OFFRE DE STAGE DE MASTER 2

Machine Learning, Deep Learning, Traitement d'images

Intitulé du stage :

Détection et classification automatiques de potentielles traces fossiles par Deep Learning : fusion de données photogrammétriques 2D et 3D

Automatic Detection and Classification of Potential Fossil Traces Using Deep Learning: Fusion of 2D and 3D Photogrammetric Data

Durée : 5 à 6 mois, de février à juillet 2027
Gratification : Selon la réglementation en vigueur
Laboratoire d'accueil : Laboratoire d'Océanologie et de Géosciences (LOG – UMR 8187 CNRS, ULCO, Université de Lille, IRD), Université du Littoral Côte d'Opale, Dunkerque.

Encadrement scientifique
Le stage sera réalisé dans le cadre d'une collaboration entre des chercheurs en géosciences et en intelligence artificielle :
- **Mouncef Sedrati** < mouncef.sedrati@univ-littoral.fr > – Professeur des Universités, LOG, ULCO - Géomorphologie, analyse des surfaces fossiles et morphométrie.
- **Ayoub Karine** < ayoub.karine@u-pariscite.fr > – Maître de Conférences, SIP/LIPADE, Université Paris Cité – Intelligence artificielle et traitement d'images.

Le co-encadrement permettra d'associer le développement de méthodes de vision par ordinateur aux questions scientifiques liées à l'identification, à la classification et à la caractérisation des structures fossiles.

Candidature : CV et lettre de candidature à transmettre aux deux responsables scientifiques.

Contexte scientifique

Un ensemble de structures circulaires et subcirculaires a récemment été découvert sur une plateforme rocheuse du littoral (Figure 1). Ces structures, dont certaines atteignent plusieurs dizaines de centimètres de diamètre, présentent des morphologies variées pouvant évoquer des impressions fossilisées d'organismes à corps mou, notamment de méduses. Leur origine reste toutefois à déterminer.

Les premières observations de terrain ont permis de distinguer trois principaux morphotypes (Figure 2) : des formes circulaires présentant une bordure denticulée associée à de petites dépressions périphériques (Figure 2, A), des structures arrondies comportant un couvercle ou remplissage sédimentaire apparent (Figure 2, B) et des dépressions circulaires à subcirculaires plus ou moins régulières (Figure 2, C). Leur nombre, leur répartition sur l'affleurement et leurs caractéristiques morphométriques restent à préciser. Certaines structures sont jointives, partiellement superposées ou dégradées par l'érosion, ce qui complique leur identification et leur délimitation.



Figure 1 : Image aérienne (à partir du levé drone du site) qui montre la diversité des formes fossiles.

