

Sélection d'images satellites pour inversion bathymétrique par réseau de neurones convolutif

Amine Chemchem^{1,2}, Guillaume Jubelin¹, Grégory Cazanave¹

¹ IRT Saint Exupery, Toulouse, France ² Atos, Pôle Data Driven Intelligence, Montpellier, France

lamine.chemchem@atos.net, guillaume.jubelin@irt-saintexupery.com,
gregory.cazanave@irt-saintexupery.com

Résumé

Les satellites optiques Sentinel-2 du programme COPERNICUS de l'ESA permettent d'envisager la production récurrente de cartes bathymétriques à partir des observations acquises. Cette production permettrait un suivi récurrent des fonds océaniques aux abords de la plupart des littoraux du globe mais nécessite pour cela des travaux d'automatisation. La sélection des images candidates à l'inversion bathymétrique reste notamment une tâche manuelle reposant sur l'expertise d'un spécialiste.

Cette sélection est réalisée sur la base de critères environnementaux photo-interprétés : couverture nuageuse, présence de sun glint¹ ou de sillages de bateaux, ou encore hétérogénéité de la masse d'eau. Nous proposons dans ces travaux de substituer cette photo-interprétation de l'expert par une approche d'intelligence artificielle. Elle s'appuie sur un réseau de neurones convolutifs développé à cet effet et entraîné sur une banque d'images labélisées à partir de propriétés optiques de la masse d'eau produites par le logiciel ACOLITE².

Le modèle proposé est performant avec une précision moyenne (f-score) de 95%. La mise en oeuvre de méthodes d'explicabilité a démontré que ces performances sont obtenues à partir d'éléments concrets et pertinents.

Mots-clés

Imagerie spatiale optique, paramètres environnementaux, classification, CNN, explicabilité.

Abstract

Sentinel-2 optical satellites of ESA's COPERNICUS programme allow the recurrent production of bathymetric maps based on acquired observations. This would allow recurrent monitoring of the ocean floor near most of the world's coasts, but requires automation work. In particular, the selection of candidate images for bathymetric inversion remains a manual task requiring the expertise of a specialist. Images selection is based on photo-interpreted environmental criteria : cloud cover, sun glint, boat wakes, or water

body heterogeneity. In this work, we propose to replace this expert photo-interpretation by an artificial intelligence approach. It is based on a convolutional neural network developed for this purpose and trained on a set of images labeled using optical properties of the water body produced by the ACOLITE software.

The proposed model performs well with an average accuracy (f-score) of 95%. The implementation of explicability methods has shown that these performances are obtained from concrete and relevant elements.

Keywords

remote sensing, environmental parameters, classification, CNN, explainability.

1 Introduction

La classification d'images satellites comme exploitable ou non pour l'inversion bathymétrique est un problème complexe nécessitant de prendre en compte aussi bien le contenu atmosphérique (couvert nuageux, aérosols) que l'état de surface de la masse d'eau (déferlante, sillage, réflexion spéculaire) ou encore de sa composition (panache turbide, remise en suspension, efflorescence algale) [1].

Alors qu'un expert est capable de qualifier la contribution dans l'image des phénomènes précités, des algorithmes de quantification de ceux-ci sont disponibles : définition d'un masque de nuage, correction du glint, estimation de la turbidité de la masse d'eau, etc.

Plutôt que d'assembler et de paramétrer ces algorithmes dans un système expert pour réaliser la classification, l'idée des travaux présentés est d'exploiter ces algorithmes pour constituer un jeu de données pour l'entraînement d'un réseau de neurones profond. En procédant de la sorte, la tâche chronophage et complexe de définition d'un système expert est supprimée. Nous évitons également de mettre en oeuvre de nombreux algorithmes sans savoir si l'image traitée sera exploitée par la suite. Enfin, nous nous appuyons sur les performances éprouvées des réseaux de neurones profonds pour cette tâche complexe de classification [2].

La création du jeu de données d'entraînement est supervisée par l'expert avec une interprétation visuelle et la mise en oeuvre de règles de décisions simples prises sur des paramètres environnementaux extraits automatiquement des

1. sun glint : réflexion spéculaire de l'éclairement solaire en direction du capteur.

2. ACOLITE : processeur d'images satellites du Royal Belgian Institute of Natural Sciences.

images. Il est ainsi possible de constituer rapidement un jeu de données assez important pour l'entraînement d'un réseau de neurones. L'effort de calcul est concentré en amont et pendant la phase d'entraînement du réseau, la simple inférence de ce dernier est réalisée au moment de la classification. La capacité à généraliser ce problème de classification ne dépend plus du paramétrage des algorithmes en fonction des littoraux traités mais de la représentativité de la diversité de cas introduite dans le jeu de données.

La substitution d'un modèle déterministe comme un système expert par un modèle probabiliste tel qu'un réseau de neurone soulève toutefois la problématique du contrôle de la prise de décision. Pour comprendre celle-ci et contrôler le fait que le réseau de neurones s'appuie sur des éléments cohérents de l'image, des méthodes d'explicabilité sont utilisées.

Cet article est organisé comme suit : la section 2 rapporte la littérature des travaux connexes, la section 3 détaille la méthodologie de cette étude, et la section 4 analyse les résultats obtenus avant de conclure.

2 État de l'art

La tâche de classification, au sens association d'un label (exploitable ou non-exploitable dans notre cas) à une image, n'est pas répandue dans le domaine de la télédétection. L'image de télédétection est surtout un support pour la cartographie et la classification dans ce domaine s'entend pixel à pixel correspondant ainsi à la tâche de segmentation dans le domaine de l'apprentissage profond. Chen et al. [3] présente tout de même une revue détaillée des méthodes de classification d'image basées sur les réseaux de neurones convolutifs tandis que [4, 5] se focalisent sur l'utilisation des réseaux de neurones sur des données de télédétection.

