Rapport de stage de fin d'études

Une approche combinant vision par ordinateur et apprentissage profond pour extraire les tailles des poissons à partir d'images de caisses débarquées en criée

Du 03/02/2024 au 03/08/2025

Réalisé par :

Ourida Ghena

Encadré par :

- Tuteur en entreprise : Mr. Jérôme Guitton, Mr .Youen Vermard, Mr. François Martignac,Mme. Marie-Pierre Etienne
- Tuteur académique : Mr. Schneider Dominique

Année universitaire : 2024/2025

Remerciements

Je tiens tout d'abord à remercier chaleureusement Jérôme Guitton, mon encadrant de stage, pour son accompagnement tout au long de cette expérience au sein de l'Institut Agro Rennes. Sa disponibilité, ses conseils clairs, et sa bienveillance ont grandement facilité mon travail et contribué à mes progrès, tant techniques que méthodologiques.

Je remercie également François Martignac, Youen Vermard et Marie-Pierre Etienne, pour leur disponibilité, leurs retours éclairés, et leur implication dans le bon déroulement de ce stage.

Un mot également pour l'ensemble de l'équipe de l'institut Agro de l'UMR DECOD, pour leur accueil, et l'environnement de travail agréable.

Un remerciement tout particulier s'adresse à Dominique Schneider, ainsi qu'à l'ensemble de mes professeurs, pour leur précieux encadrement tout au long de mon parcours académique .

Enfin, je remercie de tout cœur mes parents, mes amis, et toutes les personnes qui m'ont soutenu moralement ou intellectuellement pendant cette période. Leur présence et leur confiance m'ont été précieuses.

Résumé

Ce rapport présente un travail réalisé dans le cadre d'un stage de fin d'études de Master 2 à l'Institut Agro Rennes-Angers. Il s'inscrit dans une dynamique de recherche appliquée visant à développer des méthodes automatiques d'estimation de la taille des poissons à partir d'images capturées en criée.

L'objectif principal est de produire, à partir de données visuelles existantes, des distributions de taille individuelles fiables, contribuant ainsi à améliorer la connaissance biologique des stocks halieutiques. L'approche adoptée se veut non intrusive, automatisée et compatible avec les conditions réelles d'acquisition. Ces images, initialement destinées à la vente en ligne, sont prises selon un protocole standardisé d'un point de vue commercial, mais sans contrainte scientifique formelle. Elles présentent donc une grande variabilité dans la disposition des individus, la luminosité.

Le travail s'est concentré sur le cas du maigre du golfe de Gascogne, un stock classé parmi les stocks à données limitées et ne faisant pas l'objet d'évaluations régulières. Il s'agissait d'estimer les tailles des poissons à partir de ce qui est effectivement visible sur les images, en développant un pipeline combinant vision par ordinateur et apprentissage profond.

Ce rapport décrit la démarche adoptée, les choix méthodologiques, ainsi que les premiers résultats obtenus. Il discute également les limites rencontrées et propose des perspectives d'amélioration et d'application.

Sommaire

1. Introduction 1
2. Contexte de travail 1
2.1. Le laboratoire UMR DECOD 1
2.2. L'Institut Agro Rennes-Angers 2
2.3. Le projet de stage 2
3. Etat de l'art 2
3.1. Deep Learning 2
3.1.1. Couches de convolution : 3
3.1.2. Couches de pooling : 3
3.2. Segmentation et détection de points clés 4
3.3. Estimation de la longueur des poissons 7
4. Présentation et analyse des données 7
4.1 Données 8
4.1.1. Base de données des débarquements quotidiens 8
4.1.2. Images des caisses de poissons 9
5. Méthodologie 9
5.1 Estimation de la longueur des poissons 9
5.1.1. Estimation directe par segmentation d'instance et extraction du squelette 10
5.1.2. Estimation indirecte par détection de la tête 11
5.1.3. Modélisation statistique de la longueur à partir des repères morphométriques 12
5.1.4 Conversion des longueurs en valeurs métriques 13
5.1.5 Constitution des données d'évaluation 14
6. Résultats et analyse 16
6.1 Analyse comparative des méthodes d'estimation du facteur de conversion 16
6.2. Corrélation entre les mesures céphalique et la longueur totale 17
6.3. Évaluation des performances de la chaîne de traitement 17
6.3.1. Évaluation de la détection automatique des têtes 17
6.3.2 Évaluation de la stabilité et de la précision des mesures céphaliques 18
Figure 6.3 : Variabilité des erreurs de détection selon la mesure céphalique 19
6.4. Analyse des performances des estimations de longueur totale 19
6.4.1. Évaluation de l'estimation de la longueur totale à partir du squelette 20
Figure 6.5 – Performance de l'estimation squelettique 21
6.4.2 Évaluation des longueurs estimées à partir des mesures céphaliques 21