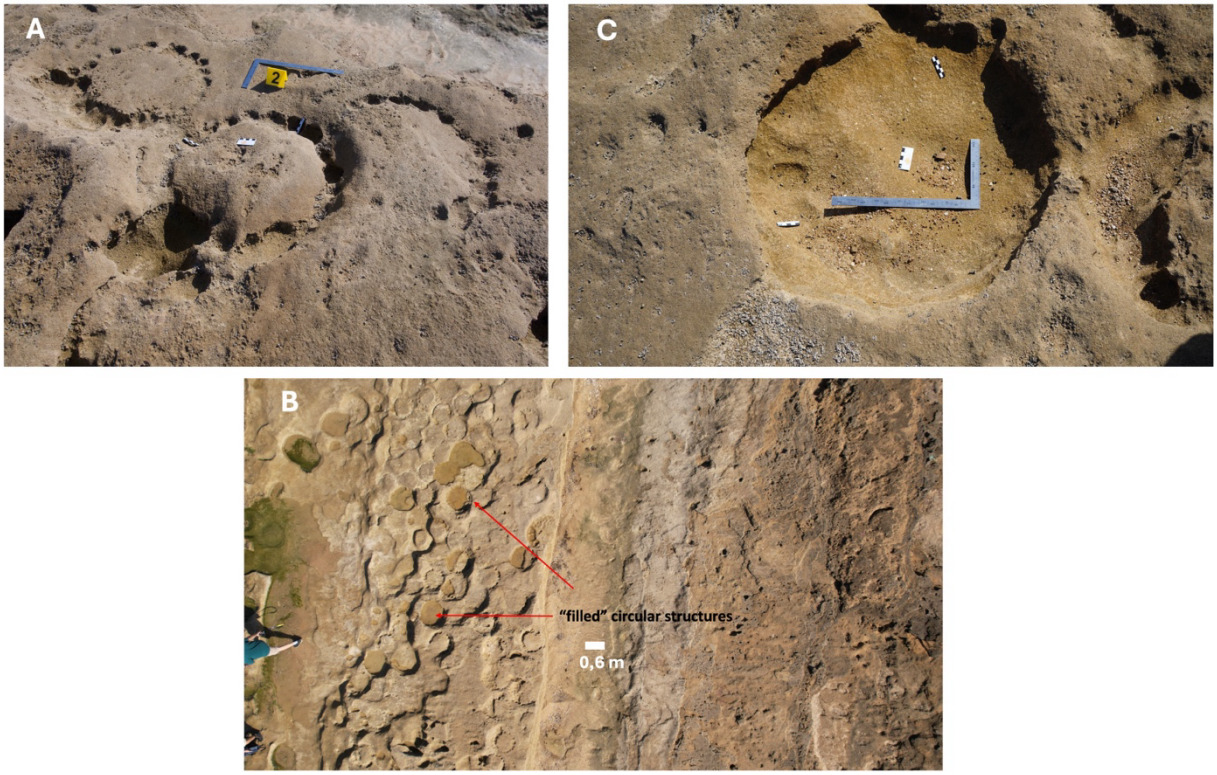


Figure 2 : Photographies des 3 principaux morphotypes de traces fossilisées : (A) des formes circulaires présentant une bordure denticulée associée à de petites dépressions périphériques, (B) des structures arrondies comportant un couvercle ou remplissage sédimentaire apparent et (C) des dépressions circulaires à subcirculaires plus ou moins régulières.

Un levé photogrammétrique haute résolution du site a permis de produire une orthophotographie et un modèle numérique d'élévation (MNE). Ces données offrent la possibilité de réaliser une cartographie détaillée des structures et d'en analyser la géométrie tridimensionnelle.

Les développements récents du Deep Learning ont démontré leur intérêt pour la reconnaissance et la classification d'empreintes fossiles, notamment de dinosaures (Lallensack et al., 2022 ; Ha et Kim, 2026

; Hartmann et al., 2026). Leur application à des surfaces naturelles présentant des formes irrégulières, parfois peu contrastées et difficiles à individualiser, reste néanmoins un défi. L'exploitation conjointe des informations photographiques et topographiques pourrait améliorer leur détection et leur caractérisation.

Objectifs du stage

Le stage vise à développer et évaluer une méthode automatisée de détection, de segmentation et de classification des structures du site de Larache, en combinant les informations issues de l'orthophotographie et du MNE.

Le travail portera principalement sur :

- **La détection et le comptage des structures**, avec le développement d'une procédure capable de les localiser et de les individualiser sur l'ensemble de la surface étudiée, y compris lorsqu'elles sont proches ou partiellement superposées.
- **La classification morphologique**, afin de distinguer les trois morphotypes identifiés sur le terrain et de caractériser les formes intermédiaires ou difficiles à attribuer.
- **L'exploitation conjointe des données 2D et 3D**, en comparant les performances des modèles utilisant les images RGB, les paramètres topographiques dérivés du MNE et leur combinaison.
- **L'extraction automatique des paramètres morphométriques** : diamètre, longueur, largeur, surface, circularité, profondeur, relief périphérique et volume relatif des structures.
- **L'analyse de leur distribution spatiale**, à partir d'un inventaire géoréférencé permettant d'examiner les variations de taille, de densité et de morphologie sur l'affleurement.

Méthodologie envisagée

L'étudiant disposera des données photogrammétriques du site ainsi que des observations morphologiques réalisées sur le terrain. Une première étape consistera à préparer les données et à construire un jeu d'apprentissage à partir de structures représentatives des différents morphotypes. Des outils de segmentation assistée, tels que Segment Anything Model (SAM), pourront être utilisés pour faciliter leurs annotations.

Des architectures de Deep Learning adaptées à la détection et à la segmentation d'instances (basées sur les convolutions ou les Transformers) seront ensuite testées. L'une des principales questions méthodologiques sera d'évaluer l'intérêt des informations topographiques pour améliorer la reconnaissance des structures. Le MNE permettra notamment de calculer des représentations du relief local, des pentes et des courbures, qui pourront être associées aux images RGB dans une approche multimodale.

Les performances des modèles seront comparées sur des secteurs indépendants des données d'entraînement. La qualité de la détection, de la segmentation et de la classification sera évaluée à partir d'indicateurs quantitatifs (précision, rappel, mAP, Intersection over Union). Une attention particulière sera portée aux formes partiellement conservées et aux confusions possibles avec les irrégularités naturelles du substrat rocheux.

