



HAL
open science

Océans et intelligence artificielle, un couple d'avenir

Sakina-Dorothee Ayata

► **To cite this version:**

Sakina-Dorothee Ayata. Océans et intelligence artificielle, un couple d'avenir. Christophe Prazuck. 30 questions sur l'océan: La Sorbonne et la mer, Sorbonne Université Presses, 2022, Sorbonne Essais, 979-10-231-0734-0. hal-03993688

HAL Id: hal-03993688

<https://hal.science/hal-03993688>

Submitted on 17 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Océans et intelligence artificielle, un couple d'avenir

Par Sakina-Dorothee Ayata,

Maitresse de conférences en écologie marine à Sorbonne Université au laboratoire LOCEAN à Paris.

Contact : sakina-dorothee.ayata@sorbonne-universite.fr

Article rédigé le 14 janvier 2022.

Référence : Sakina-Dorothee Ayata. (2022) Océans et intelligence artificielle, un couple d'avenir. *30 questions sur l'océan: La Sorbonne et la mer*. Sous la direction de Christophe Prazuck. Sorbonne, 2022, Sorbonne Essais, 979-10-231-0734-0

Texte de 6515 caractères, espaces compris.

L'intelligence artificielle, parfois appelée "IA", est un champ disciplinaire qui a le vent en poupe. Elle regroupe tout un ensemble de méthodes et techniques permettant de construire des machines "douées d'intelligence", ou tout du moins capables de faire preuve d'un comportement que l'on pourrait qualifier "d'intelligent". On peut ainsi développer des ordinateurs capables de jouer aux échecs et de défier les meilleurs joueurs du monde, construire des voitures capables de se conduire elles-mêmes, ou encore fabriquer des robots autonomes chargés de l'inspection et de la maintenance des coques de navires¹.

Mais l'intelligence artificielle est aussi au centre de nombreuses recherches en océanographie, et son rôle croissant n'est pas près de diminuer avec l'arrivée de données massives (ou "*big data*") qui nous renseignent sur l'état des océans. A l'interface entre intelligence artificielle et sciences des données, se trouve l'apprentissage automatique (ou "*machine learning*")². L'apprentissage automatique permet en particulier de résoudre trois grands types de problèmes : 1) des problèmes que l'on ne sait pas résoudre, comme prédire à partir de données de surface enregistrées par satellite la composition de l'écosystème marin en profondeur, 2) des problèmes que l'on sait résoudre sans trop savoir comment on fait pour y parvenir, comme reconnaître un poisson sur une photo, et enfin 3) des problèmes que l'on saurait en théorie résoudre mais qui demanderaient trop de ressources informatiques pour y répondre, comme simuler la turbulence à l'échelle de l'océan global. Ainsi, l'apprentissage automatique est l'outil idéal pour traiter les très grands jeux de données marines : données satellites, suivi GPS des navires, suivi automatique des écosystèmes par imagerie quantitative, enregistrements d'acoustique passive et active, etc.

En effet, les océans sont aujourd'hui scrutés en permanence et les quantités de données marines s'accumulent à vitesse grand V. D'abord par des satellites qui peuvent mesurer depuis l'espace la température, la salinité, les courants de surface ou la couleur de l'eau. Puis par des robots autonomes, comme les flotteurs-profileurs qui parcourent les océans de la surface vers le fond pour en mesurer les propriétés physico-chimiques (température, salinité, lumière) et

¹ Pour en savoir plus, voir le projet européen BugWright2 piloté par Cédric Pradalier de GoergiaTech Lorraine: <https://www.bugwright2.eu/>

² Les étudiantes et étudiants en informatique ou maths appliquées, en L3, master ou école d'ingénieurs, ainsi que les lectrices et lecteurs curieux et qui souhaitent en savoir plus pourront se référer au livre "Introduction au Machine Learning - 2e édition" de Chloé-Agathe Azencott publié aux Editions Dunod et dont une version sans les exercices est disponible gratuitement en ligne : http://cazencott.info/dotclear/public/lectures/IntroML_Azencott.pdf

biogéochimiques (fluorescence, concentration en nutriments) des masses d'eau, ou encore ceux qui enregistrent des vidéos des fonds marins. Les organismes qui vivent dans les océans sont aussi surveillés de près. L'observation des écosystèmes marins génère donc, elle aussi, de très grands jeux de données, qu'il s'agisse de données sur le blanchiment des coraux, obtenues par satellites, de la présence de bancs de poissons, estimée par les sonars des navires, ou encore de la morphologie du plancton enregistrée par des caméras miniatures.

On distingue classiquement deux principaux types d'apprentissage automatique. Le premier type est appelé apprentissage non supervisé : l'ordinateur apprend alors tout seul à partir des données. De tels algorithmes permettent par exemple d'identifier les zones de l'océan qui présentent les mêmes variations de chlorophylle mesurées par satellites, ou de regrouper des échantillons qui ont des compositions similaires en organismes marins. Le deuxième type de méthode est appelé apprentissage supervisé : pour cela, l'ordinateur doit s'entraîner pour apprendre à réaliser une tâche particulière, comme identifier des poissons puis reconnaître les espèces sur les vidéos enregistrées par les caméras situées à bord des bateaux de pêche, ou identifier la signature acoustique d'une espèce d'intérêt économique.