D'autres travaux d'apprentissage machine et d'apprentissage profond essaient de traiter des problématiques connexes sans passer par les méthodes de classification, comme par exemple le papier [6] qui propose un réseau de neurones profond capable d'évaluer des cartes de bathymétrie, en modélisant le problème sous forme de classification pixel par pixel. Le papier [7] essaie d'exploiter la corrélation spatiale locale entre les pixels avec un modèle CNN pour la prédiction de la profondeur de l'eau, en tenant compte de la relation non linéaire entre la valeur de radiance et la valeur de la profondeur de l'eau des pixels adjacents et centraux. Des données de terrain telles que des données de profondeurs d'eau mesurées et des données de sondages lidar ont été utilisées comme entrées pour construire ce modèle, les résultats démontrent l'efficacité du modèle obtenu. L'inconvénient majeur de cette approche reste la partie recueil de données terrain (avec des relevés Lidar) qui est généralement coûteuse et qui nécessite souvent un travail laborieux d'analyse et de correction d'erreurs de récupération de données avant de lancer l'apprentissage du modèle.

Nous pouvons citer d'autres travaux d'apprentissage profond appliqués aux images satellite comme les auteurs de [8] qui proposent un réseau de neurones à convolutions pour la détection de nuages et de leurs ombres portées. Ce

modèle a été capable de traiter en même temps les données en provenance de World-View 2 et de Sentinel-2. Dans [9] un modèle hybride conçu de deux classifieurs (un réseau de neurones multi-couches et un réseau de neurones convolutif) a été proposé afin de traiter la problématique de classification multi-classes. Ce modèle a surpassé les deux classifieurs individuels (MLP et CNN) ainsi que le GLCM-MLP qui inclut les caractéristiques de texture GLCM. Ces approches d'apprentissage profond rapporte à la littérature une amélioration nette en performance, par contre elle nécessitent un travail d'étiquetage minutieux en amont de la phase d'apprentissage.

Pour contourner cela, notre idée est d'utiliser l'outil ACOLITE pour automatiser cette étape tout en assurant une qualité d'étiquetage très proche du travail de l'expert. En effet, nous avons pu configurer ACOLITE de façon à extraire à extraire des paramètres pertinents de la colonne d'eau. La combinaison de ces paramètres avec des seuils appropriés nous a permis d'étiqueter de manière efficace les images Sentinel-2 en classes (exploitable /non-exploitable) comme montré dans la section 3.

De plus, nous avons mis en œuvre des méthodes d'explicabilité sur le modèle d'apprentissage profond construit afin de qualifier les caractéristiques image utilisées par ce dernier dans son processus de prise de décision. Cela permet de s'extraire de la vision "boite noire" souvent utilisée à propos des modèles neuronaux. Dans ce sens, nous pouvons citer plusieurs recherches récentes comme [10] où les auteurs proposent une méthode d'explicabilité qui démontre son efficacité sur la mise en évidence des régions de l'image contributrices à la prise de décision du modèle. Les auteurs de [11] proposent une revue pertinente résumant les méthodes d'explicabilité récentes. Dans notre étude, nous utiliserons l'outil "Xplique" développé par l'IRT Saint Exupéry dans le cadre du projet Deel.ai³. Cet outil se base sur plusieurs méthodes locales adaptées aux réseaux de neurones. Ces méthodes sont associées à différentes métriques permettant d'évaluer les résultats de l'explication.

3 Méthodologie

Notre idée dans cette étude est d'utiliser l'outil ACOLITE pour labéliser les images Sentinel-2 et de créer un dataset d'apprentissage en image exploitables et non-exploitables. Par la suite, nous allons comparer plusieurs architectures et configurations de réseaux au travers de plusieurs apprentissages afin de sélectionner le meilleur modèle en qualité de précision et en temps d'inférence. Une fois le meilleur modèle sélectionné, des tests de validation et des analyses d'explicabilité sont mis en œuvre avant de le mettre en production pour l'inférence et la classification des nouvelles images Sentinel-2 disponibles.

La méthodologie proposée dans cette étude, comme montrée à la figure 1, peut être répartie en trois étapes principales :

L'extraction des images Sentinel-2 originales et de leurs patches, la labélisation via ACOLITE de ces patches et la