6.5 Distribution des tailles..... 22
6.5.1. Extrapolation des tailles..... 23
7. Discussion générale..... 24
8. Conclusion..... 25
Bibliographie..... 26

Tables des figures

Figure 3.1 : Architecture d'un réseau neuronal convolutif..... 4

Figure 3.2 : Exemples de tâches en vision par ordinateur..... 5

Figure 3.3 : Exemple de détection de points clés sur un objet.....6

Figure 4.1 : Données descriptives associées aux caisses de poissons.....8

Figure 4.2 : Débarquement de poissons à la criée et poste de pesée.....9

Figure 5.1 : Visualisation de la mesure de la longueur du poisson en pixels.....11

Figure 5.2 : Les mesures morphométriques utilisées..... 13

Figure 5.2 : Méthodes de référence pour la conversion pixel/cm..... 14

Figure 5.3 : Exemple de sélection des poissons entièrement visible sur l'image..... 16

Figure 6.2 : Indicateurs d'erreur sur les mesures céphaliques..... 19

Figure 6.3 : Variabilité des erreurs de détection selon la mesure céphalique..... 19

Figure 6.5 – Performance de l'estimation squelettique.....21

Figure 6.6 : Indicateurs d'erreur sur l'estimation de la longueur totale à partir des mesures céphaliques.....22

Figure 6.7 : Distribution des tailles pour une catégorie commercial..... 23

1. Introduction

L'évolution rapide des technologies en intelligence artificielle et en vision par ordinateur ouvre aujourd'hui des perspectives inédites dans de nombreux secteurs professionnels, notamment dans le domaine halieutique, ces avancées permettent d'envisager l'automatisation de tâches jusqu'ici réalisées manuellement, telles que la reconnaissance d'espèces, la détection d'individus ou encore l'estimation de leur taille.

La quantification précise des tailles des poissons débarqués présente un enjeu fondamental pour les sciences halieutiques, notamment dans un contexte de gestion durable des ressources marines. Les tailles individuelles permettent de produire des indicateurs biologiques clés, comme la structure démographique, la maturité ou le potentiel reproducteur, qui alimentent les évaluations de stocks. Ces données sont généralement recueillies par des campagnes de mesures en mer ou à quai. Or, ces méthodes manuelles, bien que précises à petite échelle, sont difficilement transposables à grande échelle en raison de leur coût, de leur dépendance au personnel qualifié et de leur faible fréquence d'échantillonnage.

Dans ce contexte, la disponibilité croissante d'images produites en routine par les criées françaises dans le cadre de la vente en ligne représente une opportunité scientifique encore peu exploitée.

Ce projet vise à développer une méthode automatique d'estimation des tailles individuelles à partir d'images de caisses de débarquement. L'approche repose sur la combinaison d'outils de vision par ordinateur et d'apprentissage profond, dans le but de concevoir une solution robuste, adaptable et compatible avec les conditions réelles d'acquisition en criée.

