



MASTER 2 DE MODÉLISATION EN ÉCOLOGIE

Rapport de stage

2023-2024

Auteure : Romane DUBOIS

Taxonomie et morphométrie, peut-on remplacer l'humain par la machine ?

Application au projet DEMERSTEM

Structure d'accueil :

UMR DECOD
INRAE (Bât. 19) 65, rue de
Saint-Brieuc CS 84215 35042
Rennes

Encadrants :

François MARTIGNAC
Jérôme GUITTON
Marie-Pierre ETIENNE

SOUTENANCE : 20 JUIN 2024

Remerciements

Je remercie tout d'abord Jules Trillaud et Pierre Scala pour leur collaboration sur ce début d'étude. Je tiens également à exprimer ma gratitude aux équipes respectives de mes trois encadrants pour leur accueil chaleureux. Enfin, et ce n'est pas le moins important, je remercie mes trois encadrants pour leur accompagnement toujours marqué par la bienveillance et la bonne humeur.

Table des matières

1	Introduction	1
2	Les réseaux de neurones convolutifs (CNN)	3
3	Les images de DEMERSTEM	4
3.1	Protocole de récolte des données et ses contraintes	4
3.2	Labellisation des images	6
4	Performances de l'analyse automatique des images	8
4.1	Identification de l'espèce	8
4.1.1	Description des 3 architectures	8
4.1.2	Entraînement des modèles	10
4.1.3	Comparaison des performances des 3 modèles	11
4.2	Détection des points d'intérêts	11
4.2.1	Optimiser les performances avec le choix du jeu de données d'entraînement	12
4.2.2	Performances sur les tâches liées à l'étiquette	12
4.2.3	Performance des modèles pour le positionnement des points d'intérêt des poissons	13
5	Extraction des données d'intérêts halieutiques	14
5.1	Longueur totale du poisson	14
5.1.1	Description des 3 méthodes comparées	14
5.1.2	Comparaison des performances des 3 méthodes	16
5.2	Caractérisation automatique des mesures de morphométries	17
5.2.1	Évaluer la précision de la labellisation	17
5.2.2	Longueur prédite	17
6	L'approche automatique conduit elle aux mêmes conclusions sur l'identification entre stocks ?	18
6.1	Transformation des données	19
6.2	Démarche statistique et comparaison des résultats	19
6.2.1	Analyse discriminante linéaire (LDA)	19
6.2.2	ANOVA	20
7	Discussion	22

8 Conclusion et perspectives	24
A Annexe A	28

Table des figures

1	Fonctionnement général d'un réseau de neurones convolutif pour une tâche de classification	4
2	Carte d'échantillonnage des 6 espèces de poissons et de crustacé	5
3	Standardisation des images	6
4	Description des points d'intérêts pour les espèces avec une (gauche) et deux (droite) nageoire dorsale	6
5	Description des liens entre la base de données, l'interface et l'entraînement du modèle sur pyhton.	8
6	Description de l'architecture du modèle construit manuellement (A), de VGG16 (B) et de YOLOv8 (C) (Liu et al. [2024])	9
7	Précision de la prédiction des points de l'étiquettes en fonction du type du modèle	13
8	Précision de la prédiction des points du poisson en fonction du type de modèle entraîné pour les classes BSC et PSE, les autres classes sont dans l'annexe A	14
9	Diagramme de dispersion sur le jeu de données de DEMERSTEM (A) et le notre (B), des deux premières fonctions discriminantes de l'analyse discriminante linéaire (LDA) pour les distances morphométriques de <i>P. bellottii</i> (PAR) : Côte d'Ivoire et Ghana, par zones (Ouest et Est).	20
10	Classification des individus de <i>P. bellottii</i> (PAR) (en %) du jeux de données DEMERSTEM (A) et du notre (B), dans leur population d'origine en utilisant la matrice de classification de l'Analyse Discriminante Linéaire (LDA) basée sur la morphométrie des poissons.	21
11	Scores de l'axe 1 de l'Analyse Discriminante Linéaire (LDA) basée sur la morphométrie des individus <i>P. bellottii</i> (PAR) du jeux de données de DEMERSTEM (A) et le notre (B)	21
12	Précision de la prédiction des points du poisson en fonction du type de modèle entraîné pour les 5 classes	28

Liste des tableaux

1	Equivalence entre le nom latin de l'espèce de poisson et son code FAO	7
2	Paramètres d'augmentation modifiés et leur valeurs	11
3	Pourcentage d'individus dont l'espèces a été bien prédite à partir des 3 modèles	11
4	Précision de la mesure de l'étiquette de 3 cm pour chaque modèle	13
5	Différence moyenne en centimètre entre les mesures terrains et les mesures labellisées entre les points et sur la boîte englobante par espèces	16
6	Différence moyenne en centimètre entre les mesures labellisées et les mesures prédites par le modèle pour les 3 méthodes par classe de taille	16
7	Différence moyenne en centimètre entre les mesures labellisées et les mesures prédites par le modèle pour les 3 méthodes par classe de taille	17
8	Nombre d'individus de <i>pagellus bellotti</i> (PAR) analysés et intervalle de longueur totale (Lt) par zone de pêche étudiée du jeu de données de DEMERSTEM.	18
9	Nombre d'individus de <i>pagellus bellotti</i> (PAR) analysés et intervalle de longueur totale (Lt) par zone de pêche étudiée de notre jeu de données.	18

1 Introduction

En 2021, la proportion des stocks marins exploités de manière biologiquement durable a chuté à 62,3% (FAO [2024]). L'application de gestion efficace des pêches pour permettre la récupération de ses stocks sur-exploités, est plus que nécessaire pour inverser les tendances de déclin de la durabilité. C'est d'autant plus vrai là où la pêche prend une part importante du système économique national, tel qu'au Sénégal où la pêche représente 14,1% du PIB. C'est politique de gestion de pêche durable repose sur une évaluation pertinente de ces stocks.

C'est dans ce contexte que le projet européen DEMERSTEM (DEMERSal ecosySTEM) (García-Isarch et al. [2023]) a été mis en place. Il est le fruit d'une collaboration entre 10 organismes scientifiques spécialisés en halieutique (européens et africains) et vise à améliorer le processus d'évaluation des stocks en Afrique de l'Ouest. Les stocks de poissons sont définis au sein de ce contexte ouest-africain sur la base des Zones Économiques Exclusives (ZEE) ou de regroupements de ZEE. Or, pour évaluer correctement un stock, il convient en premier lieu de bien définir la population concernée, sur des critères écologiques et non, politiques. En effet, les dynamiques de rendement de chaque stock peuvent différer en fonction de leur productivité biologique et de l'impact relatif de la pêche sur la croissance, la survie et la capacité de reproduction. La définition des stocks a été l'objet de multiples études (Marini et al. [2017] ; Turan [2004] ; Pawson and Jennings [1996]) et méthodes au cours des dernières décennies , soulignant ainsi l'importance de cette démarche pour réguler efficacement les activités de pêche et assurer une exploitation plus durable des ressources halieutiques (Begg et al. [1999]).

L'un des axes de DEMERSTEM a donc été de mieux caractériser les stocks de 6 espèces démersales afin de permettre une meilleure évaluation , et *in fine*, une meilleure régulation des activités de pêche qui impactent ces stocks. Cette régulation est d'autant plus importante que le risque de surpêche est élevé sur ces espèces démersales, dont la capacité de reproduction est plus lente que celle des espèces pélagiques (durée de vie plus longue et maturité plus tardive des espèces démersales). Pour définir des stocks à partir des données collectées, les individus ont été caractérisés en fonction de leur morphologie, de leurs traits d'histoire de vie et de leur correspondance génétique. Pour étudier la morphologie des individus, l'équipe de DEMERSTEM a mené une étude par méthode de réseaux de truss (Strauss [1982]). Ceux-ci sont définis comme une série de distances mesurées entre des points anatomiques d'un individu, ces distances sélectionnées forment un motif régulier. Les différents protocoles d'acquisition, définis par l'Institut de recherche océanographique espagnol (IEO), coordinateur de cette activité au sein du projet DEMERSTEM, ont été déployés par les partenaires africains (Mauritanie, Sénégal,

Guinée-Bissau, Guinée, Ghana, Côte d’Ivoire), et les collectes de données ont été réalisées dans les 6 pays ciblés par les instituts de recherche halieutique nationaux.

Cette étude a nécessité l’acquisition d’un très grand nombre de données pour caractériser ces individus en fonction de leurs paramètres biologiques. Par le nombre important de données récoltées, la question de l’automatisation d’une partie de ce protocole de recherche se pose. En effet, avec l’augmentation de la disponibilité de données, allant des images de pièges photographiques à celles des satellites, en passant par les données des campagnes en mer, le traitement d’images et de vidéos devient de plus en plus crucial. Des méthodes comme l’apprentissage profond (*deep learning*) a déjà démontré son potentiel en écologie (Miele et al. [2021]), notamment grâce à sa capacité à automatiser des tâches répétitives.

On peut citer, l’identification automatique des espèces par analyse d’image qui est une application courante de l’apprentissage profond. Par exemple, des algorithmes ont été développés et utilisés pour des mammifères (Willi et al. [2019]; Norouzzadeh et al. [2018]), des insectes (Ärje et al. [2020]) et des poissons (Villon et al. [2018]). Villon et al. ont utilisé un modèle pour identifier 20 espèces de poissons différentes, et les résultats montrent une précision de 94,9% contre 89,3% pour l’identification par des humains.

Pour la mesure de la longueur totale, diverses méthodes ont été explorées. Des approches comme l’érosion morphologique (Huang et al. [2016]) ont été utilisées pour extraire la ligne médiane de la zone du poisson préalablement détectée par apprentissage profond. D’autres ont utilisé différents types d’architectures, comme un Mask R-CNN (Álvarez Ellacuría et al. [2020]) ou un Faster R-CNN (Furuta et al. [2021]), pour détecter la tête et la queue des poissons.

Pour détecter ce réseau automatiquement, il faut d’abord détecter les points d’intérêt, puis mesurer les distances morphométriques. Cette approche pour un réseau complet de longueurs a encore très peu été étudiée. Cependant, Yu et al. ont développé en 2023 un modèle pour la détection de points d’intérêt afin de mesurer la longueur totale d’un poisson et ont obtenu une précision de 85,1% pour la détection des points d’intérêt.

Le premier objectif de cette étude est d’évaluer la possibilité de détecter l’espèce d’un poisson photographié de façon automatique. Le deuxième est de pouvoir obtenir la longueur totale du poisson. Cette longueur est utilisée dans le protocole pour l’analyse des traits d’histoire de vie des individus. Le troisième objectif vise à établir le profil morphologique des individus de la même manière que le protocole DEMERSTEM le définit, et ce, de façon automatique à partir d’une image.