3. DEpendable & EXplainable Learning : <https://www.deel.ai>

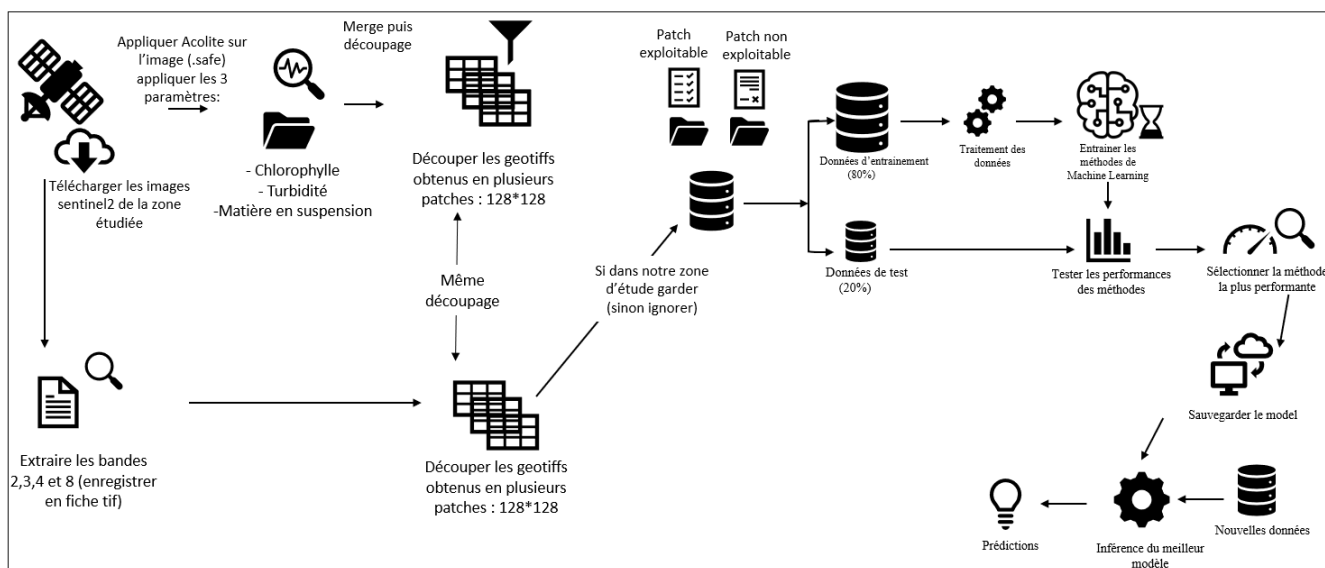


FIGURE 1 – Schéma du workflow global

phase d'affectation des labels.

Une fois l'apprentissage réalisé, nous avons par la suite approfondi l'analyse des résultats avec une étude d'explicabilité, notamment sur les mauvaises classifications. Elle nous a permis de mettre en avant le fait que les décisions du réseau sont majoritairement prises à partir des pixels du domaine continental et non du domaine océanique. De ce fait, nous avons décidé de masquer toutes les images avec un masque de terre pour contraindre l'apprentissage sur le domaine océanique. L'étude d'explicabilité sur ce nouvel apprentissage a bien permis de valider l'efficacité de ce masquage quant à la prise de décision.

3.1 Phase d'extraction de l'image originale et de ses patchs

Dans cette phase, l'image originale est extraite à partir des produits Sentinel-2, en exploitant les bandes 2,3,4, et 8 (bleu, vert, rouge et la proche-infrarouge). Ces spectres ont été sélectionnés afin d'exploiter la meilleure résolution possible offerte par Sentinel-2 sans avoir de traitement de sur ou sous échantillonnage à réaliser (10m). Une fois que l'image (raster) contenant ces 4 bandes est extraite, nous la découpons en plusieurs tuiles (patchs) de taille égale (128*128) afin de ne conserver que les tuiles faisant partie de notre zone d'étude : la bande littorale du département de l'Hérault sur 250km de largeur, voir la figure2.

3.2 Phase de passage d'ACOLITE et génération de ses patchs

Comme expliqué précédemment, nous utilisons ACOLITE pour labéliser les images Sentinel-2 en deux classes : exploitables et non-exploitables. Pour ce faire, nous nous intéressons aux produits de sortie mesures physiques ACOLITE qui contribuent directement ou indirectement à évaluer la qualité de l'image satellite pour le traitement d'inversion bathymétrique. Dans ce sens, nous avons identifié trois



FIGURE 2 – Tuiles étudiées à partir d'une image S2 (la bande littoral héraultaise)

paramètres principaux qui sont : **la turbidité, la chlorophylle et la matière en suspension** ; paramètres calculables et produits au format geotiff pour chaque image Sentinel-2 de notre base de données d'apprentissage.

La turbidité : désigne la teneur d'un fluide en matières qui le troublent. Dans les cours d'eau, elle est généralement causée par des matières en suspension et des particules colloïdales qui absorbent, diffusent ou réfléchissent la lumière. Quand un fleuve turbide se jette en mer, il crée généralement un bouchon vaseux, parfois bien visible depuis un satellite, ceci veut dire qu'une eau trop turbide rend l'image inexploitable pour le calcul de la bathymétrie. A noter que dans ACOLITE, la turbidité est calculé en suivant l'algorithme de Nechad et al[12]. Nous avons fixé par expérimentations le seuil de turbidité acceptable inférieure à 20 mg/L par conséquent, les patchs qui dépassent cette valeur sont classés comme non exploitables.

La chlorophylle : En plus de la turbidité, il est important d'étudier dans les zones non turbides l'indicateur chlorophylle. En effet, cet indicateur reflète la biomasse phytoplanctonique présente dans la colonne d'eau. La chloro-

phylle est un pigment photosynthétique très répandu, il en existe différents types et le plus représenté dans les végétaux marins est la chlorophylle *a* [13]. Elle est de plus représentative de la matière organique végétale, vivante ou fraîchement morte. Sachant que la présence de cet indice impacte négativement l'étude de la bathymétrie, nous avons fixé le seuil de 1.6 µg chl_a/L par expérimentations pour différencier les images exploitables de celles non exploitables.

La matière en suspension (MES) : autre facteur qui a un impact négatif sur le calcul de la bathymétrie. En effet, MES désigne les matières solides insolubles visibles à l'œil nu présentes en suspension et peut atteindre jusqu'à 3 kg/m² mais est surtout confinée près de la côte, pour des profondeurs inférieures à 20 mètres on constate que la zone où la concentration en MES est la plus forte correspond à la zone de déferlement des vagues [14]. Par expérimentations nous avons fixé le seuil de 10 mg/L au générateur ACOLITE pour déterminer les tuiles acceptables de celles qui ne le sont pas.

