

These OceaniX / NavalGroup

**Méthodes IA multimodales dans des contextes
d'observation océanographique et de
surveillance maritime multi-capteurs hétérogènes**

Encadrement: D. Cazau, N. Farrugia, R. Fablet (Chaire IA OceaniX, IMT Atlantique/Lab-STICC)

Références NavalGroup: Alexandre Gensse, Fabien Chaillan (NavalGroup)

Equipe d'accueil: Lab-STICC, Chaire IA OceaniX ([link](#))

Résumé

Le développement de nouvelles méthodes de fusion et de reconnaissance de données multi-capteurs adaptées aux données volumineuses et hétérogènes est nécessaire afin de mieux structurer et exploiter la richesse des *Big Ocean Data*. En s'inspirant de modèles du type réseaux de neurones profonds avec schémas d'apprentissage variationnel, cette thèse traitera de problématiques méthodologiques spécifiques aux approches multimodales (par ex, explicabilité de la décision multimodale, reconstruction d'observations manquantes/partielles) à travers deux cas d'application liées à l'observation océanographique (prédiction des conditions météocéanographiques de surface) et la surveillance des espaces maritimes (séparation et reconnaissance de sources acoustiques sous-marines spécifiques à certaines activités marines telles que bateaux et sous-marins).

Cette thèse réalisée dans le cadre de la chaire IA OceaniX fait l'objet d'un partenariat industriel avec NavalGroup.

Contexte

De par sa dynamique spatio-temporelle complexe et sa nature multi-physique, l'océan est un objet complexe à étudier, nécessitant des programmes d'observation interdisciplinaires et multi-échelles pour être finement analysés. Au sein de tels programmes, de nombreux types de capteurs d'observation co-existent pour la surveillance d'un même site. Une particularité des *Big Ocean Data* qui en découlent est leur nature fortement multimodale (c.a.d différentes natures de signaux, par ex séries temporelles, images, etc) et hétérogène (c.a.d différentes résolutions et échelles temporelles et spatiales). Face à l'expansion de ces *Big Ocean Data*, rendues massivement accessibles au plus grand nombre, le développement de nouvelles méthodes de fusion et de reconnaissance de données multi-capteurs adaptées aux données volumineuses et hétérogènes est nécessaire afin de mieux structurer et exploiter la richesse des *Big Ocean Data*.

Objectifs

Dans cette thèse, nous nous proposons d'explorer le potentiel des approches d'apprentissage de représentation multimodale hétérogène pour la surveillance maritime multi-capteur, pouvant notamment servir à de l'aide à la décision sur nos différents cas d'application. Deux cas d'applications nous intéresseront en particulier, un premier visant à caractériser et prédire les conditions météocéanographiques de surface, et un second sur une tâche de séparation et reconnaissance de sources acoustiques sous-marines spécifiques à certaines activités marines telles que bateaux et sous-marins. Dans ces deux cas, différents types de capteurs / modalités de signaux seront utilisés (acoustique passive / séries temporelles, SAR / Images, bouées in-situ / séries temporelles, radar / Images, échosondeur multifaisceaux / Images, AIS / "four-hot encoding"), constituant de larges réseaux de capteurs multimodaux distribués dans l'espace, très hétérogènes en termes d'échantillonnages espace-temps, mais dont les différentes composantes portent potentiellement des informations utiles complémentaires pour la surveillance d'un objet cible, les biais respectifs étant indépendants.

Méthodes et résultats attendus

A l'issue d'une phase initiale d'état de l'art portant à la fois sur l'apprentissage profond et l'apprentissage multimodal ainsi que le contexte thématique de la thèse (monitoring des océans, typologies et caractéristiques des capteurs d'intérêt), le plan de travail proposé comprend deux parties principales: la première dédiée au développement théorique et algorithmique de modèles d'apprentissage et de reconnaissance à partir de données multimodales hétérogènes, la deuxième seconde visant à leur application sur des cas d'étude réels en monitoring l'environnement marin.