2. Contexte de travail

2.1. Le laboratoire UMR DECOD

Créé en janvier 2022, l'UMR DECOD Dynamique et durabilité des écosystèmes de la source à l'océan rassemble des équipes issues de l'INRAE, de l'Ifremer et de l'Institut Agro Rennes-Angers. Ce collectif de recherche s'attache à étudier les milieux aquatiques, en s'appuyant sur l'observation, l'expérimentation

et la modélisation. L'objectif est de mieux comprendre la biodiversité et les dynamiques écologiques, afin de proposer des outils pour leur gestion durable. [3]

2.2. L'Institut Agro Rennes-Angers

L'Institut Agro Rennes-Angers est un établissement public d'enseignement supérieur et de recherche, spécialisé dans les domaines de l'agriculture, de l'alimentation et de l'environnement. Il dispose de plusieurs campus, dont celui de Rennes.

2.3. Le projet de stage

Le stage s'inscrit dans un projet porté par l'Ifremer. Il vise à explorer les capacités de l'intelligence artificielle à extraire automatiquement des informations biologiques, comme la taille des poissons, à partir d'images collectées en criée. Ces travaux s'insèrent dans une perspective plus large de valorisation de données issues de sources non scientifiques mais accessibles en routine, à des fins de suivi des stocks halieutiques.

3. Etat de l'art

3.1. Deep Learning

L'apprentissage profond est un domaine de recherche essentiel dans le traitement d'images, la vision par ordinateur, et la bioinformatique, entre autres. Cela a conduit au développement et à l'application de diverses techniques d'apprentissage profond dans ces domaines. En général, les techniques d'apprentissage profond sont classées en trois grandes catégories : les réseaux de neurones convolutifs (CNN), les machines de Boltzmann restreintes (RBM) et les auto-encodeurs. Dans cette section, nous nous concentrons sur les CNN, car ils constituent la base des modèles utilisés dans ce projet.[1][2]

L'architecture des CNN, illustrée à la Figure 2.1, se compose principalement de trois types de couches : les couches de convolution, les couches de pooling et les couches entièrement connectées. L'apprentissage du réseau peut être divisé en deux étapes : directe (forward) et rétroactive (backward). Dans l'étape forward, on classe d'abord l'image d'entrée en fonction des poids et des biais de chaque couche. Le coût de la perte est calculé à partir des données d'entrée à l'aide de la sortie prédite. Dans

l'étape backward, en fonction du coût de la perte mesuré, les gradients sont calculés pour chaque paramètre. Grâce à ces gradients, les paramètres sont ensuite mis à jour pour l'itération suivante. La procédure d'entraînement peut être arrêtée après un nombre satisfaisant de répétitions. Les fonctionnalités de ce réseau sont décrites comme suit :

3.1.1. Couches de convolution :

Dans cette couche, le CNN utilise de nombreux filtres pour effectuer une convolution sur l'image entière, y compris les cartes de caractéristiques intermédiaires, afin de générer différentes cartes de caractéristiques. [4]

3.1.2. Couches de pooling :

La couche de pooling réduit la dimension des cartes de caractéristiques issues des convolutions, ce qui permet de diminuer le nombre de paramètres et la complexité du modèle. Elle conserve les informations essentielles en sélectionnant les attributs dominants. Les deux méthodes courantes sont le max-pooling et l'average pooling.

3.1.3. Couches entièrement connectées :

Les couches entièrement connectées constituent la phase finale d'un réseau CNN. Elles reçoivent un vecteur aplati issu des cartes de caractéristiques générées par les couches de convolution. Chaque neurone est connecté à tous les neurones de la couche précédente, permettant l'apprentissage de combinaisons non linéaires entre les attributs extraits [5].