3.3 Phase d'étiquetage

La phase d'expérimentations de la labélisation nous a permis d'affiner la construction du critère final de seuillage permettant de discriminer au mieux les 2 classes pour chaque patch. Nous nous sommes notamment appuyés sur la comparaison avec une labélisation manuelle réalisée par un expert en bathymétrie côtière. Le critère optimal retenu prend en considération en plus de la moyenne des concentrations des constituants de la colonne d'eau (matière en suspension (TSM), la turbidité (TUR) et la Chlorophylle (CHLA)), le nombre de pixels satisfaisant et ne satisfaisant pas ces conditions (pixel ok - pixel nok) afin de labéliser au patch. Il peut être décrit comme suit :

Algorithm 1 Méthode d'étiquetage des patches

Pour chaque "i" dans patches **faire** :

Si $0 < TSM < 10$ et $0 < TUR < 20$ et $0 < CHLA < 1.6$ et Pixel_ok > Pixel_nok **alors** :

patches[*i*] ← *acceptable*

Sinon :

patches[*i*] ← *non_acceptable*

Une fois que les patches des images ACOLITE sont labélisés, nous étiquetons chaque patch correspondant de l'image originale (issue des bandes 2, 3, 4 et 8) avec le même label. L'enjeu est de réaliser l'apprentissage sur les images originales et par la suite (une fois le modèle entraîné) d'appliquer des prédictions directement sur des images originales sans avoir à passer par l'étape ACOLITE.

3.4 Phase d'apprentissage

TABLE 1 – Recherche de la meilleure architecture / hyperparamètres

Architecture	Hyperparamètres	f1-score
Conv2D_1(64) max_pooling2d_1 Conv2D_2(64) max_pooling2d_2 Flatten () Dense_1 (64) Dense_2 (2)	epochs=80 batch_size=16 optimizer=Adam learning_rate=0.00001	76%
Conv2D_1(64) max_pooling2d_1 Conv2D_2(64) max_pooling2d_2 Conv2D_3(64) max_pooling2d_3 Flatten () Dense_1 (64) Dense_2 (64) Dense_3 (2)	epochs=80 batch_size=16 optimizer=Adam learning_rate=0.00001	95%
Conv2D_1(64) max_pooling2d_1 Conv2D_2(64) max_pooling2d_2 Conv2D_3(128) max_pooling2d_3 Conv2D_4(128) max_pooling2d_4 Flatten () Dense_1 (128) Dense_2 (64) Dense_3 (2)	epochs=80 batch_size=16 optimizer=Adam learning_rate=0.00005	92%

Dans cette phase, nous avons implémenté et testé plusieurs modèles personnalisés conçus depuis le début tout en variant les hyperparamètres. L'idée ici est de commencer par l'architecture la plus simple possible et de rajouter des couches et/ou des neurones par couches à chaque expérimentation jusqu'à l'obtention des résultats satisfaisants. L'objectif de cette démarche est de trouver le modèle avec les meilleures performances tout en gardant l'architecture du modèle la moins complexe possible afin d'éviter un besoin excessif en matière de ressources de calcul.

De plus, pour chaque architecture, nous cherchons les hyperparamètres optimaux, ceci est facilement implémentable avec Keras_tuner⁴. Les résultats de cette expérimentation rapportés dans le tableau 1 nous ont permis de sélectionner l'architecture la plus simple et la plus efficace en termes de f1-score. Elle est constituée de 3 couches de convolutions suivies de 3 couches denses.

La Figure 3 résume l'architecture du modèle retenu.

Dans cette étude, nous avons rassemblé un dataset contenant 41 images Sentinel-2 datant de 2017 à 2021 de la zone côtière héraultaise, ces 41 images sont découpées en patches de 128*128 pixels. L'idée ici est de traiter des tuiles de petites taille pour épouser les contours des nuages et autres panaches turbides et ainsi exploiter une plus grande partie des images. Ce découpage nous a donné un ensemble d'apprentissage de 17690 patches. Nous nous sommes assurés