Apprentissage et de reconnaissance à partir de données multimodales hétérogènes: Dans cette partie, l'objectif sera de développer des modèles d'apprentissage et de reconnaissance à partir de données multimodales hétérogènes. Nous nous intéresserons en particulier à des modèles de type réseaux de neurones profonds, et plus particulièrement à l'apprentissage profond multimodal basé sur des schémas d'apprentissage variationnel. Ceux-ci peuvent être utilisés pour capturer les correspondances entre les modalités à travers la création d'espaces de dimension réduite, dits "espaces latents sous-jacents", où les observations des modalités sont projetées plus ou moins proches les unes des autres selon les «concepts sémantiques» qu'elles représentent, c'est-à-dire les processus physiques qui font les générer. Nous privilégierons notamment des stratégies d'apprentissage adaptées à un contexte faiblement supervisé, soit par la partialité des observations (par ex, modalité manquante ou bruitée), soit par le (très) faible nombre d'échantillons exemplaires vus par l'algorithme durant son apprentissage. Il s'agira alors d'obtenir de nos algorithmes des comportements "intelligents", tels que générer ou prédire une modalité à partir d'une autre, identifier des informations complémentaires inter-modales, traiter des données manquantes ou non observées ou bruyantes dans l'une des modalités, comme déjà été obtenus dans différents domaines d'application tels que la reconnaissance vocale audiovisuelle (Ngiam, 2011) et plus récemment en océanographie (Rao, 2017). Au terme de cette période, une première phase de validation "expérimentale" de ces méthodes théoriques est prévue à partir de jeux de données existants de la communauté internationale (par exemple, NYU Depth Dataset V2¹, MNIST-multimodal², Google Audio Set³). Nous y développerons notamment toute la méthodologie d'évaluation nécessaire à répondre à différentes questions scientifiques clés pour les modèles multimodaux telles que

- quantifier les gains de performance de modèles multimodaux sur des tâches de reconnaissance en les comparant à des modèles de référence unimodaux. Un cas particulièrement intéressant sera d'évaluer le transfert d'apprentissage entre un système d'apprentissage multimodal et un système de reconnaissance unimodal cible non-colocalisés;
- rendre les modèles explicables afin de comprendre les transferts de connaissance entre modalités et ainsi d'évaluer finement les contributions des différentes modalités dans une décision de reconnaissance. Cela permettra de mieux évaluer nos capacités à établir des "situations" de l'environnement perçu ou des "vues" de l'évènement à reconnaître;
- qualifier les avantages du multimodal dans des contextes d'apprentissage faiblement supervisé contenant des observations partielles et/ou en faible quantité. Globalement, ces méthodes devraient permettre de mieux exploiter les larges volumes de données non-labellisées et à mieux extrapoler leurs prédictions au-delà de leur espace d'apprentissage.

¹ https://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html

² <https://github.com/Jakobovski/decoupled-multimodal-learning>

³ <https://research.google.com/audioset/>

Application au monitoring de l'environnement marin: L'objectif sera ici de démontrer et évaluer les concepts et algorithmes développés dans la première partie sur les cas d'application réels suivants:

1. caractérisation et prédiction des conditions météocéanographiques de surface (vent, pluie, état de mer). Les différents capteurs utilisables dans ce contexte seront des images SAR, du radar, de l'acoustique passive et des sorties de modèles numériques (voir détails en Annexe). Ce *framework* d'observation sur l'eau / sous l'eau est hautement multimodal, croisant séries temporelles haute-résolution in-situ (acoustique passive) et imagerie satellitaire large-échelle;
2. Séparation de sources acoustiques sous-marines, avec un focus particulier pour la reconnaissance de bateaux et sous-marins. Les capteurs envisagés ici seront : acoustique passive, données environnementales (modèles numériques, données satellitaires brutes), images SAR, données AIS, observations visuelles de sources cibles et notes de commandement de bord, labels temporels d'annotation par acoustique passive, voir aussi images satellitaires haute-résolution. Un *framework* ici encore hautement multimodal, mêlant images et séries temporelles à différentes résolutions et fréquence d'échantillonnage, ainsi que données textuelles non-structurées et symboliques. Une problématique d'identification de comportements singuliers pourra notamment être traitée, comme déjà réalisés par images SAR (Nguyen, 2018).