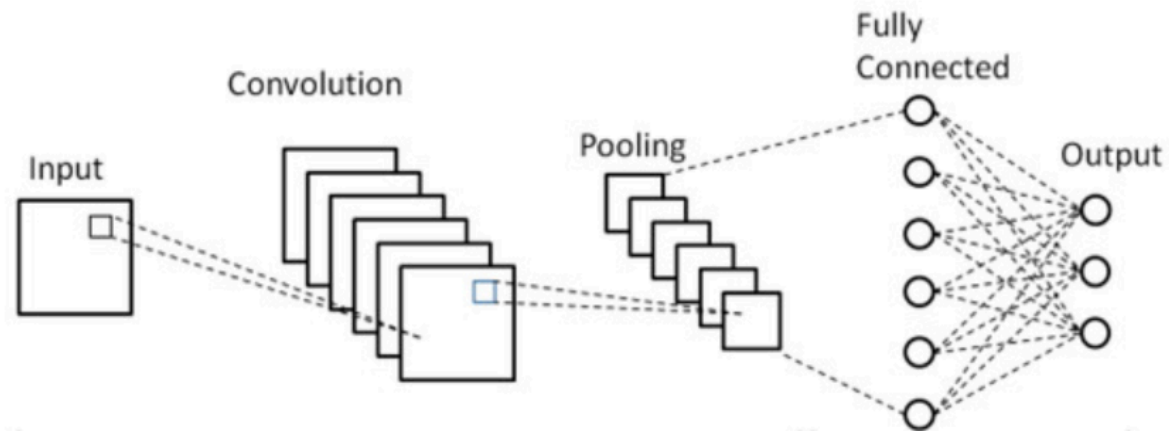

Figure 3.1 : Architecture d'un réseau neuronal convolutif

3.2. Segmentation et détection de points clés

L'apprentissage profond a connu une application importante dans les tâches de vision par ordinateur telles que la classification d'images, la détection d'objets, la recherche d'images et la segmentation d'images.

La détection d'objets ou localisation va au-delà de la classification d'image en identifiant non seulement les classes des objets présents sur une image, mais aussi leurs emplacements, généralement à l'aide de boîtes englobantes. La segmentation sémantique vise un niveau d'analyse plus précis en attribuant un label à chaque pixel d'une image, selon l'objet ou la région qu'il représente. La segmentation d'instances va encore plus loin en attribuant des labels différents à chaque instance d'un objet, même si elles appartiennent à la même classe, combinant ainsi la détection d'objets et la segmentation sémantique [6][7].