4. https://www.tensorflow.org/tutorials/keras/keras_tuner

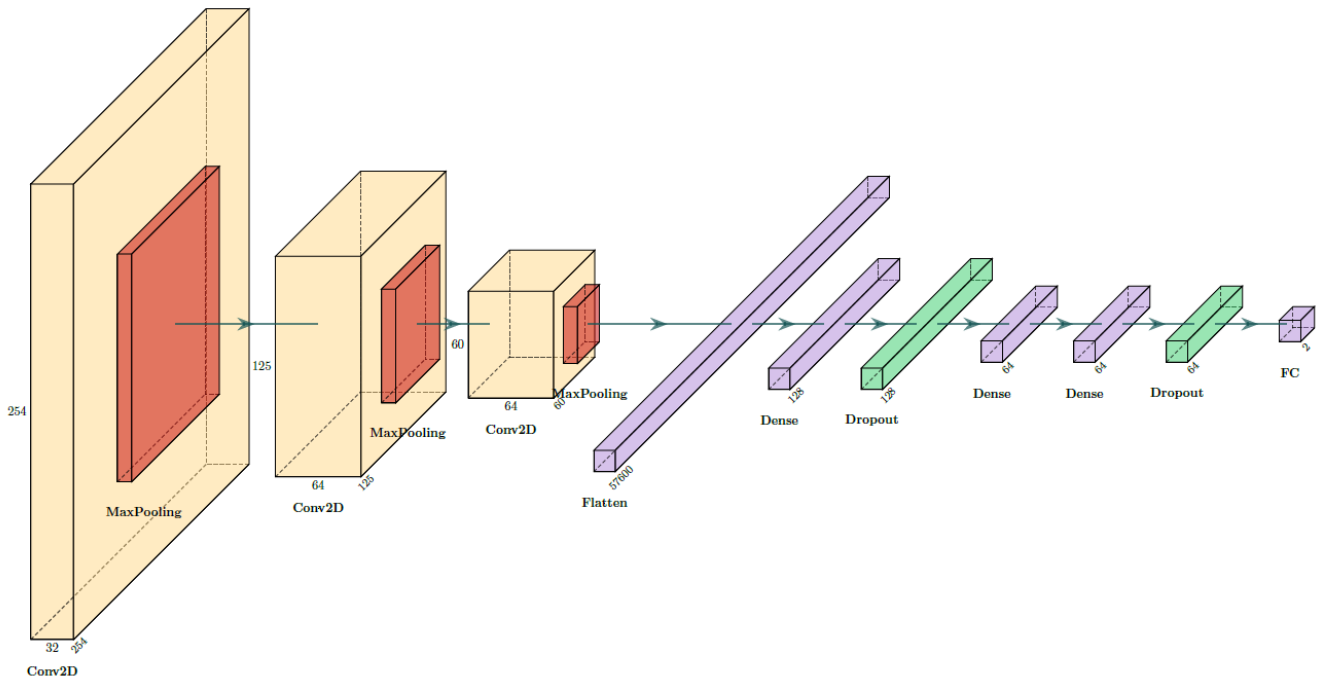


FIGURE 3 – Schéma de l'architecture du modèle CNN retenu

de constituer un dataset équilibré (autant d'images exploitables que non exploitables). Après labélisation, nous avons obtenu 9646 patches labélisés en classe 0 (exploitable) et 8044 labélisés en classe 1 (non exploitable). La répartition entre l'ensemble d'entraînement et de test est de 85%, 15% respectivement comme montré sur la figure 4 :

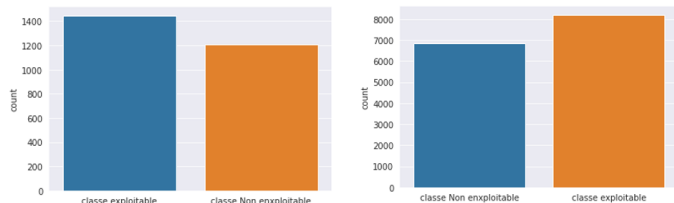


FIGURE 4 – Distribution de patches d'entraînement / patches de test

4 Résultats et discussion

Comme décrit précédemment, nous avons fixé les hyperparamètres du modèle par expérimentations par les méthodes de recherche automatique des hyperparamètres, les paramètres retenus sont :

- Nombre d'époques = 80,
- L'optimiseur = Adam,
- l'indice d'apprentissage = 0.00001,
- La taille du batch = 16.

Avec ces paramètres, l'apprentissage du modèle se réalise de façon optimale sans sur-apprentissage apparent (voir la figure 5).

En testant le modèle de classification obtenu sur notre en-

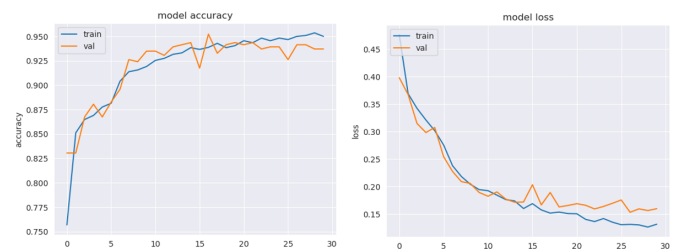


FIGURE 5 – Évolution de la précision et de la loss lors de l'entraînement

semble de test, nous obtenons la matrice de confusion schématisée à la figure 6 ainsi que les évaluations des performances comme mentionnées à la figure 7.

Pour rappel les formules d'évaluation sont définies mathématiquement comme suit :

$$Accuracy(Taux_de_justesse) = \frac{VP + VN}{VP + VN + FP + FN}$$

$$Precision = \frac{VP}{VP + FP} \quad , \quad Rappel = \frac{VP}{VP + FN}$$

$$F1-score = \frac{2 * Precision * Rappel}{Precision + Rappel} = \frac{2 * VP}{2 * VP + FP + FN}$$

Sachant que :

VP : représente les vrais positifs,

VN : les vrais négatifs,

FP : les faux positifs,

FN : les faux négatifs.

Ces résultats montrent bien que le modèle obtenu est très efficace , nous remarquons bien à partir de la matrice de

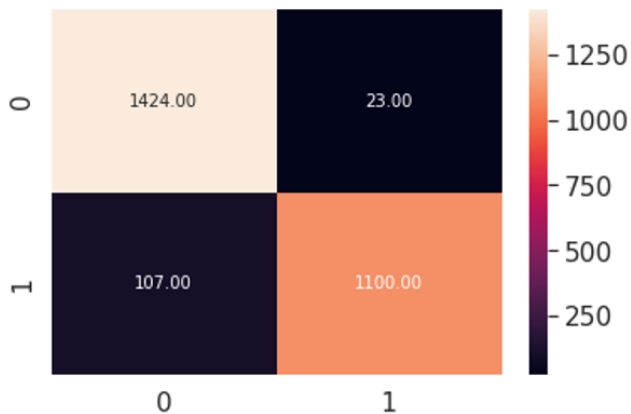


FIGURE 6 – Matrice de confusion

	precision	recall	f1-score	support
0	0.93	0.98	0.96	1447
1	0.98	0.91	0.94	1207
accuracy			0.95	2654
macro avg	0.95	0.95	0.95	2654
weighted avg	0.95	0.95	0.95	2654

FIGURE 7 – Évaluation et performance du modèle

confusion, que sur les 2654 patches de test, le modèle ne s’est trompé que sur 130 patches : 107 patches prédits comme exploitables alors qu’ils ne le sont pas et sur 23 patches prédits en non-exploitables alors qu’ils le sont.

En évaluant le modèle, ce dernier obtient une précision f1-score moyenne de 95% de précision, score très satisfaisant par rapport à notre objectif de mise en production régulière de la chaîne bathymétrique.

4.1 Explicabilité du modèle retenu

Les résultats obtenus paraissent satisfaisants, néanmoins nous avons approfondi l’analyse avec des méthodes d’explicabilité (avec la librairie Xplique⁵ issue du projet DEEL réalisé par les équipes de l’IRT Saint Exupéry) afin de qualifier notamment d’où les erreurs de prédiction peuvent provenir. L’explicabilité permet de connaître les pixels qui influencent le plus la prise de décision du modèle obtenu. Voici pour exemple, 4 tuiles pour l’étude de l’explicabilité, ces 4 tuiles ont été classées correctement par notre modèle de la manière suivante : en Non-exploitables (patch1792_1536, patch1536_1536) et en Exploitables (patch768_2816, patch768_2560) et les résultats de l’explicabilité avec les méthodes Rise, Lime, KernelShape et Occlusion sont schématisés à la figure 8.