Les contours obtenus seront ensuite exploités pour développer une extraction automatisée des paramètres morphométriques, en utilisant les informations altimétriques du MNE. Des méthodes statistiques et de Machine Learning pourront compléter l'analyse afin d'examiner la variabilité des formes et la pertinence de leur regroupement en trois morphotypes.

La méthode retenue sera appliquée à l'ensemble de l'affleurement. Les résultats seront intégrés dans un système d'information géographique pour produire un inventaire spatial et morphométrique des structures.

Résultats attendus

Le stage devra permettre de disposer d'une chaîne de traitement reproductible, associant détection, segmentation, classification et extraction des paramètres morphométriques.

Les principaux livrables/résultats comprendront une cartographie des structures identifiées, une base de données géoréférencée contenant leurs caractéristiques géométriques et topographiques, ainsi qu'une évaluation comparative des performances des modèles testés. Le modèle d'apprentissage et une interface web avec import d'images et prédictions sont attendus également.

L'analyse fournira les premiers éléments quantitatifs permettant de comparer les morphotypes et d'examiner leur organisation sur l'affleurement. Les méthodes développées pourront également être adaptées à d'autres surfaces fossilifères ou archéologiques présentant des objets morphologiques complexes.

Profil recherché

Le stage s'adresse à un(e) étudiant(e) de Master 2 ou en dernière année de cycle d'ingénieur en Intelligence artificielle, Vision par ordinateur ou Traitement d'images avec un intérêt particulier aux géosciences.

Une bonne maîtrise de Python et des bibliothèques de Deep Learning (PyTorch notamment) est attendue. Des compétences en segmentation d'images, reconnaissance de formes, apprentissage multimodal ou traitement de données 3D seront particulièrement appréciées.

Le sujet conviendra à un(e) étudiant(e) souhaitant appliquer et adapter des méthodes de vision par ordinateur à des données scientifiques réelles, dans le cadre d'une collaboration entre spécialistes de l'intelligence artificielle et des géosciences.

Aucune connaissance préalable en géologie ou en paléontologie n'est requise. Les éléments scientifiques nécessaires à la compréhension du site seront apportés par l'équipe encadrante.

Quelques références bibliographiques

Budka, M., Bennett, M.R., Reynolds, S.C., Barefoot, S., Reel, S., Reidy, S., & Walker, J. (2021). Sexing white 2D footprints using convolutional neural networks. *PLOS ONE*, 16(8), e0255630. https://doi.org/10.1371/journal.pone.0255630

Ha, Y., & Kim, S.-S. (2026). Deep learning enables automated detection of dinosaur footprints with high accuracy. *Scientific Reports*, 16, 25415. https://doi.org/10.1038/s41598-026-56031-5

Hartmann, G., Blakesley, T., dePolo, P.E., & Brusatte, S.L. (2026). Identifying variation in dinosaur footprints and classifying problematic specimens via unbiased unsupervised machine learning. *Proceedings of the National Academy of Sciences*, 123(5), e2527222122. https://doi.org/10.1073/pnas.2527222122

Karine, A., Napoléon, T., & Jridi, M. (2025). I2CKD: Intra- and inter-class knowledge distillation for semantic segmentation. *Neurocomputing*, 649, 130791. https://doi.org/10.1016/j.neucom.2025.130791

Lallensack, J.N., Romilio, A., & Falkingham, P.L. (2022). A machine learning approach for the discrimination of theropod and ornithischian dinosaur tracks. *Journal of the Royal Society Interface*, 19, 20220588. https://doi.org/10.1098/rsif.2022.0588

Marques, C.S., Mota, A., Belvedere, M., Castanera, D., Díaz-Martínez, I., Malafaia, E., Pereira, S., Rosalino, L.M., Santos, V.F., Sciscio, L., & Dufourq, E. (2026). Deep tracks: Using deep learning and procedurally simulated data for automated vertebrate footprints classification. *Ecological Informatics*, 93, 103523. https://doi.org/10.1016/j.ecoinf.2025.103523

Sedrati, M., Morales, J.A., Duveau, J., El M'rini, A., Mayoral, E., Díaz-Martínez, I., Anthony, E.J., Bulot, G., Sedrati, A., Le Gall, R., Santos, A., & Rivera-Silva, J. (2024). A Late Pleistocene hominin footprint site on the North African coast of Morocco. *Scientific Reports*, 14, 1962. https://doi.org/10.1038/s41598-024-52344-5