Plan de travail

Le plan de travail proposé est le suivant:

M1-M2	Familiarisation avec modèles IA multimodaux et capteurs d'acquisition
M3-M12	Développement de modèles originaux répondant à nos problématiques méthodologiques (par ex, observations partielles / manquantes), validation sur <i>toy datasets</i>
M13-M24	Application 1: prédiction conditions météocéanographiques
M25-M30	Application 2: séparation de sources acoustiques sous-marines
M31-M36	Valorisation (publications et conférences) et rédaction du manuscrit

Encadrement

Cette thèse sera co-encadrée par D. Cazau (ENSTA Bretagne, Lab-STICC), R. Fablet et N. Farrugia (IMT Atlantique, Lab-STICC) dans le cadre de la chaire en IA OceaniX dirigée par R. Fablet. Elle bénéficiera de l'environnement interdisciplinaire développée par la chaire sur le campus brestois au sein de l'EUR Isblue. Cette thèse s'inscrit également dans le cadre d'une collaboration industrielle au sein de la chaire OceaniX avec Naval Group pour le monitoring et la surveillance de l'environnement marin.

Dans le cadre de la chaire OceaniX, le(la) candidate participera également à des actions de formation et animation à l'IA pour les sciences et technologies marines et bénéficiera d'un soutien à la mobilité internationale au sein d'équipes internationales partenaires de la chaire OceaniX (e.g., S. Brunton, Univ. Washington, Seattle; S. Matwin, Dalhousie Univ., Halifax, A. Mahalevan, Woods Hole Inst., Mass.; L. Bertino, NERSC, Bergen; F. Doblaz-Reyez, BCS, Barcelone).

Bibliographie

- Ngiam, J. et al. (2011) Multimodal deep learning Proc. 28th Int. Conf. Mach. Learn., pp. 689–696
- Nguyen, D. et al. (2018). A Multi-task Deep Learning Architecture for Maritime Surveillance using AIS Data Streams. arXiv:1806.03972
- Rao, D. et al. (2017) Multimodal learning and inference from visual and remotely sensed-data The International Journal of Robotics Research, Vol. 36(1) 24–43

Annexe sur jeux de données disponibles

Détails supplémentaires sur les jeux de données spécifiques à chaque application:

1. Les conditions météocéanographiques telles que vent, état de mer, pluie font partie des produits satellitaires les plus couramment traités, donc de nombreux jeux de données existent en accès libre. Nous sommes déjà en possession des suivants: par images SAR (par ex, Sentinel-1 ESA, dataset annoté par Wang (2018)), par modèles d'assimilation de données (par ex, ERA-INTERIM ECMWF ~ 20 ans de données), bouées NOAA et météo France... D'autres jeux de données sont issus de différents consortiums de recherche: données d'acoustique passive avec stations météo et radar co-localisés (collaboration avec NRC⁴ (~2 ans de donnée) et GéoAzur (~1 an de donnée) en Méditerranée);
2. Plusieurs jeux de données d'acoustique passive en accès libre (par ex, workshop DCLDE) ou issus de collaboration scientifique (par ex, Sorbonne Universités, Géoazur, CEBC Chizé, AFB) pourront être exploités dans cette thèse, au sein de différents environnements (petits/grands fonds) et paysages acoustiques. Nous disposons aussi de plusieurs partenariats déjà en place pour bénéficier de données AIS (par ex, CLS).

⁴ <http://www.w1m3a.cnr.it/O11/modules/system/home.php>