Ce rapport débute par une courte introduction sur le fonctionnement de l’apprentissage profond, puis par une description des données utilisées. Ensuite, chaque section présente les méthodes et les résultats des différents objectifs d’automatisation. Enfin,

une comparaison est effectuée entre les résultats obtenus à partir des jeux de données prédits par l'apprentissage profond et ceux obtenus via le protocole de DEMERSTEM. Cette comparaison vise à évaluer la concordance des résultats lorsque la même méthode statistique est appliquée à ces deux ensembles de données. L'objectif est de déterminer si les prédictions issues de l'apprentissage profond sont cohérentes avec les observations recueillies selon le protocole établi par DEMERSTEM, ce qui permettrait de valider l'efficacité de l'approche automatisée dans ce contexte spécifique.

2 Les réseaux de neurones convolutifs (CNN)

L'apprentissage profond est un procédé d'apprentissage automatique, il se caractérise par l'utilisation de réseaux de neurones artificiels (ANN). Ce type de réseau est composé de plusieurs couches de neurones artificiels, chacune connectées entièrement à la suivante. Les connexions entre les neurones sont associées à des poids, qui sont ajustés lors du processus d'apprentissage pour améliorer la précision du modèle. Le passage d'une couche à la suivante se fait par la combinaison d'une fonction de régression et d'une fonction d'activation. Certains réseaux de neurones artificiels sont associés à des couches de convolution présentes en amont, pour le traitement des images ou d'autres types de données visuelles.

C'est LeCun et al. qui, en 1998, utilisent pour la première fois ce terme de « convolution » dans ce contexte avec le développement d'un réseau de neurones convolutif (ou *Convolutional Neural Network*, CNN) pour la classification de nombres écrits à la main. Une couche de convolution a pour but d'extraire les caractéristiques de l'image. Elle est constituée de plusieurs filtres de convolution. Tous les filtres d'une couche partagent la même largeur et la même hauteur. Ils sont définis comme des matrices de poids qui détectent des motifs spécifiques dans l'image, tels que les bords, les textures, ou d'autres caractéristiques visuelles. Ces poids sont également optimisés lors de la phase d'entraînement. Les filtres opèrent en parallèle sur la même image, et à chaque portion de l'image rencontrée, un calcul de convolution s'effectue, produisant en sortie une carte de caractéristiques (*feature map*) qui met en avant les caractéristiques dans l'image.

Là où d'autres méthodes de classification nécessitaient une extraction des caractéristiques à la main, les CNN permettent d'automatiser cette étape. Les caractéristiques d'une image, dans ce sens, correspondent à ce que « voit » le modèle. Ce sont ces caractéristiques qui définissent les variables d'entrée de la tête de classification, constitué de plusieurs couches de neurones entièrement connectées (*dense*). Les catégories prédites sont les variables de sorties de cette tête de classification. (cf Figure 1)

Afin de construire un modèle robuste et de bien évaluer ses performances de prédiction le jeu de données est divisé en trois parties. La première partie est le jeu d'entraînement,

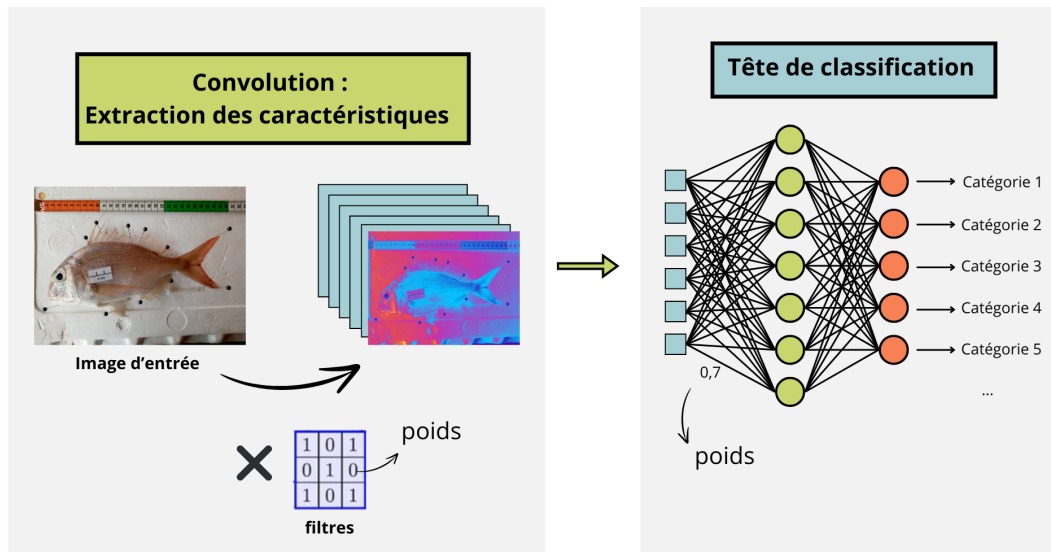


Figure 1 – Fonctionnement général d'un réseau de neurones convolutif pour une tâche de classification

qui va servir à entraîner le modèle au sens propre. Ce jeu de données contient la majorité des échantillons et est utilisé pour ajuster les poids du réseau de neurones à chaque itération. Pendant l'entraînement, le modèle apprend en minimisant une fonction de perte à l'aide d'un algorithme d'optimisation. La deuxième partie est le jeu de validation, utilisé pendant la phase d'entraînement pour ajuster les hyperparamètres (nombre et taille des filtres, augmentation de données...) et évaluer la performance du modèle sur des données indépendantes de l'entraînement. Le jeu de validation aide à prévenir le surapprentissage en fournissant une mesure de performance intermédiaire et permet de décider quand arrêter la phase d'entraînement. La troisième partie est le jeu de test, utilisé uniquement après l'entraînement complet du modèle. Il permet de calculer les performances finales du modèle sur des données totalement nouvelles, ce qui permet de fournir une évaluation indépendante de l'entraînement, de la capacité du modèle à généraliser. Au cours de ce projet, le jeu de données a été réparti selon les proportions suivantes : 70% pour l'entraînement, 10% pour la validation et 20% pour le test.

3 Les images de DEMERSTEM

3.1 Protocole de récolte des données et ses contraintes

Le projet DEMERSTEM a permis de recueillir 3 970 images de cinq espèces de poissons et d'une espèce de crustacé dans les eaux relevant de la juridiction nationale (ZEE)

des pays partenaires (Figure 2). Cependant, seules les données relatives aux poissons ont été utilisées au cours de cette étude, ce qui diminue le nombre d'images à 3 259. Pour récolter ces données, l'IEO (Instituto Español de Oceanografía) a mis en place un protocole d'échantillonnage (García-Isarch et al. [2020]) afin de réaliser des photos standardisées lors des campagnes scientifiques et sur les lieux de débarquement des pêcheurs professionnels artisans (pirogues). Le protocole impose une standardisation de chaque image avec un individu à plat de profil, une règle graduée et une étiquette de 3 cm standardisée posée près de la nageoire pectorale. (Figure 3).

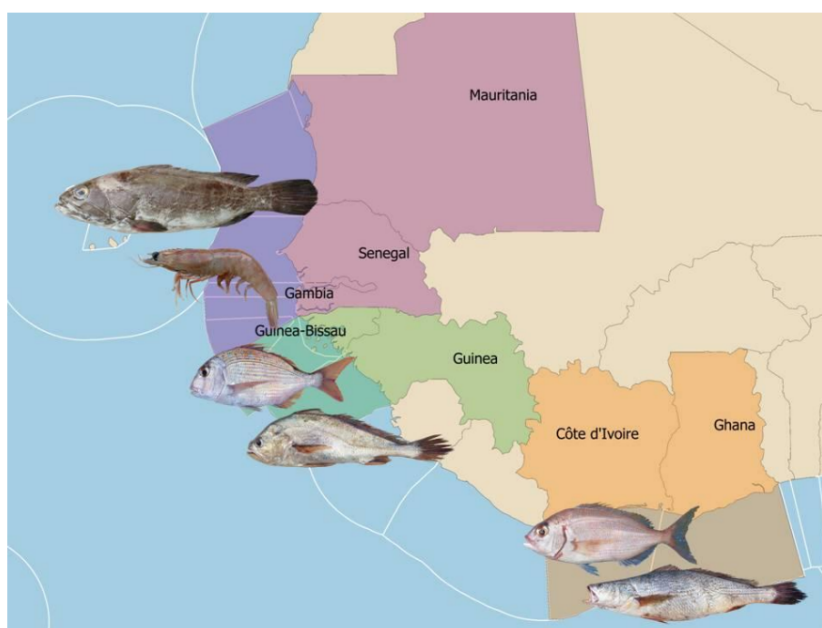


Figure 2 – Carte d'échantillonnage des 6 espèces de poissons et de crustacé

La diversité des partenaires impliqués dans ce protocole a nécessité la mise en place d'un dispositif simple. Aucun appareil photo spécifique n'a donc été imposé, ce qui a entraîné une variété de résolutions dans les images capturées. En conséquence, la relation pixel-centimètre diffère d'une photo à l'autre. C'est pour pallier ce problème de résolution que l'étiquette standardisée de 3 cm a été ajoutée aux images.

Afin de définir les individus morphologiquement et de construire le réseau de *truss*, des points d'intérêts ont été sélectionnés. La figure 4 montre les emplacements des points d'intérêts utilisés pour chaque groupe d'espèces considéré, avec *E.aeneus*, *P.caeruleostictus* et *P.bellottii* dans un premier groupe et *P.elongatus* et *P.senegalensis* dans le deuxième. Ils sont similaires à ceux utilisés dans d'autres analyses de réseau de *truss* (Rawat et al. [2017]). Deux groupes d'espèces sont identifiés en fonction du nombre des points d'intérêt

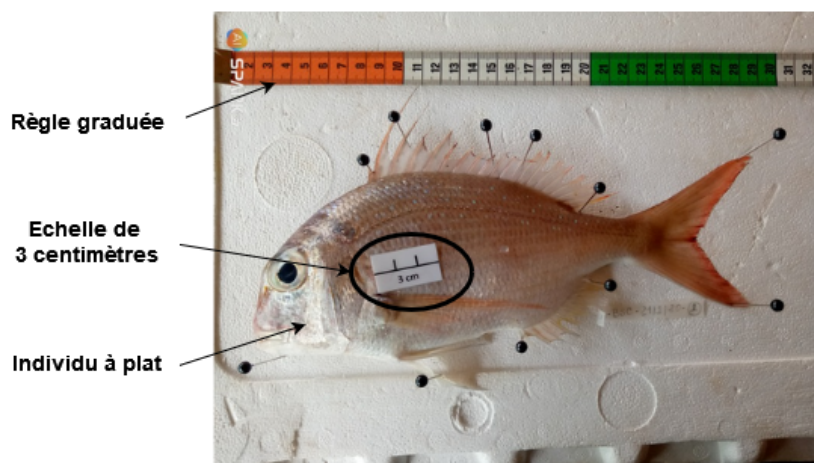


Figure 3 – *Standardisation des images*

à utiliser. Ces points sont situés sur le côté du poisson et sont des caractéristiques clairement identifiables de l'espèce, correspondant notamment aux points d'insertion des nageoires. Après avoir identifié les points d'intérêt sur chaque image, le réseau de *truss* sur le corps du poisson a été construit en les interconnectant pour former un ensemble complet de mesures.