Grâce à l'utilisation de systèmes d'imagerie automatique, en particulier ceux développés au Laboratoire d'Océanographie de Villefranche-sur-Mer (LOV, SU/CNRS) près de Nice, plus de 160 millions d'images individuelles d'organismes planctoniques ont pu être collectées à travers l'ensemble des océans du globe. Grâce à des méthodes d'apprentissage supervisé, comme les réseaux de neurones profonds, un algorithme de reconnaissance automatique a pu être entraîné et plus de 70 millions d'images ont pu être classées au sein de différents groupes taxonomiques. Au-delà de l'amélioration de la classification automatique de ces images, les recherches actuellement menées, en particulier à Sorbonne Université, visent à développer de nouveaux algorithmes frugaux, c'est-à-dire moins consommateurs d'énergie et de mémoire afin d'équiper de manière embarquée des robots autonomes, mais aussi à identifier - indépendamment des espèces - les caractéristiques individuelles des organismes planctoniques qui font leur succès écologique, comme par exemple la présence de réserves lipidiques ou le taux de reproduction estimé grâce à l'identification automatique de la présence d'œufs (cf. figure).

Enfin, le développement d'outils d'intelligence artificielle dit "hybrides", c'est-à-dire couplés à des modèles numériques, offrent de nouvelles perspectives pour modéliser la physique des océans, et ainsi produire des "jumeaux numériques", mais aussi pour mieux représenter la dynamique des écosystèmes marins. C'est par exemple ce à quoi s'évertuera le projet Smart-Biodiv (Technologies d'Intelligence Artificielle pour la recherche en biodiversité) qui vient d'être sélectionné dans le cadre du challenge IA-Biodiversité de l'Agence Nationale de la Recherche (ANR)³.

Océans et intelligence artificielle forment donc un couple d'avenir en particulier pour quantifier l'impact du changement climatique sur les océans à partir des observations et des modèles, mais aussi pour mieux informer les décideurs, par exemple pour définir des quotas de pêche durable ou mettre en place et suivre des aires marines protégées et ainsi préserver la biodiversité marine.

³ <https://anr.fr/fr/detail/call/challenge-ia-biodiv-recherche-en-intelligence-artificielle-dans-le-champ-de-la-biodiversite/>

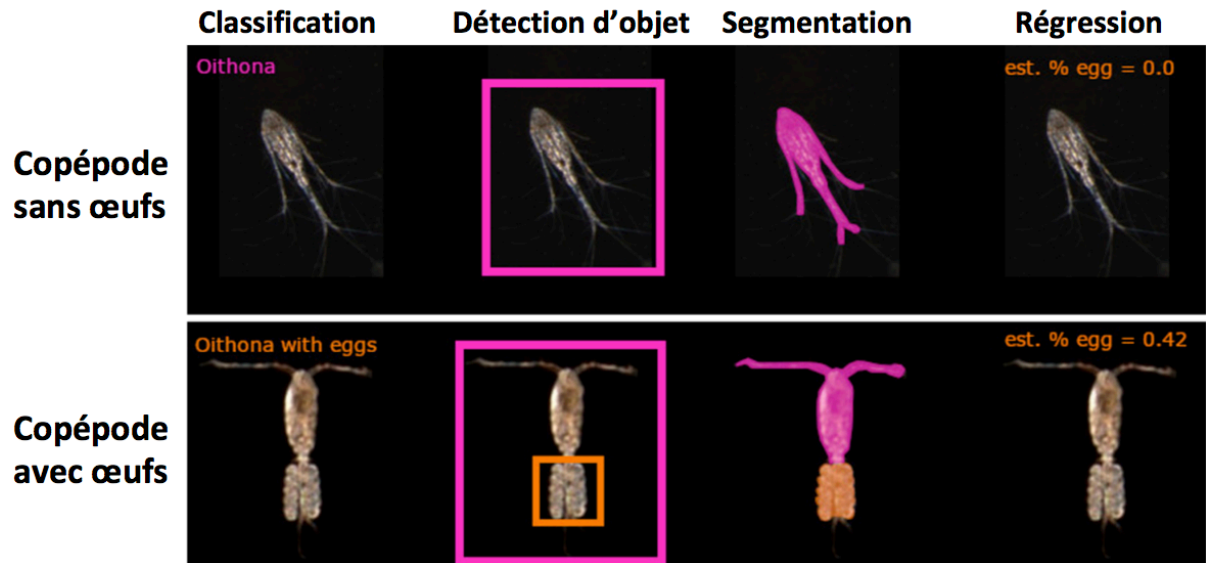


Figure : Exemples de méthodes d'apprentissage automatique (supervisé) pour apprendre à l'ordinateur à reconnaître automatiquement la présence d'œufs sur des images de copépodes. Les copépodes sont de petits crustacés planctoniques de quelques millimètres qui sont probablement les organismes les plus abondants de notre planète et qui forment un maillon clef des réseaux trophiques marins entre les producteurs primaires et les poissons. La classification permet de ranger les images dans des classes prédéfinies (copéptide sans œufs, copéptide avec œufs). La détection d'objet permet de localiser des objets d'intérêt dans les images (ici le copéptide est délimité dans un rectangle rose et les œufs dans un rectangle orange). La segmentation permet d'associer chaque pixel d'une image à une classe (en rose les pixels correspondant à un copéptide, en orange ceux correspondant à des œufs). Enfin, la régression permet d'estimer la probabilité de présence d'œufs dans l'image (0% en haut, 42% en bas). Figure issue de l'article Orenstein et al. *Machine learning techniques to characterize functional traits of plankton from image data*. 2021. ([hal-03482282](https://hal.archives-ouvertes.fr/hal-03482282))

Sakina-Dorothee Ayata, Maître de conférences en écologie marine à Sorbonne Université au laboratoire LOCEAN.