D’après ces résultats, nous pouvons constater que les méthodes donnent des cartes de chaleur assez similaires, surtout lorsqu’il s’agit de la méthode ‘Lime’ et ‘KernelSha-

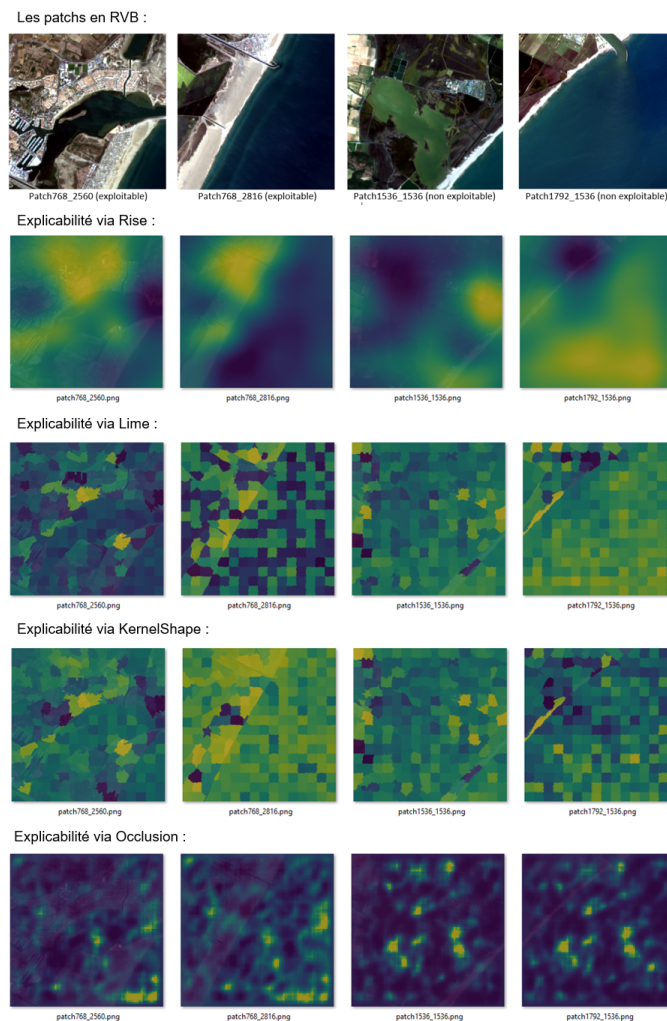


FIGURE 8 – Explicabilité du modèle

pe’ ; elles se basent toutes les deux sur des super-pixels pour schématiser les zones contributrices à la prise de décision lors de l’inférence du modèle. Nous remarquons aussi que la plupart des décisions sont prises à partir des pixels représentant des zones habitées (zone terre) au lieu de se concentrer sur les pixels de la zone côtière comme c’était le cas lors de la phase de labélisation avec les calculs des paramètres ACOLITE. La conclusion est donc que notre apprentissage est biaisé par l’environnement alors que nous souhaitons seulement nous concentrer sur des paramètres physiques maîtrisés de la colonne d’eau. Par conséquent, nous avons décidé de refaire l’apprentissage du modèle en appliquant un masque de terre (land mask) pour contraindre l’apprentissage à se faire exclusivement sur les pixels ne contenant que la colonne d’eau (représentant que des pixels mer).

4.2 Expérimentations avec des images masquées (par le masque terre)

Notre idée en appliquant le masque terre (land mask) sur notre base d’apprentissage issue d’images Sentinel-2 est de pousser le modèle à ignorer les pixels représentant la

5. <https://github.com/deel-ai/xplique>

terre et de raisonner uniquement sur les pixels représentant la mer. Pour ce faire, nous avons extrait une partie du masque terre depuis OpenStreetMap⁶ qui correspond à notre zone d'intérêt. Nous montrons en figure 9 un exemple d'une image de notre base d'apprentissage en appliquant ce masque.

Après l'application du masque, nous suivons toutes les étapes précédentes comme pour les images non masquées : découpage en patches 128*128 pixels, puis labélisation avec ACOLITE et découpage entre ensemble d'apprentissage et ensemble de test pour arriver à l'étape d'implémentation et d'expérimentation du modèle CNN. L'architecture du meilleur modèle retenu pour les images masquées en termes de précision f1-score est identique à celui retenu pour les images non masquées. La matrice de confusion et les résultats d'évaluation de ce modèle sont rapportés respectivement sur la figure 10 et la figure 11.

Nous remarquons que les résultats du modèle entraîné sur les images masquées sont légèrement supérieurs à ceux du modèle entraîné sur les images non masquées. D'après la matrice de confusion, nous pouvons noter que sur l'ensemble de 2707 images de test, il ne se trompe que sur 130 patches : 77 patches non exploitables prédits comme exploitables et 53 exploitables prédits comme non exploitables. Ceci donne un f-score moyen de 95% qui est très satisfaisant. Ces résultats illustrent que non seulement la prédic-