(1) la pointe antérieure du museau au niveau de la mâchoire supérieure ; (2) l'origine de la première nageoire dorsale ; (3) l'origine de la deuxième nageoire dorsale ; (4) le point antérieur-dorsal de la nageoire caudale ; (5) le point antérieur-ventral de la nageoire caudale ; (6) l'origine de la nageoire anale ; (7) l'origine de la nageoire ventrale ; (8) le point le plus postérieur de la mâchoire ; (9) l'extrémité postérieur du poisson

Figure 4 – *Description des points d'intérêts pour les espèces avec une (gauche) et deux (droite) nageoire dorsale*

3.2 Labellisation des images

La labellisation du jeu de données durant cette étude vise à :

- Renseigner l'espèce du poisson, nommée par son code FAO (Tableau 1)
- Placer une boîte englobante au tour du poisson et de l'étiquette

— Placer les points d'intérêts sur le poisson (Figure 4) et l'étiquette

Afin de reproduire le même protocole que celui de DEMERSTEM, les points d'intérêt placés sont les mêmes que ceux choisis pour le projet. Les points de l'étiquette permettent de calculer une longueur de référence pour convertir ces distances de pixel à centimètres.

Table 1 – *Equivalence entre le nom latin de l'espèce de poisson et son code FAO*

Nom de l'espèce	Code FAO	Effectifs
<i>Epinephelus aeneus</i>	GPW	357
<i>Pagrus caeruleostictus</i>	BSC	1102
<i>Pagellus bellottii</i>	PAR	603
<i>Pseudolithus elongatus</i>	PSE	478
<i>Pseudolithus senegalensis</i>	PSS	694

Pour l'annotation des images, l'utilisation d'un outil existant est une option envisageable, car il en existe une multitude répondant à divers besoins. Toutefois, afin de mutualiser les données de labellisation et d'optimiser ce processus, il a été décidé, dans le cadre de cette étude, de développer une application sous R Shiny (Chang et al. [2023]). Cette application est accessible via le lien suivant : https://halieut.agrocampus-ouest.fr/discardless_app/Image_label/.

Les informations relatives aux images, telles que leur dimension, leur localisation ou encore le projet associé, sont stockées dans une base de données. Ces informations sont ensuite transmises à l'application, qui télécharge uniquement l'image à afficher. Cela permet d'éviter le téléchargement en bloc de toutes les images, ce qui peut être très long en fonction de leur nombre et de leur taille. Une fois les données labellisées, elles sont stockées dans une table de la base de données avec toutes les informations associées à chaque objet (Figure 5). Les labels ainsi que les informations liées aux images labellisées sont accessibles sur la base et importable sur un logiciel de programmation tel que Python, afin de réaliser l'entraînement du modèle. L'entraînement de chaque modèle a été réalisé sur un ordinateur avec GPU (*Graphics Processing Unit*) pour optimiser le temps de calcul.

L'application prend en charge plusieurs fonctionnalités : l'ajout de nouvelles images à la base de données, la labellisation par un opérateur des codes FAO, des boîtes englobantes et des points d'intérêt. Elle possède également une partie utilisation de l'outil qui permet la prédiction de l'espèce, de la longueur totale et des distances du réseau de *truss* en centimètres. L'accès aux sections de labellisation et d'ajout de nouvelles images est sécurisé par un identifiant et un mot de passe, ce qui permet de contrôler qui effectue les annotations et quelles images sont annotées.

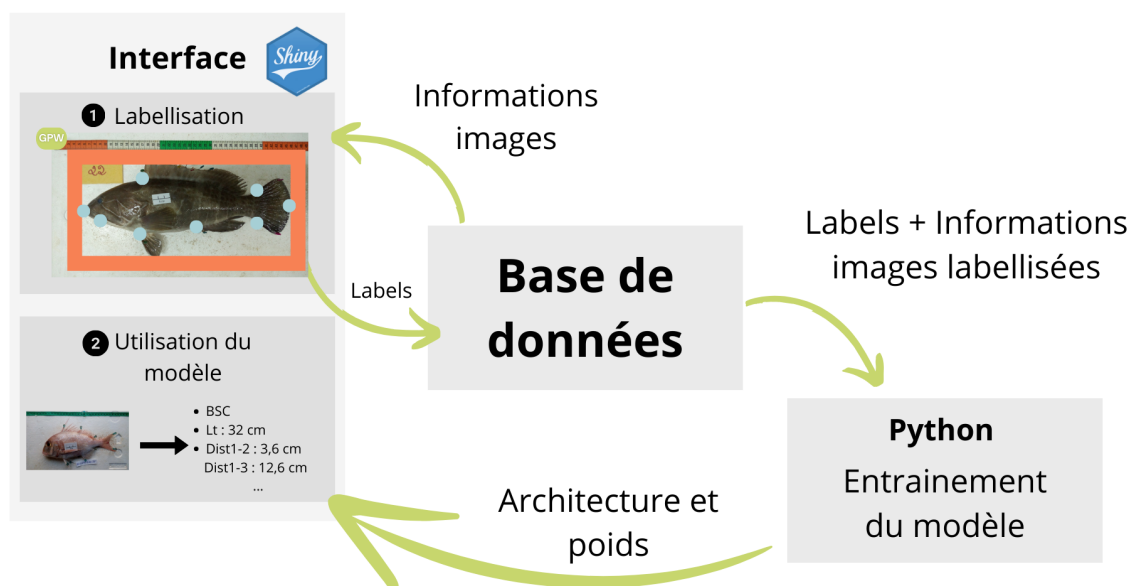


Figure 5 – Description des liens entre la base de données, l’interface et l’entraînement du modèle sur pyhton.

4 Performances de l’analyse automatique des images

Rappelons tout d’abord que ce traitement automatique à deux objectifs : identifier l’espèce et retrouver les points d’intérêts.

4.1 Identification de l’espèce

L’identification automatique des espèces constitue un problème classique de classification. Nous avons comparé plusieurs approches pour aborder ce problème. La première consiste à concevoir intégralement (*from scratch*) une architecture en assemblant différentes couches. Les deux autres approches utilisent le transfert d’apprentissage (*transfer learning*), une méthode qui exploite l’architecture d’un modèle déjà entraîné. Ce modèle est ensuite modifié en réentraînant uniquement les couches de la tête de classification. Les deux architectures comparées sont VGG16 (Simonyan and Zisserman [2015]) et YOLOv8(Ultralytics [2023]).

4.1.1 Description des 3 architectures

Modèle construit intégralement

Le modèle *from scratch* (Figure 6 A) est constitué de quatre blocs. Ils contiennent tous une couche de convolution et une couche de *max-pooling* auquel est associé un

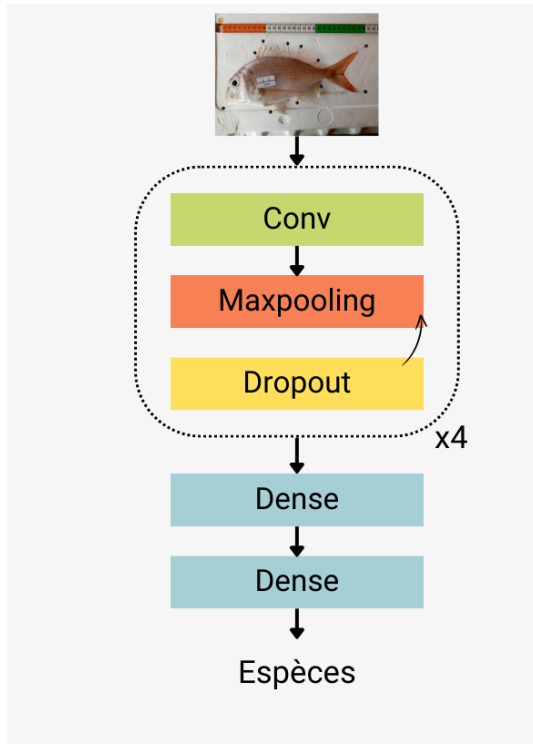
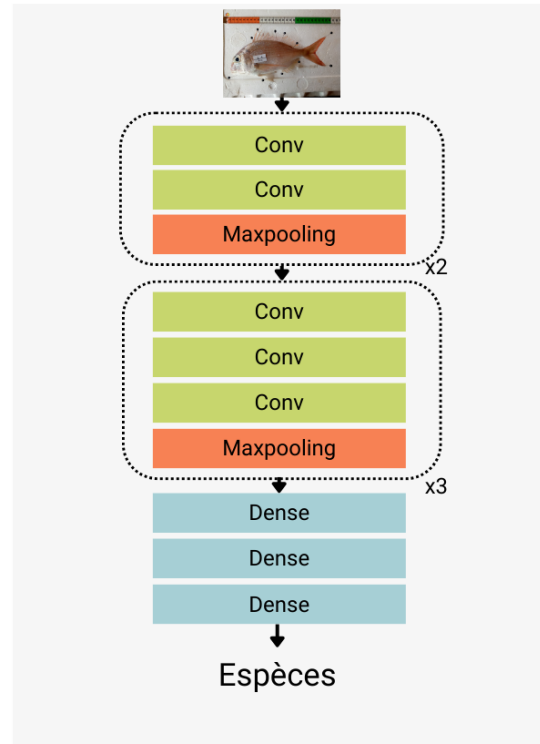
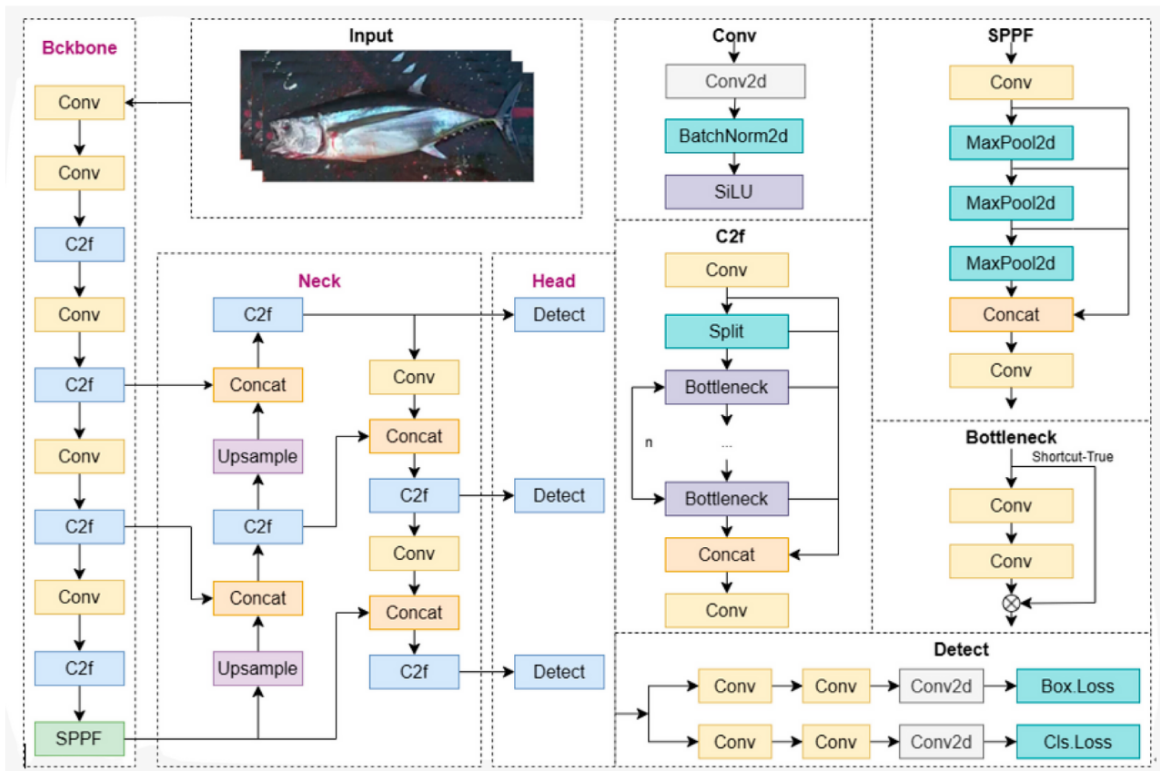
A**B****C**

