
Classification automatique des diatomées par apprentissage profond pour l'amélioration du diagnostic écologique des milieux aquatiques

Pierre Faure-Giovagnoli¹, Souhila Founas², David Heudre³, Thibault De Garidel-Thoron⁴, Philippe Usseglio-Polatera^{5,6}, Cédric Pradalier^{7,6}, and Martin Laviale^{*6,8}

¹GeorgiaTech Lorraine-UMI 2958 – GeorgiaTech, CNRS : UMI2958 – France

²Laboratoire Interdisciplinaire des Environnements Continentaux – CNRS : UMR7360, Université de Lorraine – France

³DREAL Grand Est – Ministère de l'écologie de l'Energie, du Développement durable et de l'Aménagement du territoire – France

⁴Centre européen de recherche et d'enseignement de géosciences de l'environnement (CEREGE) – Aix Marseille Université, INSU, Institut de recherche pour le développement [IRD], CNRS : UMR7330 – Europôle Méditerranéen de l'Arbois - Avenue Louis Philibert - BP 80 - 13545 Aix-en-Provence cedex 4, France

⁵Laboratoire Interdisciplinaire des Environnements Continentaux (LIEC) – Université de Lorraine, Centre National de la Recherche Scientifique : UMR7360 – Campus Bridoux - Rue du Général Delestraint, Université de Lorraine F-57070 METZ - France, France

⁶LTSER-ZAM – LTSER – France

⁷Georgia Tech - CNRS [Metz] (UMI2958) – Georgia Institute of Technology [Atlanta], Georgia Tech Lorraine, Centre National de la Recherche Scientifique : UMI2958 – Metz Technopôle 2-3 rue Marconi 57070 METZ, France

⁸Laboratoire Interdisciplinaire des Environnements Continentaux (LIEC) – CNRS : UMR7360, Université de Lorraine – France

Résumé

Les diatomées sont des microalgues présentes dans tous les milieux aquatiques. Ces organismes sont utilisés en routine comme bioindicateur de la qualité écologique des eaux douces dans le cadre de la mise en œuvre de la Directive Cadre européenne sur l'Eau (DCE). Les indices biologiques actuels basés sur les diatomées reposent sur des critères morphologiques (forme et ornementation de l'exosquLETTE siliceux, le frustule) pas toujours faciles à caractériser en routine (i.e. à l'aide de méthodes optiques conventionnelles). L'identification est donc chronophage, souvent sujette à de multiples biais (expérience de l'opérateur, qualité de l'image) et nécessite un niveau élevé d'expertise.

Ceci justifie le développement d'un outil plus robuste, basé sur une classification automatique des diatomées. Cet objectif est toujours un défi d'actualité, depuis les premières tentatives datant des années 90. Dans ce contexte, le développement récent des approches

*Intervenant

d'apprentissage profond pour identifier et quantifier les traits des organismes à partir d'images semble prometteur pour résoudre les problèmes rencontrés jusqu'à présent. Notre objectif est donc de proposer un nouvel outil d'identification des diatomées basé sur des algorithmes de reconnaissance automatique de formes à partir d'images individuelles. Par rapport à l'approche classique d'identification des diatomées, cet outil se veut notamment plus robuste car indépendant de l'opérateur et permettra d'améliorer (gain de temps, coût) ceux actuellement disponibles dans le cadre du suivi réglementaire de l'état écologique des cours d'eau. Le développement de cet outil implique 1/l'acquisition d'une base de données représentative du milieu naturel (banques d'images de diatomées) qui permette 2) le développement des algorithmes d'identification des diatomées. En première approche, une banque simplifiée d'images représentatives de 209 espèces de diatomées (environ 9 000 images au total) a été créée à partir de guides d'identification en libre accès. Cette base d'images a permis de générer environ 30 000 images composites simulant des images d'échantillons naturels. Ceci a permis le développement de premiers algorithmes de détection (diatomées vs. débris) et de classification (espèces présentes). Les premiers résultats obtenus montrent que l'outil de reconnaissance permet de distinguer les diatomées présentes sur un échantillon (précision de 90%) ainsi que les débris (précision de 99%). L'outil a également été testé sur quelques images réelles avec une précision pouvant atteindre 86% pour les débris et 73% pour les diatomées présentes. Ces résultats encourageants démontrent la faisabilité de notre approche (preuve de concept). La banque d'images sera donc consolidée afin de la rendre plus représentative de l'ensemble des espèces indicatrices de diatomées retrouvées dans une zone géographique donnée (e.g. bassin de la Moselle, bassin Rhin-Meuse) et ainsi d'améliorer la performance des algorithmes de reconnaissance.

Mots-Clés: qualité de l'eau, bioindication, diatomées, analyse d'images, reconnaissance automatique, apprentissage machine