6. <https://osmdata.openstreetmap.de/data/land-polygons.html>

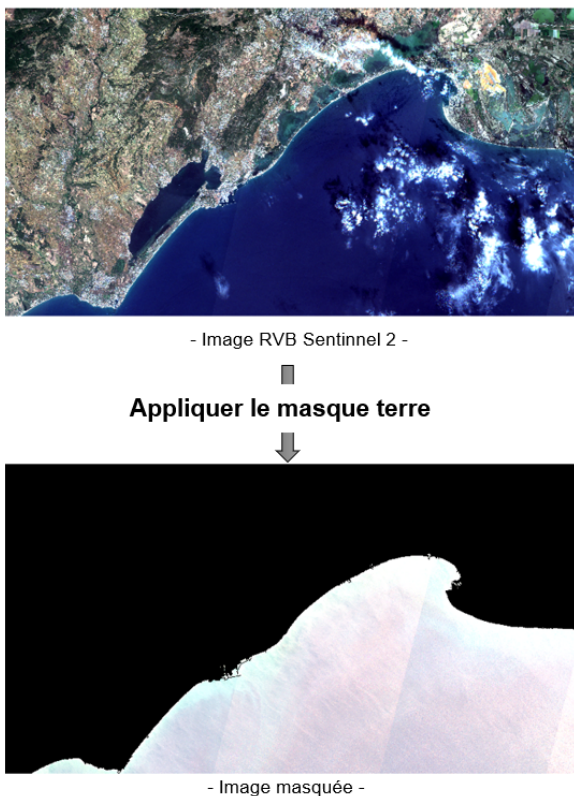


FIGURE 9 – Procédure de masquage d'image Sentinel-2 par le masque terre

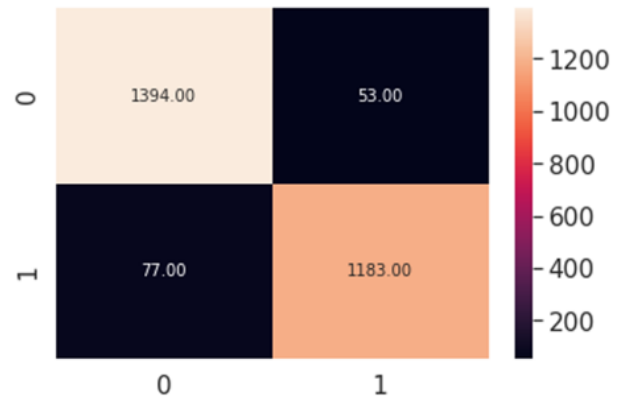


FIGURE 10 – Matrice de confusion du modèle sur les images masquées

	precision	recall	f1-score	support
0	0.95	0.96	0.96	1447
1	0.96	0.94	0.95	1260
accuracy			0.95	2707
macro avg	0.95	0.95	0.95	2707
weighted avg	0.95	0.95	0.95	2707

FIGURE 11 – Performances du modèle sur les images masquées

tion est plus explicable, plus maîtrisée mais également que le masquage permet de réduire le nombre de prédictions en erreur.

De la même manière que pour les images non masquées, en appliquant la méthode d'explicabilité "Rise" sur un ensemble de tuiles, nous obtenons les résultats affichés à la figure 12.

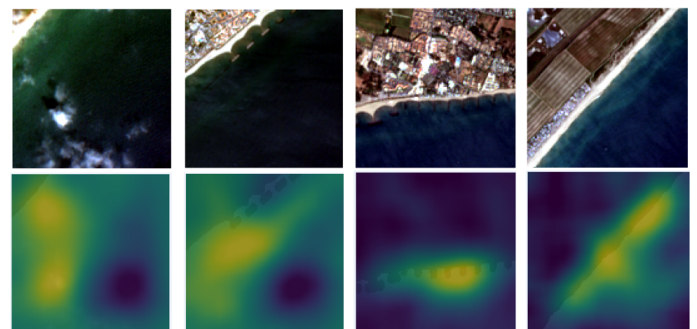


FIGURE 12 – Explicabilité du modèle via la méthode 'Rise' sur les images masquées

Malgré l'apport du masquage qui permet de s'absoudre des pixels continentaux dans la prise de décision comme illustré dans la figure 12 au travers de la méthode d'explicabilité Rise, il reste des cas en erreur et notamment des faux positifs. Étant donné que ces faux positifs engendrent des cal-

culs inutiles, il nous paraît intéressant de pousser l'analyse de l'explicabilité en la corrélant aux images de paramètres physiques calculés afin de tenter de réduire leur occurrence. Il existe également un intérêt sur l'optimisation des faux négatifs : l'optimisation de la couverture spatiale mais celle-ci apparaît comme moins dimensionnante car moins de cas d'erreur et un délai de revisite de 5 jours avec les satellites Sentinel-2.

5 Conclusion et perspectives

Dans cette étude, nous avons pu développer un modèle CNN capable d'identifier de manière très précise si une image satellite est exploitable ou pas par le processus d'inversion bathymétrique. En effet, nous avons pu réaliser un étiquetage très efficace à l'aide de l'outil ACOLITE en utilisant des paramètres physiques de la colonne d'eau tels que : la turbidité, la chlorophylle et les matières en suspension. Grâce à cet étiquetage, nous avons pu proposer un modèle CNN à 6 couches que nous avons entraîné sur des images masquées dans le domaine continental pour le pousser à raisonner uniquement sur les pixels de l'image représentant la mer et à ignorer les pixels terre. Avec cette approche, nous avons obtenu un modèle de classification d'image satellitaire efficace et explicable, car comme montré avec l'étude d'explicabilité, il génère des cartes de caractéristiques à partir des pixels mer et a obtenu un f1 score moyen de 95%.

Comme perspectives, nous prévoyons de poursuivre l'analyse d'explicabilité en étudiant les corrélations entre celle-ci et les paramètres physiques produit par ACOLITE. Nous prévoyons également d'exploiter lors de l'apprentissage du modèle les masques de nuages et de leurs ombres produits par ACOLITE. En effet, les travaux d'explicabilité ont montré des erreurs de classification liées à ceux-ci (voir la figure 13). Ces erreurs sont en particulier liées à la génération de valeurs aberrantes des paramètres physiques sur les nuages et leurs ombres. Les performances du modèle devraient être ainsi améliorées en masquant ces derniers. Par ailleurs, nous comptons évaluer les performances de ce modèle sur des images provenant d'autres satellites tels que les satellites Landsat de la NASA. Ces derniers possèdent des bandes spectrales comparables à celles des Sentinel-2 mais proposent des résolutions spatiales moins fines permettant d'évaluer les capacités du modèle à généraliser selon ce dernier critère. Enfin l'intégration dans la chaîne automatique d'inversion bathymétrique du modèle développé sera réalisée, elle permettra la sélection automatique d'images Sentinel-2 dans un contexte opérationnel.