Figure 6 – Description de l'architecture du modèle construit manuellement (A), de VGG16 (B) et de YOLOv8 (C) (Liu et al. [2024])

paramètre de *dropout*. C'est une technique qui permet d'inactiver aléatoirement et temporairement un certain pourcentage de valeurs dans les cartes de caractéristiques nouvellement définies, c'est-à-dire d'ignorer leur contribution pendant cette itération spécifique de l'apprentissage. Cela force le réseau à ne pas dépendre excessivement de certaines caractéristiques particulières, améliorant ainsi sa robustesse et sa capacité à généraliser sur de nouvelles données, et évitant qu'il ne se spécialise dans des motifs particuliers des données d'entraînement. Ici, ce pourcentage est de 30%.

VGG16

C'est K. Simonyan et A. Zisserman de l'équipe de recherche Visual Geometry Group (VGG) de l'Université d'Oxford qui ont développé cette architecture en 2015 (Figure 6 B). Le chiffre "16" dans VGG16 fait référence aux 16 couches qui contiennent des poids à ajuster. VGG16 comporte au total 21 couches, comprenant 13 couches de convolution, 5 couches de *max pooling* et 3 couches de neurones entièrement connectées.

YOLOv8

YOLOv8 est un modèle développé par Ultralytics. C'est une architecture plus complexe développée sur la figure 6 C . En plus d'une tête de classification YOLOv8 possède une tête de détection qui lui permet de réaliser des tâches de détection en plaçant une boîte englobante au tour de l'objet à détecter. C'est l'utilisation de cette architecture qui induit la labellisation de boîte englobante dans notre méthode.

4.1.2 Entraînement des modèles

Pour prendre en compte l'effet aléatoire de la répartition des données dans les 3 groupes. L'entraînement est effectué 10 fois, avec la répartition en 3 groupes renouvelés à chaque itération. À l'issue de chaque entraînement, une précision de prédiction est calculée sur les prédictions du jeu de données de test, ce qui permet d'obtenir une précision moyenne sur les 10 entraînements.

La précision est définie comme étant une mesure du pourcentage de prédictions correctes par rapport à l'ensemble des prédictions.

Augmentation des données

Afin d'augmenter la variabilité du jeu de données et d'éviter le surajustement, durant la phase d'entraînement, à chaque itération (*epoch*), pour l'architecture développée manuellement et VGG16, les images sont retournées horizontalement ou verticalement et tournées de manière aléatoire avec une probabilité de 0,3. Pour YOLOv8, les paramètres sont légèrement différents en raison de la spécificité de son implémentation en Python.

Les paramètres modifiés par rapport à leur valeur par défaut sont présentés dans le Tableau 2. Pour déterminer les valeurs des paramètres, les critères suivants ont été pris en compte : l'incurvation du poisson, les variations de luminosité et la rotation des images.

Table 2 – Paramètres d'augmentation modifiés et leur valeurs

Paramètres	Valeur
<i>degrees</i>	-180
<i>translate</i>	0
<i>flipud</i>	0.5
<i>mosaic</i>	0
<i>erasing</i>	0.2

4.1.3 Comparaison des performances des 3 modèles

Les résultats sont présentés dans le tableau 3. Le modèle le plus performant parmi les trois est YOLOv8, avec une précision moyenne de 99,95% et un écart-type de 0,07. Cette architecture ainsi que les poids du modèle entraîné sont importés dans l'interface développée pour construire la partie « Utilisation des modèles ».

Table 3 – Pourcentage d'individus dont l'espèce a été bien prédite à partir des 3 modèles

Modèles	Moyenne	Sd
<i>From.Scratch</i>	94.41	0.03
<i>VGG16</i>	95.06	0.01
<i>YOLOv8</i>	99.95	0.07

4.2 Détection des points d'intérêts

La deuxième tâche que nous avons cherchée à automatiser par des méthodes d'apprentissage profond est la détection des points d'intérêt morphologiques du poisson ainsi que des points d'intérêt de l'étiquette. Ces points d'intérêt serviront ensuite à calculer les distances morphométriques et à évaluer la pertinence des stocks tels qu'ils sont actuellement définis.

Cette tâche est un problème de régression et non de classification, donc les architectures utilisées pour l'identification des espèces ne peuvent pas être utilisés. Le développement manuel d'une autre architecture est possible, mais comme mentionné précédemment, ce n'est pas une solution efficace et pertinente. Nous avons donc décidé de choisir une architecture déjà existante d'un modèle déjà entraînée pour réaliser du transfert d'apprentissage. L'architecture utilisée est YOLOv8-pose.

YOLOv8-pose est un réseau de neurones convolutif profond basé sur YOLOv8, qui ajoute l'estimation des points d'intérêts du poisson, en plus de sa détection (Maji et al. [2022]). La totalité de l'architecture de YOLOv8 est reprise et une tête de détection des points d'intérêt est ajoutée.

4.2.1 Optimiser les performances avec le choix du jeu de données d'entraînement

Pour placer automatiquement ces points d'intérêt sur l'image, il a fallu décider quels modèles construire en fonction des espèces étudiées. Il y a plusieurs options. La première est d'entraîner un modèle par espèce, et de construire les modèles suivants : M_{GPW} , M_{PAR} , M_{BSC} , M_{PSE} , M_{PSS} . La deuxième option est d'entraîner un modèle par groupe d'espèces qui se ressemblent morphologiquement, ce qui correspond à un modèle pour les espèces ayant huit points d'intérêt (M_8) et un modèle pour les espèces ayant neuf points d'intérêt (M_9). La dernière option étudiée est d'entraîner un même modèle pour toutes les espèces (M_{global}). Au cours de l'entraînement de chaque modèle les données sont augmentées de la même façon que pour l'identification de l'espèce avec YOLOv8.

4.2.2 Performances sur les tâches liées à l'étiquette

Avant de regarder la performance des modèles sur la prédiction des points d'intérêt des poissons, il est important de s'intéresser en premier lieu à l'échelle de 3 centimètres, car elle nous sert de référence pour la conversion pixel-centimètre. Pour ce faire la distance euclidienne entre les coordonnées des points prédits et annotés est calculée et convertie en centimètre par le calcul suivant :

$$d_{cm} = \frac{d_{pix} * 3}{taille_{pix}}$$

Avec d_{pix} la distance à convertir et $taille_{pix}$ la distance euclidienne en pixel entre les deux points annotés de l'étiquette. Toutes les conversions de pixel à centimètre dans le reste du rapport sont réalisées grâce à ce calcul.

Sur la figure 7 on peut voir que les points de l'étiquette sont bien détectés avec les deux points prédits à environ 0,1 cm (3%) de leur point labellisé respectif. Il n'y pas de différences de prédiction entre les deux points. Concernant la longueur prédite de l'étiquette en centimètre, la MAE (*mean absolute error*) entre la longueur labellisée et prédite, est calculée pour chaque modèle et est détaillée dans le tableau 4. Au mieux la longueur est prédite à $\pm 0,061$ cm (2%) ce qui correspond à la prédiction du modèle M_{global} . Les moins bonnes estimations correspondent aux prédictions des modèles réparti

par espèces, à l'exception du modèle M_{BSC} dont le nombre d'images est supérieur à celles des 4 autres classes.

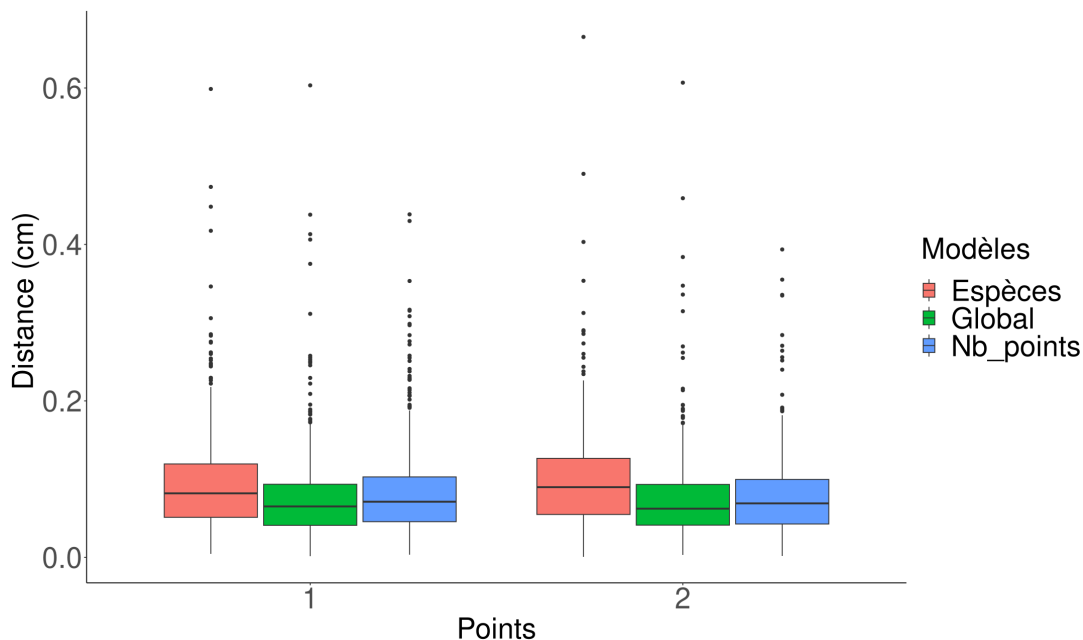


Figure 7 – Précision de la prédiction des points de l'étiquettes en fonction du type du modèle

Table 4 – Précision de la mesure de l'étiquette de 3 cm pour chaque modèle

Modèles	M_{GPW}	M_{PAR}	M_{BSC}	M_{PSE}	M_{PSS}	M_8	M_9	M_{global}
MAE (cm)	0.109	0.101	0.065	0.085	0.084	0.064	0.072	0.061

4.2.3 Performance des modèles pour le positionnement des points d'intérêt des poissons

Afin d'évaluer la précision de la prédiction des points d'intérêts anatomique du poisson de chaque modèle, la métrique de performance m est calculée comme suit :

$$m_i = \frac{d_i}{\text{Aire de la boite englobante}}$$

Avec d_i la distance euclidienne entre le point prédit i et le point labellisé correspondant. La standardisation par l'aire de la boite englobante du poisson où se trouve le point i permet de prendre en compte les différences de tailles des différentes espèces.

