagnétiques.

L'interprétation de ces données géophysiques doit intégrer des objets d'échelles variables de quelques mètres à quelques centaines de mètres sur des profils pouvant atteindre plusieurs kilomètres. De plus, les données géophysiques doivent être interprétées en les couplant avec des données géologiques issues de cartes et de sondages destructifs ponctuels. L'interprétation est donc délicate à réaliser, longue, subjective et dépend de l'expert.

Le BRGM s'est déjà doté d'un outil d'analyse automatique des données pour l'interprétation conjointe de la résistivité électrique et des vitesses sismiques (Samyn et al, 2014) mais son utilisation est limitée à la détection d'anomalies géophysiques dans le corps de digue ou dans les terrains sous-jacents dus à la présence de karst sur les levés de Loire. Depuis quelques années, des études montrent qu'il est possible d'utiliser l'intelligence artificielle via des algorithmes d'apprentissage en géoscience. Ce type de méthode commence à être répandue pour la sismique réflexion, par exemple en améliorant les processus de traitement des données (Jia et Ma, 2017), en analyse de faciès (Wrona et al, 2018) ou structurale (Fehmers et Höcker, 2003). Des études récentes utilisent ce type d'approche pour l'interprétation des données issues de différentes thématiques. Par exemple, Abedi et al. (2012) utilise une méthode de classification supervisée appelé SVM pour explorer les gisements de porphyry-CU. Il utilise à la fois la carte géologique, des données géophysiques et géochimiques pour évaluer le gisement. Cracknell et Reading (2014), quant à eux, utilisent des algorithmes d'apprentissage automatique pour reconstruire la carte géologique à partir de données de géophysique aéroportée (Gamma spectrométrie et magnétisme), et des données satellitaires multi-spectrales. Ces études montrent qu'il est possible d'analyser conjointement des données issues d'acquisitions différentes à l'aide d'algorithme d'apprentissage.

L'objectif de ce stage est d'explorer le potentiel des approches d'intelligence artificielle dont l'apprentissage automatique pour traiter ces données et faciliter leur analyse. On s'intéressera en particulier aux liens existants entre les différentes mesures géophysiques réalisées sur les digues entre elles, ainsi qu'avec les autres types de données géologiques complémentaires. Au moins deux tâches sont ainsi envisagées à partir de ces données multivariées : la recherche de motifs caractéristiques d'éléments de structure ou d'anomalies et la classification. La première tâche peut être abordée par exemple grâce à l'usage de distances élastiques entre courbes comme le DTW (Albrecht, 2007 ; Mondal et al., 2016), en considérant les données comme une séquence dans une



direction donnée (profondeur, axe longitudinal de la digue).

L'avantage est que cette approche ne nécessite pas ou peu d'apprentissage. Alternativement si les données le permettent (et notamment la quantité de données annotées), on pourra envisager d'utiliser des approches d'apprentissage machine comme les CNN (Kiranyaz et al., 2019, LeCun et al., 1998). Pour la tâche de classification, là encore en fonction des données, plusieurs approches seront étudiées (SVM, Réseaux de neurones profonds notamment). Cette étude permettra de déterminer la pertinence de la construction d'un modèle de classification des digues, et potentiellement la prédiction d'anomalies.

La description détaillée de cette analyse, les propositions et recommandations éventuelles qui peuvent en découler sur la pertinence des approches de *machine learning* vis-à-vis de l'objectif visé, sont une première contribution, qui est une des attentes principales du projet. En fonction du temps disponible, une deuxième contribution attendue est de participer au développement d'un modèle de prédiction de la décompaction du sol, en s'appuyant sur cette analyse. Pour réaliser ce travail, l'étudiant aura à sa disposition les données obtenues lors des campagnes d'acquisitions pertinentes réalisées par le BRGM depuis les années 2000.

Références

- Abedi, M., Torabi, S. A., Norouzi, G. H., & Hamzeh, M. (2012). ELECTRE III: A knowledge-driven method for integration of geophysical data with geological and geochemical data in mineral prospectivity mapping. *Journal of applied geophysics*, 87, 9-18.
- Albrecht, T. (2007). Dynamic Time Warping (DTW). In M. Müller (Ed.), *Information Retrieval for Music and Motion* (pp. 69–84). Springer.
- Cracknell, M. J. & Reading, A. M. (2014). Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput. Geosci.* 63, 22–33).
- Fehmers, G. C., & Höcker, C. F. (2003). Fast structural interpretation with structure-oriented filtering. *Geophysics*, 68(4), 1286-1293.
- Kanevski, M., Timonin, V., & Pozdnukhov, A. (2009). *Machine learning for spatial environmental data: theory, applications, and software*. EPFL press.
- S. Kiranyaz, T. Ince, O. Abdeljaber, O. Avci and M. Gabbouj, "1-D Convolutional Neural Networks for Signal Processing Applications," *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, 2019, pp. 8360-8364, doi: 10.1109/ICASSP.2019.8682194.
- Jia, Y., & Ma, J. (2017). What can machine learning do for seismic data processing? An interpolation application. *Geophysics*, 82(3), V163-V177.
- Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278-2324, 1998.
- Mondal, T., Ragot, N., Ramel, J.-Y., & Pal, U. (2016). Flexible Sequence Matching Technique: An Effective Learning-free Approach For word-spotting. *Pattern Recognition*, 60, 596–612. <https://doi.org/10.1016/j.patcog.2016.05.011>
- Samyn, K., Mathieu, F., Bitri, A., Nachbaur, A., & Closset, L. (2014). Integrated geophysical approach in assessing karst presence and sinkhole susceptibility along flood-protection dykes of the Loire River, Orléans, France. *Engineering geology*, 183, 170-184.
- Wrona, T., Pan, I., Gawthorpe, R. L., & Fossen, H. (2018). Seismic facies analysis using machine learning. *Geophysics*,