Remerciements

Cette étude a été réalisée dans le cadre du projet de recherche SB : Monitoring environnemental à l'échelle locale & aide à la décision – Bassin de Thau. Mené par l'IRT Saint Exupery, il est co-financé par les sociétés TELESPAZIO, ATOS, INATYSCO, INSIDE, SUEZ EAU FRANCE (Rivages Pro Tech) et réalisé en collaboration avec le Syndicat Mixte du Bassin de Thau et Sète Agglopolé.

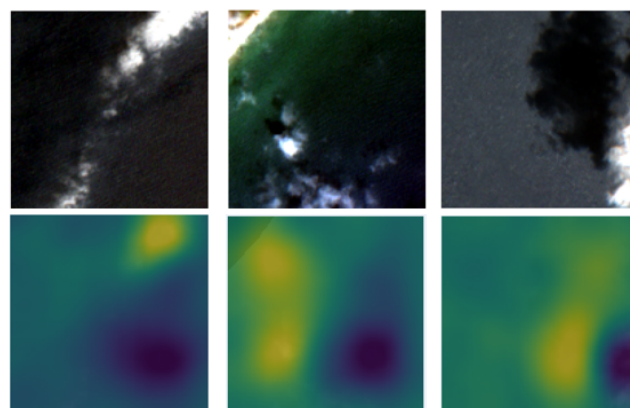


FIGURE 13 – Explicabilité du modèle via la méthode 'Rise' sur les images nuageuses

Références

- [1] E. Salameh, F. Frappart, R. Almar, P. Baptista, G. Heygster, B. Lubac, D. Raucoules, L. P. Almeida, E. W. J. Bergsma, S. Capo, M. De Michele, D. Idier, Z. Li, V. Marieu, A. Poupardin, P. A. Silva, I. Turki, and B. Laignel, "Monitoring Beach Topography and Nearshore Bathymetry Using Spaceborne Remote Sensing : A Review," *Remote Sensing*, vol. 11, p. 2212, Jan. 2019. Number : 19 Publisher : Multidisciplinary Digital Publishing Institute.
- [2] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications : A meta-analysis and review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 152, pp. 166–177, June 2019.
- [3] L. Chen, S. Li, Q. Bai, J. Yang, S. Jiang, and Y. Miao, "Review of Image Classification Algorithms Based on Convolutional Neural Networks," *Remote Sensing*, vol. 13, p. 4712, Jan. 2021. Number : 22 Publisher : Multidisciplinary Digital Publishing Institute.
- [4] L. Zhang, L. Zhang, and B. Du, "Deep Learning for Remote Sensing Data : A Technical Tutorial on the State of the Art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, pp. 22–40, June 2016. Conference Name : IEEE Geoscience and Remote Sensing Magazine.
- [5] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, "Deep Learning in Remote Sensing : A Comprehensive Review and List of Resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, pp. 8–36, Dec. 2017. Conference Name : IEEE Geoscience and Remote Sensing Magazine.
- [6] B. Wilson, N. C. Kurian, A. Singh, and A. Sethi, "Satellite-derived bathymetry using deep convolutional neural network," in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, pp. 2280–2283, 2020.

- [7] B. Ai, Z. Wen, Z. Wang, R. Wang, D. Su, C. Li, and F. Yang, "Convolutional neural network to retrieve water depth in marine shallow water area from remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 2888–2898, 2020.
- [8] M. Segal-Rozenhaimer, A. Li, K. Das, and V. Chirayath, "Cloud detection algorithm for multi-modal satellite imagery using convolutional neural-networks (cnn)," *Remote Sensing of Environment*, vol. 237, p. 111446, 2020.
- [9] C. Zhang, X. Pan, H. Li, A. Gardiner, I. Sargent, J. Hare, and P. M. Atkinson, "A hybrid mlp-cnn classifier for very fine resolution remotely sensed image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 140, pp. 133–144, 2018. Geospatial Computer Vision.
- [10] M. Brahim, S. Mahmoudi, K. Boukhalfa, and A. Moussaoui, "Deep interpretable architecture for plant diseases classification," in *2019 Signal Processing : Algorithms, Architectures, Arrangements, and Applications (SPA)*, pp. 111–116, 2019.
- [11] Q.-s. Zhang and S.-C. Zhu, "Visual interpretability for deep learning : a survey," *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 1, pp. 27–39, 2018.
- [12] B. Nechad, K. G. Ruddick, and Y. Park, "Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters," *Remote Sensing of Environment*, vol. 114, no. 4, pp. 854–866, 2010.
- [13] A. Minghelli-Roman and C. Dupouy, "Influence of Water Column Chlorophyll Concentration on Bathymetric Estimations in the Lagoon of New Caledonia, Using Several MERIS Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, pp. 739–745, Apr. 2013. Conference Name : IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
- [14] Y. Leredde, H. Michaud, E. Berthebaud, C. Lauer-Leredde, P. Marsaleix, C. Estournel, B. Guerinel, S. Thorin, T. Schvartz, and C. Richard, "Modélisation numérique de l'hydrodynamique sédimentaire dans la baie d'Aigues-Mortes (Languedoc-Roussillon, France). Application à la géomorphodynamique et à la remise en suspension des sédiments rechargés sur les plages," in *XIIIèmes JNGCGC Dunkerque*, pp. 447–458, Editions Paralia, 2014.