Le point qui correspond à l'extrémité postérieure du poisson (point 9) est le moins bien détecté pour les espèces BSC et PAR, et ce, pour tous les modèles (Figure 12). Ces deux classes correspondent aux espèces avec une queue en fourche. Les modèles M_8 M_9 et

M_{global} donnent globalement de meilleurs résultats. L'utilisation d'un modèle par espèce n'est donc pas une bonne solution pour détecter les points d'intérêts. Pour la suite de la méthode, ce sont les données de prédiction des modèles M_8 et M_9 qui seront utilisés.

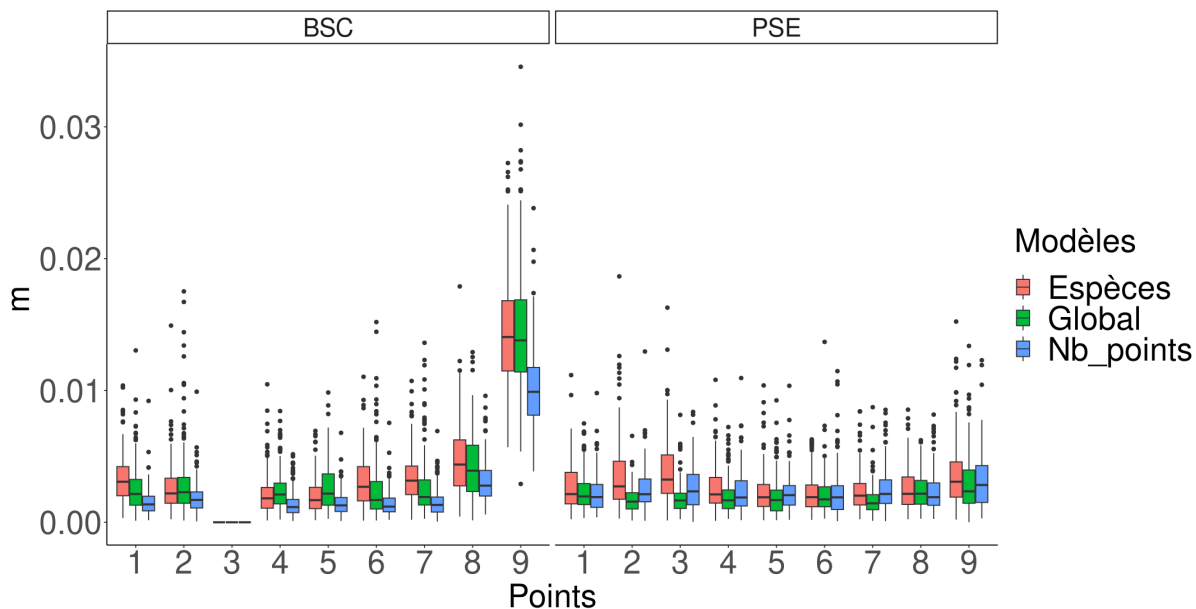


Figure 8 – Précision de la prédiction des points du poisson en fonction du type de modèle entraîné pour les classes BSC et PSE, les autres classes sont dans l'annexe A

5 Extraction des données d'intérêts halieutiques

Rappelons que les données d'intérêt halieutique que l'on cherche à déterminer automatiquement par analyse d'image sont la longueur totale du poisson et les distances morphométriques. La longueur totale du poisson est utilisée dans ce contexte pour l'analyse des traits d'histoire de vie des individus. Les distances morphométriques permettent de former un réseau de *truss* pour chaque individu et participent directement à l'identification des stocks.

Une fois que les modèles M_8 et M_9 sont entraînés, les variables prédites (en-dehors du code FAO) sont : la longueur (en pixels), la largeur (en pixels) et les coordonnées du centre de la boîte englobante ainsi que les coordonnées des points d'intérêt du poisson et de l'étiquette.

5.1 Longueur totale du poisson

5.1.1 Description des 3 méthodes comparées

Pour mesurer la longueur totale du poisson, trois méthodes ont été sélectionnées. Ces trois méthodes utilisent des variables prédites par les modèles M_8 et M_9 .

Méthode 1.9 : la distance entre le point 1 (mâchoire supérieure) et 9 (extrémité postérieure de la queue) est utilisée. Cette distance est obtenue en calculant en pixel la distance euclidienne entre les deux points. Elle est convertie en centimètre grâce à la distance mesurée sur l'étiquette donc on sait qu'elle fait 3 centimètres.

Méthode BB : comme deuxième proxy, la longueur de la boîte englobante est utilisée. Cette longueur s'exprime en pourcentage de la taille de l'image. Elle est également convertie en centimètre grâce à la distance mesurée sur l'étiquette.

Méthode LM : contrairement aux deux autres méthodes qui sont dérivées directement du réseau de neurones, pour cette deuxième méthode, on ajuste un modèle de régression multiple sur les distances prédites entre les point 1 à 8. De la même manière que pour la méthode 1.9 l'ensemble de ces distances sont converties en centimètre. Le modèle se définit comme suit :

$$Y_{ik} = \mu + \alpha_i + \sum_{j=1}^J \beta_j d_{ik}^j + \varepsilon_{ik}$$

Avec Y_{ik} la longueur totale prédite, μ représentant le terme constant, α_i l'effet de l'espèce i , et β_j l'effet de la distance j , d_{ik}^j la distance j de l'individu k de l'espèce i .

Pour déterminer la capacité de prédiction de la méthode LM sur un nouveau jeu de données, on utilise la statistique PRESS (*Prediction Error Sum of Squares* (Allen [1971])). En effet ayant utilisé les valeurs du jeu de données test pour construire le modèle linéaire, il ne reste plus de données neutres pour tester la capacité de prédiction du modèle. Pour ce faire, la statistique PRESS va, à chaque itération, mettre de côté un individu du jeu de donnée, redéfinir le modèle et calculer la valeur prédite de l'individu retiré. Ce qui permet de définir les résidus comme suit :

$$e_{(i)} = Y_i - \hat{Y}_{(i)}$$

Avec Y_i la longueur totale mesurée de l'individu i et $\hat{Y}_{(i)}$ la longueur totale prédite pour l'individu i quand il est exclu de la construction du modèle. Calculer l'ensemble des résidus est long et coûteux en ressource, ils peuvent également être définis par :

$$e_{(i)} = \frac{e_i}{1 - h_{ii}}$$

Avec h_{ii} la mesure d'influence de la mesure observé de l'individu i (Y_i) sur la mesure prédite ($\hat{Y}_{(i)}$)

5.1.2 Comparaison des performances des 3 méthodes

Pour évaluer et comparer ces trois méthodes, deux références de comparaison, issues de mesures réalisées manuellement par un opérateur, sont possibles : la longueur labellisée sur l'image ou la longueur mesurée sur le terrain. Afin de s'assurer que nos labellisations de la longueur des boîtes et la distance entre les points ne sont pas totalement éloignées de la réalité terrain, une différence moyenne est calculée pour chaque espèce (Tableau 5). Les individus ont également été répartis en 4 classes de taille : 0-20cm, 20-40cm, 40-60cm, 60-80cm.

Table 5 – *Différence moyenne en centimètre entre les mesures terrains et les mesures labellisées entre les points et sur la boîte englobante par espèces*

Classe de tailles	Points	Boîtes
0-20 cm	-0,541	-0,150
20-40 cm	-2,04	-1,42
40-60 cm	-5,12	-4,39
60-80 cm	-7,67	-6,47

On observe une tendance à sous-estimer la longueur totale labellisée par rapport à la longueur totale mesurée sur le terrain, aussi bien pour la distance entre les points que pour la longueur de la boîte englobante. De plus, cette sous-estimation augmente avec la taille des individus.

Nous avons ensuite examiné la différence moyenne entre la longueur prédite et la longueur labellisée par classe de taille (Tableau 6).

Table 6 – *Différence moyenne en centimètre entre les mesures labellisées et les mesures prédites par le modèle pour les 3 méthodes par classe de taille*

Classe de tailles	1.9	BB	LM
0-20 cm	-0,023	0,518	-0,081
20-40 cm	-0,233	0,408	-0,003
40-60 cm	-0,701	0,510	0,223
60-80 cm	-0,392	0,874	0,357

Nous remarquons qu'en utilisant la longueur de la boîte comme méthode, il y a une légère surestimation de la taille. À l'inverse, il y a une légère sous-estimation avec la longueur entre les points 1 et 9. Pour le modèle linéaire, cela dépend de la taille du poisson. Parmi les trois méthodes, c'est le modèle linéaire qui se rapproche le plus de la longueur labellisée.

Dans un second temps, lorsque nous examinons la différence moyenne entre la longueur totale prédite et les mesures prises sur le terrain (Tableau 7). Nous remarquons une sous-estimation des longueurs mesurées sur le terrain pour la majorité des classes de taille,

Table 7 – Différence moyenne en centimètre entre les mesures labellisées et les mesures prédites par le modèle pour les 3 méthodes par classe de taille

Classe de tailles	1.9	BB	LM
0-20 cm	-0,564	0,369	-0,226
20-40 cm	-2,270	-1,010	0,031
40-60 cm	-5,820	-3,880	0,347
60-80 cm	-8,060	-5,600	0,684

avec les méthodes BB et 1.9 . Cette sous-estimation s'accroît avec l'augmentation de la taille des individus. Cet effet n'est pas visible dans les résultats de la méthode LM.

5.2 Caractérisation automatique des mesures de morphométries

Au cours du projet DEMERSTEM, l'acquisition des mesures de morphométrie sélectionnées a été faite sur ImageJ (Schneider et al. [2012]). Afin d'évaluer la précision de la labellisation réalisée par nos opérateurs au cours de cette étude, une première comparaison entre nos labellisations et les mesures acquises par DEMERSTEM. Pour simplifier la démarche nous avons utilisé uniquement les données prédites disponibles pour l'espèce PAR (n = 83 individus).

5.2.1 Évaluer la précision de la labellisation

Afin de déterminer si la précision des mesures réalisées par le modèle est satisfaisante, une première comparaison est effectuée entre les distances obtenues par notre labellisation et celles de l'équipe de DEMERSTEM. Cette différence servira de référence pour évaluer les prédictions du modèle.

L'ensemble des 14 distances de ces individus mesurées sur ImageJ a été comparé aux mêmes distances labellisées via notre application. La RMSE (*Root Mean Square Error*) calculée pour l'ensemble de ces distances est de **0,4 cm**. Il y a donc une différence d'environ 0,4 centimètre entre notre labellisation et celle réalisée par l'équipe de DEMERSTEM.

5.2.2 Longueur prédite

Nous avons ensuite calculé la distance euclidienne entre les points d'intérêt prédits par le modèle M_8 et M_9 pour obtenir les 14 distances composant le *truss network* du poisson. Ces distances ont ensuite été converties en centimètres. Ce jeu de données prédit a été comparé au jeu de données labellisé par nos soins. La RMSE calculée pour l'ensemble de ces distances est de **0,2 cm**. L'erreur du modèle est donc inférieure à la variabilité humaine dans notre cas.

6 L'approche automatique conduit elle aux mêmes conclusions sur l'identification entre stocks ?

Au cours du projet DEMERSTEM, une fois que les distances morphométriques ont été mesurées sur les images de poissons, elles ont été analysées statistiquement grâce à des analyses multivariées. L'objectif est de caractériser chaque stock d'un point de vue morphologique. Concernant l'espèce *pagellus bellotti*(PAR), actuellement un seul et même stock est défini sur quatre pays : Côtes d'Ivoire, Ghana, Togo et le Bénin. L'échantillonnage n'a cependant été réalisé qu'au Ghana et en Côte d'Ivoire, sur 2 zones différentes par pays. Le but de ces analyses multivariées est de savoir si à partir des individus de ces deux pays, on retrouve un seul stock défini en fonction de l'analyse morphologique ou plusieurs.

L'équipe de DEMERSTEM a analysé 221 individus de l'espèce *pagellus bellotti*(PAR). Notre jeu de données est composé des données prédites du jeu de données test du modèle M_8 , filtrées pour inclure uniquement les individus de l'espèce *pagellus bellotti* et capturés dans les mers de la Côte d'Ivoire et du Ghana ce qui représente 83 individus. Cette différence importante d'individus est due à la nécessité de subdiviser le jeu de données pour l'apprentissage profond.

Les tableaux 8 et 9 détaillent le nombre d'individus par port de pêche, la taille moyenne des poissons de chaque groupe, pour le jeu de données de DEMERSTEM (8) et le nôtre (9).

Table 8 – Nombre d'individus de *pagellus bellotti* (PAR) analysés et intervalle de longueur totale (L_t) par zone de pêche étudiée du jeu de données de DEMERSTEM.

Zone de pêche	Effectif	Longueur moyenne	Ecart-type
Côte d'Ivoire Ouest (C.Iv_0)	31	17	2,26
Côte d'Ivoire Est (C.Iv_E)	69	17	2,80
Ghana Ouest (Gha_0)	60	17	1,41
Ghana Est (Gha_E)	60	20	0,97

Table 9 – Nombre d'individus de *pagellus bellotti* (PAR) analysés et intervalle de longueur totale (L_t) par zone de pêche étudiée de notre jeu de données.

Zone de pêche	Effectif	Longueur moyenne	Ecart-type
Côte d'Ivoire Ouest (C.Iv_0)	8	18	0,99
Côte d'Ivoire Est (C.Iv_E)	16	17	2,46
Ghana Ouest (Gha_0)	21	18	1,38
Ghana Est (Gha_E)	38	19	2,03

6.1 Transformation des données

Avant toute analyse statistique les données sont transformées. Pour prendre en compte l'effet de tailles des poissons, les données ont été standardisées selon cette équation :

$$Dtrans_i = \log D_i - b(\log Lt_j - \log Lmean)$$

Avec D la distance i mesurée en centimètres, Lt_j la longueur totale de l'individu j et $Lmean$ la longueur totale moyenne de tous les poissons du jeu de données. Pour le jeu de données de DEMERSTEM, Lt_j représente la longueur à la fourche du poisson, tandis que dans notre jeu de données, Lt_j correspond à la longueur mesurée avec la méthode LM.

6.2 Démarche statistique et comparaison des résultats

L'ensemble des analyses a été réalisé sur le jeu de données de DEMERSTEM et sur les données prédites, afin de comparer les résultats. Le but de cette partie est de déterminer si les prédictions du modèle permettent d'obtenir des résultats équivalents sur au moins une partie de la méthode statistiques.

Dans un premier temps, une analyse discriminante linéaire est réalisée, elle consiste à identifier les combinaisons linéaires de variables qui séparent au mieux les groupes. Autrement dit dans notre contexte, ils ont cherché à déterminer si on peut identifier l'origine d'un poisson à partir des mesures morphométriques. Ces combinaisons linéaires correspondent aux fonctions discriminantes ; chaque individu possède donc un score pour chacune de ces fonctions. C'est ce score qui permet ensuite de classer les poissons dans les différentes zones de pêche.

Dans un second temps, ils ont cherché à savoir si les groupes formés grâce aux fonctions discriminantes sont significativement différents. Pour ça, une ANOVA a été réalisée suivie d'un test post-hoc sur les scores de chaque individu par groupe obtenu grâce à la première fonction discriminante.

6.2.1 Analyse discriminante linéaire (LDA)

Les deux premières fonctions discriminantes pour chaque jeu de données sont représentées sur la figure 9. La première fonction discriminante explique 82,0% de la variance intergroupe pour le jeu de données de DEMERSTEM et 79,7% pour le nôtre. Sur la figure 9 A on peut voir deux groupes se former entre la Côte d'Ivoire et le Ghana, mais les zones au sein d'un même pays ne semble pas se distinguer. À l'inverse pour notre jeu de données, on ne peut pas distinguer de groupes différents.

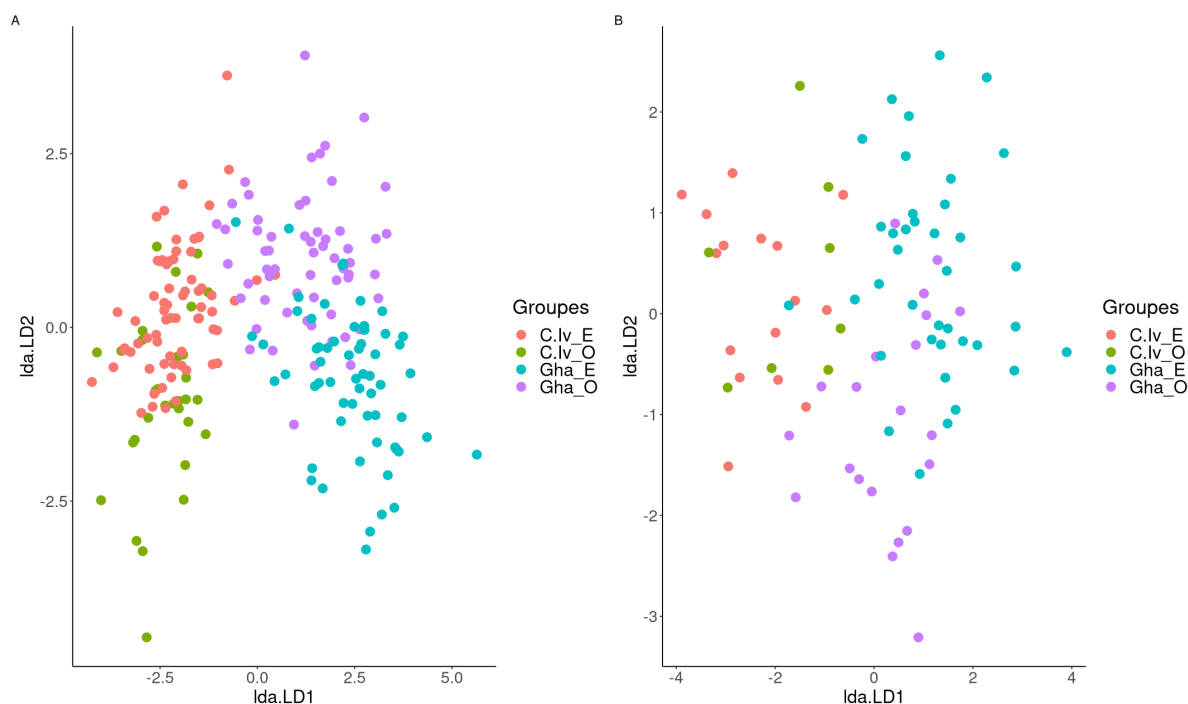


Figure 9 – Diagramme de dispersion sur le jeu de données de DEMERSTEM (A) et le notre (B), des deux premières fonctions discriminantes de l’analyse discriminante linéaire (LDA) pour les distances morphométriques de *P. bellottii* (PAR) : Côte d’Ivoire et Ghana, par zones (Ouest et Est).

Une fois le modèle construit, des prédictions sont réalisées sur les mêmes jeux de données pour vérifier la classification. La figure ?? représente les deux matrices de confusion. 80% des individus sont bien classés dans le cas de DEMERSTEM, et 77% pour notre jeu de données. Sur la figure ?? A, on distingue assez bien les deux groupes qui se formaient sur le diagramme de dispersion. Avec notre jeu de données, les groupes sont moins nettement défini, cependant un groupe pour le Ghana se distingue.

6.2.2 ANOVA

Le score de chaque individu pour ce premier facteur est calculé et tracé par groupe de pêche sur la figure 11.

Les tests sur les scores des individus pour la première fonction discriminante n’ont montré aucune différence significative ($P > 0,05$) entre les individus de l’ouest et de l’est de la Côte d’Ivoire, suggérant qu’ils peuvent appartenir à un même stock. Cependant, les deux zones du Ghana ont montré des valeurs proches mais significativement différentes ($P < 0,001$) entre elles, et nettement distinctes de celles de la Côte d’Ivoire. Les résultats sont les mêmes pour les deux jeux de données.

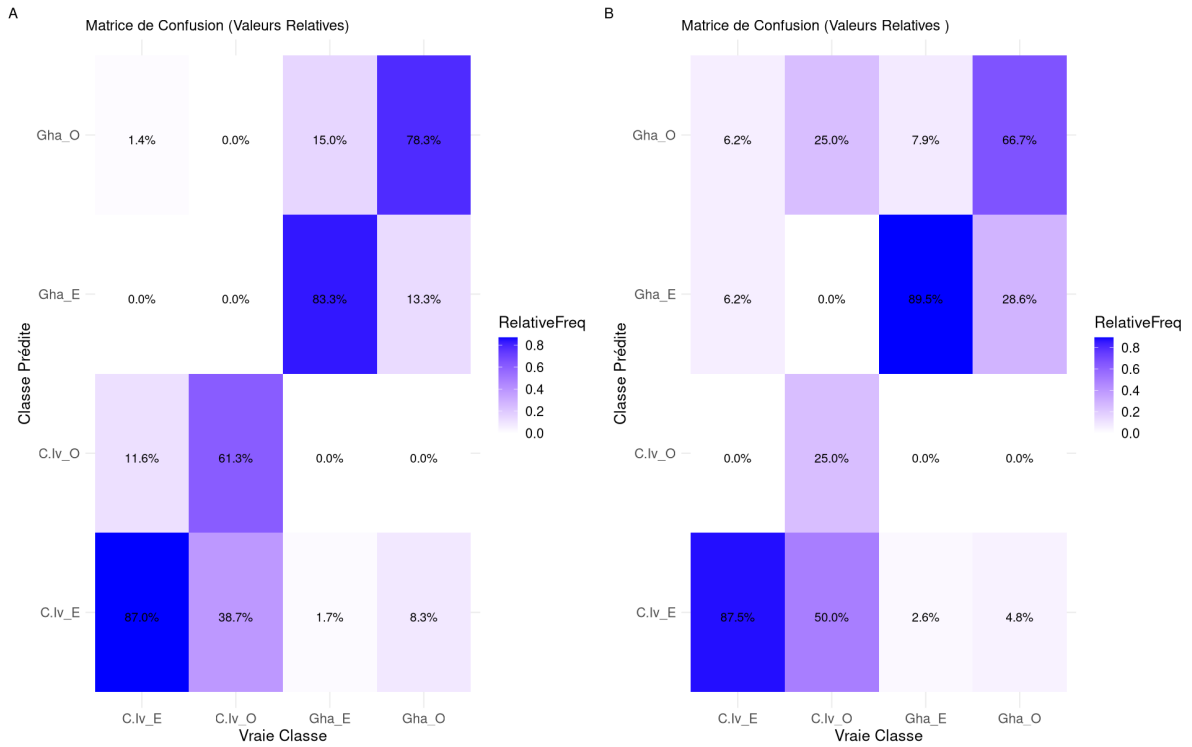


Figure 10 – Classification des individus de *P. bellottii* (PAR) (en %) du jeux de données DEMERSTEM (A) et du notre (B), dans leur population d’origine en utilisant la matrice de classification de l’Analyse Discriminante Linéaire (LDA) basée sur la morphométrie des poissons.

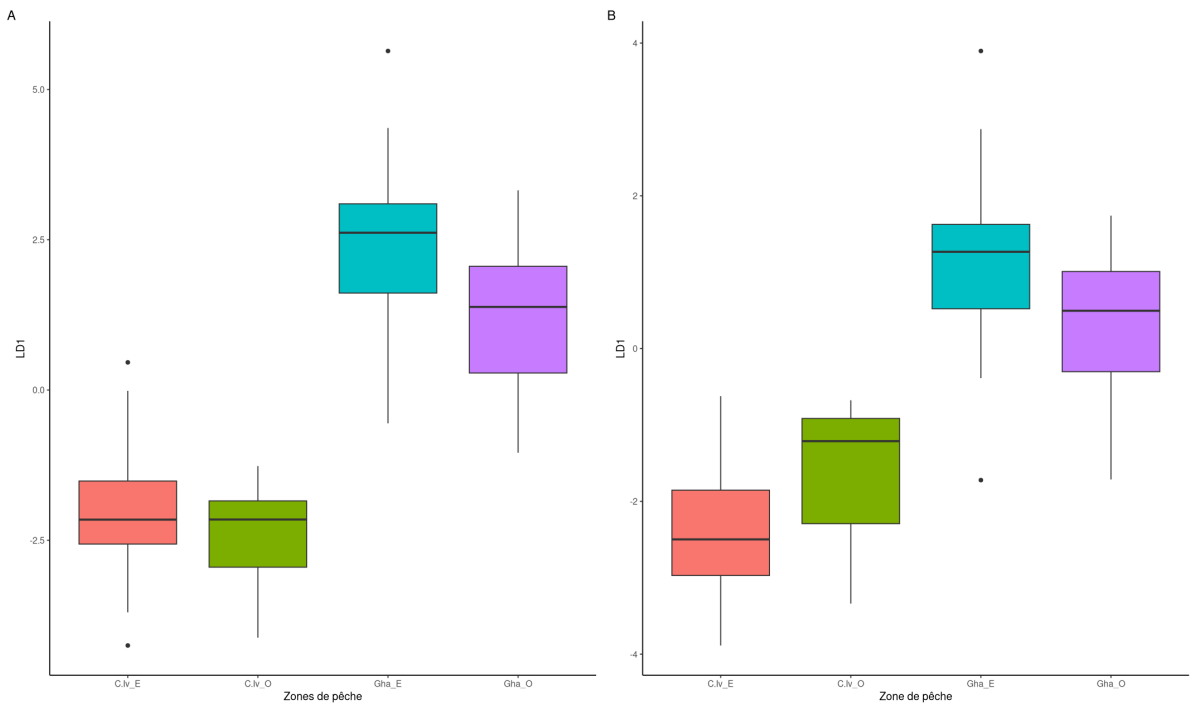


Figure 11 – Scores de l’axe 1 de l’Analyse Discriminante Linéaire (LDA) basée sur la morphométrie des individus *P. bellottii* (PAR) du jeux de données de DEMERSTEM (A) et le notre (B)

7 Discussion

Au cours de cette étude, nous nous sommes demandé s'il était possible de remplacer l'expertise humaine par celle d'un ordinateur pour une partie du protocole du projet DEMERSTEM. La première étape consistait à détecter automatiquement l'espèce du poisson sur des photographies prises au port lors des débarquements, à mesurer sa longueur totale et, enfin, à établir le profil morphologique des poissons en mesurant des longueurs préalablement choisies. La deuxième étape consistait à vérifier la robustesse des données en comparant les résultats des données collectées par DEMERSTEM avec ceux prédits par les modèles M_8 et M_9 obtenu avec la même méthode statistique. Cela permet de vérifier que l'utilisation de données prédites n'empêche pas de retrouver les mêmes conclusions.

Concernant le premier objectif d'automatisation, à savoir l'identification de l'espèce, il est difficile d'obtenir de meilleurs résultats que ceux fournis par l'architecture YOLOv8, qui affiche une précision de 99,98%. Cette performance, plus que satisfaisante, souligne l'efficacité de l'architecture YOLOv8 pour cette tâche spécifique. Cependant, le fait qu'une architecture développée manuellement ait également obtenu de bons résultats (94,41%) indique que la tâche d'identification n'est pas excessivement complexe, notamment parce qu'elle se limite à seulement cinq espèces. De plus, malgré des poissons mesurant jusqu'à 8 cm, aucune influence de la taille du poisson n'a été observée sur les performances du modèle pour distinguer les espèces. Cela vaut également pour la variabilité des photos (résolution, couleur du fond du bac, etc.). Dans ce contexte, on peut considérer que l'objectif d'identifier l'espèce par analyse d'image est atteint.

Pour la mesure de la longueur totale, les trois méthodes utilisées présentent des biais. La méthode BB, qui repose sur la longueur de la boîte englobante, est limitée par l'orientation du poisson, étant moins pertinente lorsque le poisson n'est pas horizontal. La méthode 1.9 pose un problème pour les poissons à queue en fourche, car le point 9 à l'extrémité postérieure du poisson est mal détecté. Cette limite est due en partie au fait que YOLOv8-pose optimise le placement des points en fonction de l'ensemble des points et non point par point. Pour améliorer la précision du placement des points, il serait intéressant d'explorer une autre architecture qui optimise potentiellement le placement des points individuellement, ou de modifier YOLOv8-pose en ce sens. D'autre part, cette limite met en évidence une erreur de choix de labellisation en plaçant ce point 9 aléatoirement sur une des deux extrémités de la queue. Il conviendrait de choisir un point d'intérêt plus prédictible pour les espèces à nageoire caudale en fourche, comme le creux de la fourche, qui est bien détecté par d'autres modèles (Yu et al. [2023]; Liu et al. [2024]). La méthode LM, contrairement aux deux autres, nécessite un deuxième

ajustement sur les données. Il serait donc intéressant de tester cette méthode sur un jeu de données différent pour s'assurer de sa généralisation.

Pour les trois méthodes, le modèle parvient à mesurer automatiquement une taille proche de la longueur totale labellisée par les opérateurs lors de la phase de labellisation. Cependant, la mesure automatique diffère des mesures effectuées sur le terrain, montrant une sous-estimation par rapport aux mesures terrain et aux mesures labellisées. Cette sous-estimation peut s'expliquer par deux facteurs principaux. Premièrement, lorsque l'individu est grand, l'étiquette placée sur le poisson se rapproche de l'objectif de l'appareil photo, augmentant ainsi sa taille en pixels et biaisant la conversion des pixels en centimètres. Deuxièmement, il pourrait y avoir un biais humain tendant à surestimer la taille du poisson lorsque l'individu est plus grand lors de mesures sur le terrain. Il serait pertinent d'explorer cet effet plus en profondeur pour mieux comprendre ces biais et améliorer la précision des mesures. Toutefois, la méthode LM reste la plus performante pour la prédiction de la longueur totale, avec une précision de $\pm 1,15$ cm.

Dans l'optique d'automatiser une partie du protocole, notre étude a mis en évidence des ajustements et des améliorations à apporter à l'acquisition des images. Tout d'abord, il est important de placer l'étiquette à côté du poisson et non sur lui. Ensuite, il est préférable d'éviter autant que possible l'utilisation d'étiquettes faites à la main, et d'utiliser la même étiquette pour toutes les images. Bien que cela induise une spécialisation et limite la généralisation du modèle pour la détection de l'étiquette, cela permet également de rendre cette détection plus précise et permet donc d'estimer plus précisément la taille des poissons sur les photos.

En ce qui concerne l'identification des stocks, le troisième objectif d'automatisation était de mesurer les distances morphologiques des poissons en formant un réseau de *truss* pour chaque individu. Ces distances ont été mesurées grâce à la prédiction de points par les modèles M_8 et M_9 , qui correspondent aux modèles entraînés en fonction du nombre de points d'intérêt à détecter. Un placement précis des points est essentiel pour garantir l'exactitude des distances mesurées. Cependant, il existe une variabilité humaine, ou un effet opérateur, dans le placement des points labellisés et donc dans la mesure des distances. Cette variabilité est de 0,4 cm entre les opérateurs de DEMERSTEM et ceux de notre étude, ce qui peut s'expliquer par une différence d'expertise et de méthodes.

La méthode statistique de DEMERSTEM a été appliquée sur les données prédites par les modèles M_8 et M_9 ainsi que sur les données mesurées par les opérateurs de DEMERSTEM. On observe certaines différences dans la composition des groupes (Figure 9 et 10). Cela peut d'une part s'expliquer par le fait que l'analyse n'est pas réalisée sur les mêmes individus et surtout, il y a une différence importante dans le nombre d'individus par groupe. C'est d'autant plus vrai que la LDA est sensible à la proportion d'indivi-

des par groupes qui ici n'est pas équivalente entre les deux jeux de données. Ce manque d'individus du côté du jeu des données prédites induit un manque d'information significatif pour construire les fonctions discriminantes de la LDA. Cependant, et malgré ces limites, l'analyse par ANOVA identifie de façon similaire, dans les deux jeux de données, deux groupes distincts : un groupe Côtes d'Ivoires et un groupe Ghana. Ces résultats étant très encourageants pour une seule des cinq espèces étudiées, il serait intéressant de compléter l'analyse avec les 4 autres espèces. En particulier, l'espèce *Pagrus caeruleostictus*, qui dispose d'un plus grand nombre d'images, permettrait de vérifier si le manque de précision observé dans nos résultats avec *Pagellus belottii* est dû à un manque d'information ou à d'autres facteurs.

8 Conclusion et perspectives

Malgré l'effort de recherche important autour de l'automatisation de tâches répétitives en écologie par méthode d'apprentissage profond, la valorisation d'objectifs opérationnels de ces méthodes est moins présente. Au cours de notre étude, la possibilité d'automatiser les parties ciblées du protocole de DEMERSTEM a été étudiée. Les résultats qui en résultent sont encourageants malgré leurs limites. Pour une application concrète de notre méthode développée, il serait intéressant d'alimenter l'étude avec de nouvelles architectures pour augmenter la précision des longueurs prédites. YOLO-NAS Pose par exemple qui est développée par Deci en 2023 () donne de meilleurs résultats pour l'estimation de la pose des points d'intérêt, elle serait donc intéressante à utiliser pour comparaison. Cette étude a également servi de preuve de concept pour l'automatisation de protocoles impliquant le placement de points d'intérêt sur une image, de la labellisation à l'utilisation du modèle, en passant par l'application développée. L'automatisation de ces tâches permettrait non seulement d'accroître le volume de données à analyser, mais également de le faire en minimisant les effets opérateurs relatifs à l'identification des espèces, à la mesure des individus ou à la saisie des données.

Références

- Allen, D. M. (1971). Mean Square Error of Prediction as a Criterion for Selecting Variables. *Technometrics*, 13(3) :469–475. Publisher : Taylor & Francis _eprint : <https://www.tandfonline.com/doi/pdf/10.1080/00401706.1971.10488811>.
- Begg, G. A., Friedland, K. D., and Pearce, J. B. (1999). Stock identification and its role in stock assessment and fisheries management : an overview. *Fisheries Research*, 43(1) :1–8.
- Chang, W., Cheng, J., Allaire, J., Xie, Y., and McPherson, J. (2023). *shiny : Web Application Framework for R*. R package version 1.7.4.
- Deci (2023). `super-gradients/YOLONAS-POSE.md` at master · Deci-AI/super-gradients.
- FAO (2024). *The State of World Fisheries and Aquaculture 2024*. FAO.
- Furuta, N., Tanaka, T., and Komeyama, K. (2021). Method for automating fish-size measurement and camera calibration using a three-dimensional structure and an optical character recognition technique. , 87 :100–107.
- García-Isarch, E., Landa, J., González, J., Teresa, M., Santamaría, G., Pérez, M., and Muñoz, E. (2020). Demerstem protocoles d'Échantillonnage biologique. Technical report.
- García-Isarch, E., Landa, J., Pérez, M., Partida, B., Thiaw, M., Meissa, B., Diallo, I., Tapé, J., Barri, I., Jallow, M. S., Bakari, C., Czerwinski, I., Camara, O. I., Ansong, E., and Johnstone, C. (2023). STOCK IDENTIFICATION : ADVANCING IN THE KNOWLEDGE OF STOCK STRUCTURE AS A REQUIREMENT FOR STOCK ASSESSMENT. Technical report.
- Huang, T.-W., Hwang, J.-N., and Rose, C. S. (2016). Chute based automated fish length measurement and water drop detection. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1906–1910, Shanghai. IEEE.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11) :2278–2324. Publisher : Institute of Electrical and Electronics Engineers.
- Liu, Y., Fu, Z., Song, L., Sui, H., and Li, B. (2024). Improved YOLOv8-Pose Algorithm for Albacore Tuna (*Thunnus alalunga*) Fork Length Extraction and Weight Estimation. *Journal of Marine Science and Engineering*, 12(5) :784.

- Maji, D., Nagori, S., Mathew, M., and Poddar, D. (2022). YOLO-Pose : Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 2636–2645, New Orleans, LA, USA. IEEE.
- Marini, M., Suman, A., Farajallah, A., and Wardiatno, Y. (2017). Identifying *Penaeus merguensis* de Man, 1888 stocks in Indonesian Fisheries Management Area 573 : a truss network analysis approach. 10(4).
- Miele, V., Dray, S., and Gimenez, O. (2021). R95 : Images, écologie et deep learning.
- Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., and Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*, 115(25) :E5716–E5725. Publisher : Proceedings of the National Academy of Sciences.
- Pawson, M. G. and Jennings, S. (1996). A critique of methods for stock identification in marine capture fisheries. *Fisheries Research*, 25(3) :203–217.
- Rawat, S., Benakappa, S., and Kumar, J. (2017). Identification of fish stocks based on Truss Morphometric : A review.
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). Nih image to imagej : 25 years of image analysis. *Nature Methods* 2012 9 :7, 9 :671–675.
- Simonyan, K. and Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv :1409.1556 [cs].
- Strauss, R. E. (1982). THE TRUSS : BODY FORM RECONSTRUCTIONS IN MORPHOMETRICS¹2. *SYSTEMATIC ZOOLOGY*, 3.
- Turan, C. (2004). Stock identification of Mediterranean horse mackerel (*Trachurus mediterraneus*) using morphometric and meristic characters. *ICES Journal of Marine Science*, 61(5) :774–781.
- Ultralytics (2023). Yolov8. <https://github.com/ultralytics/ultralytics>. Accessed : 2023-06-03.
- Villon, S., Mouillot, D., Chaumont, M., Darling, E. S., Subsol, G., Claverie, T., and Villéger, S. (2018). A Deep learning method for accurate and fast identification of coral reef fishes in underwater images. *Ecological Informatics*, 48 :238–244.

- Willi, M., Pitman, R. T., Cardoso, A. W., Locke, C., Swanson, A., Boyer, A., Veldthuis, M., and Fortson, L. (2019). Identifying animal species in camera trap images using deep learning and citizen science. *Methods in Ecology and Evolution*, 10(1) :80–91.
_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.13099>.
- Yu, Y., Zhang, H., and Yuan, F. (2023). Key point detection method for fish size measurement based on deep learning. *IET Image Processing*, 17(14) :4142–4158.
- Álvarez Ellacuría, A., Palmer, M., Catalán, I. A., and Lisani, J.-L. (2020). Image-based, unsupervised estimation of fish size from commercial landings using deep learning. *ICES Journal of Marine Science*, 77(4) :1330–1339.
- Ärje, J., Melvad, C., Jeppesen, M. R., Madsen, S. A., Raitoharju, J., Rasmussen, M. S., Iosifidis, A., Tirronen, V., Gabbouj, M., Meissner, K., and Høye, T. T. (2020). Automatic image-based identification and biomass estimation of invertebrates. *Methods in Ecology and Evolution*, 11 :922–931.

A Annexe A

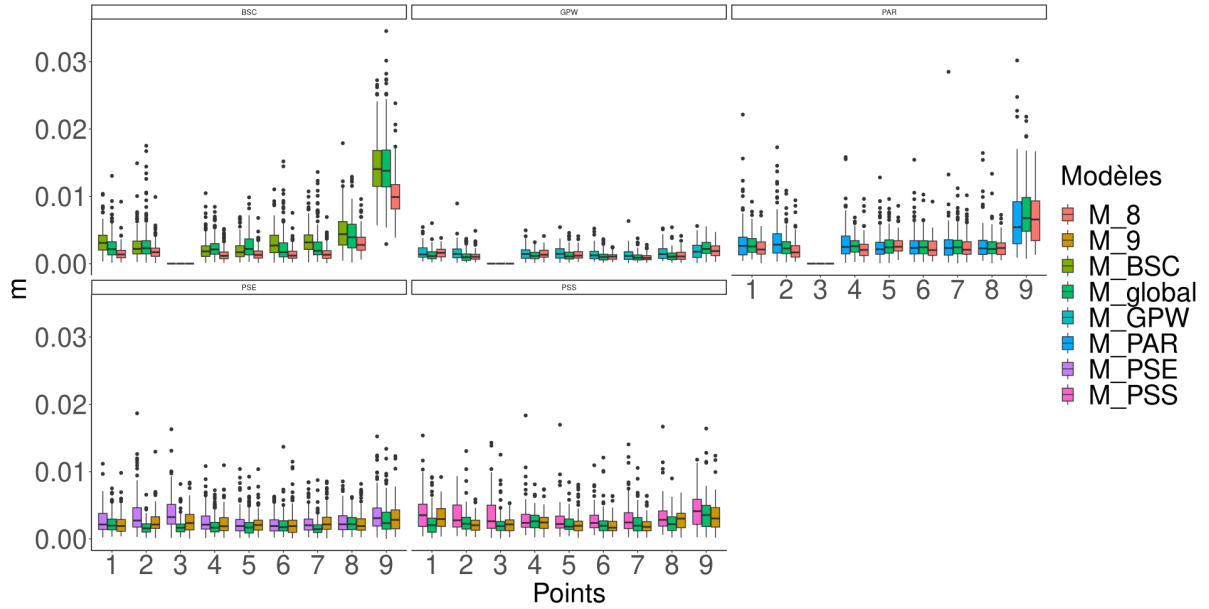


Figure 12 – Précision de la prédiction des points du poisson en fonction du type de modèle entraîné pour les 5 classes

Master de Modélisation en écologie
de l'Université de Rennes 1
2023-2024
Romane Dubois

Résumé L'identification des stocks est un enjeu majeur pour une régulation des pêches durables. Dans le contexte du projet DEMERSTEM, un protocole important a été mis en place sur plusieurs pays, demandant l'expertise de plusieurs collaborateurs. La question se pose donc, est-ce qu'on peut remplacer cette expertise humaine par celle d'une machine? Notre étude a permis de montrer une identification automatique de l'espèce avec une précision moyenne de 99,89% en utilisant l'architecture de YOLOv8. Pour la prédiction des points clés, YOLOv8-pose est utilisé. La longueur totale est mesurée à $\pm 1,15$ cm et la RMSE entre les distances morphométriques prédites et labellisée est 0,2 cm. L'analyse statistique de ces distances prédites, permet de retomber sur des similitudes importantes avec les résultats obtenus avec les mêmes analyses menées par l'équipe de DEMERSTEM. Les résultats de cette étude sont encourageants quant à l'automatisation de protocoles utilisant l'identification d'espèce, la mesure de longueurs et le placement de point d'intérêt.

Mots-clés : Apprentissage profond, Taxonomie, Morphométrie, Points d'intérêt, Stock

Abstract Stock identification is a major issue for the sustainable regulation of fisheries. In the context of the DEMERSTEM project, an important protocol has been used in several countries, which need the expertise of numerous collaborators. The question then arises : can we replace this human expertise with a machine? Our study has shown that automatic species identification with an average accuracy of 99.89% with the YOLOv8 architecture. For the prediction of the keypoints, the YOLOv8-pose architecture is used. The total length is measured at ± 1.15 cm, and the RMSE between the predicted and labeled morphometric distances is 0.2 cm. Statistical analysis of these predicted distances shows significant similarities with the results obtained by the DEMERSTEM team. The results of this study are encouraging for the automation of protocols using species identification, length measurement, and the placement of keypoints.

Keywords : Deep learning, Taxonomy, Morphometry, Keypoints, Stock