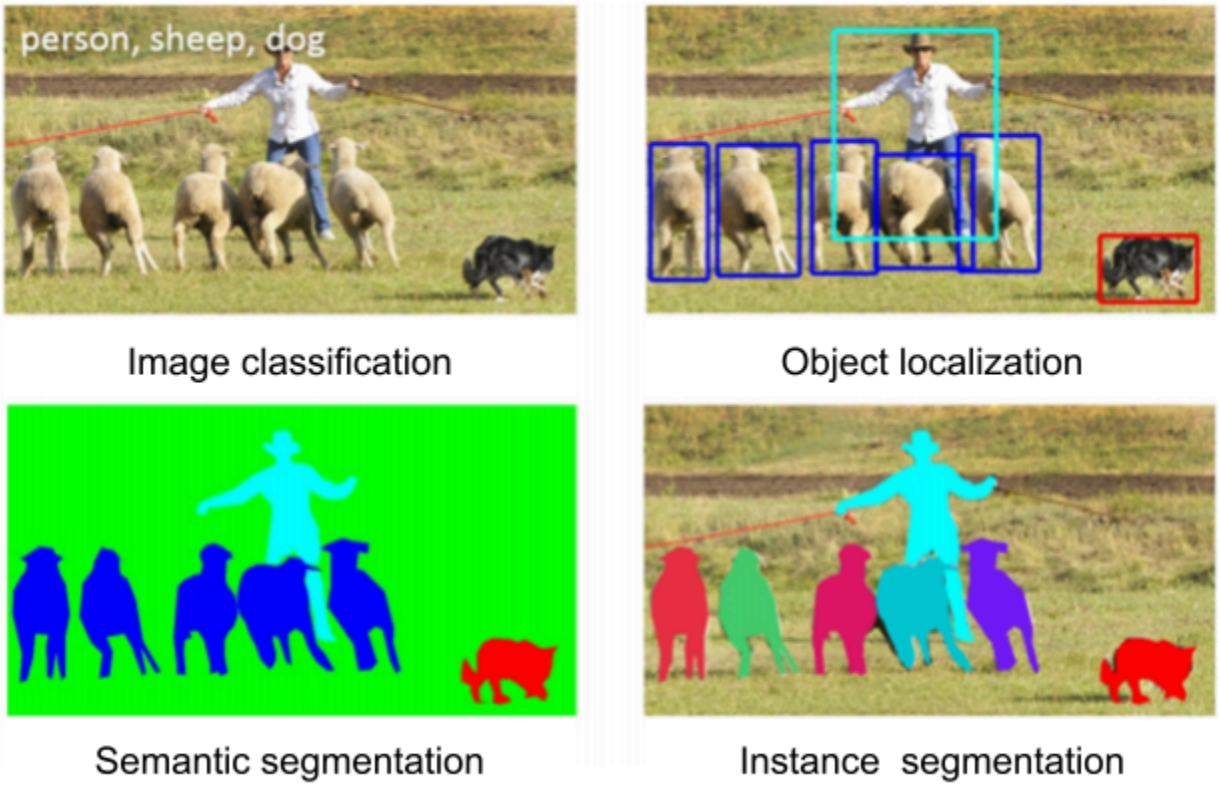

Figure 3.2 : Exemples de tâches en vision par ordinateur

La segmentation peut porter sur l'intégralité de l'image ou se concentrer sur des parties spécifiques telles que la tête ou la queue [8,9], voire sur chaque poisson individuellement. Dans le cas où plusieurs individus sont présents dans une même image, il est pertinent de considérer la tâche comme un problème de segmentation d'instances. Plusieurs approches avancées ont été développées pour cela, parmi lesquelles Mask R-CNN, SAM, et YOLO figurent parmi les plus connues.

Mask R-CNN est une méthode étendue de Faster R-CNN, intégrant un module supplémentaire pour prédire les masques de segmentation. Cette double capacité détection et segmentation précise des pixels est un outil adapté pour identifier chaque instance d'un objet. Toutefois, ce modèle est coûteux en ressources computationnelles et requiert des techniques telles que l'augmentation de données pour gagner en robustesse.

SAM est un modèle d'intelligence artificielle développé par Meta AI, conçu pour extraire n'importe quel objet à partir d'une image en un seul clic. Il s'agit d'un système de segmentation piloté par des instructions, capable de généraliser à des objets ou images inconnus sans apprentissage supplémentaire. Cependant, il repose fortement sur la sélection de boîtes englobantes pour effectuer la segmentation, ce qui ajoute une étape supplémentaire au processus et en augmente la complexité.

YOLO s'est imposé comme une solution prometteuse et efficace pour la détection et la segmentation d'objets. Réputé pour sa rapidité et sa précision, il surpasse les modèles classiques tels que R-CNN. Le principe de YOLO consiste à diviser une image en une grille et à prédire, [6, 12, 13] pour chaque cellule, la présence, la classe et la position des objets. En combinant un réseau de neurones convolutif pour l'extraction des caractéristiques avec des couches entièrement connectées pour la prédiction, YOLO permet une détection et une localisation en une seule étape, ce qui réduit les calculs redondants et accélère le traitement. La série a évolué jusqu'à YOLOv12, intégrant des techniques d'augmentation de données.

Dans certaines situations, il est nécessaire d'identifier des parties spécifiques d'un objet, dans notre cas, des régions caractéristiques sur le corps du poisson (œil, bouche, nageoire, etc.). Parmi les modèles les plus utilisés, Keypoint R-CNN étend l'architecture Mask R-CNN en ajoutant une branche dédiée à la régression des coordonnées des points clés. Ce modèle est particulièrement robuste, mais reste coûteux en ressources et en annotations. Récemment, les modèles de la famille YOLO ont été adaptés à la détection de points proposant une alternative plus légère et rapide, en intégrant directement la régression des keypoints dans leur tête de détection.

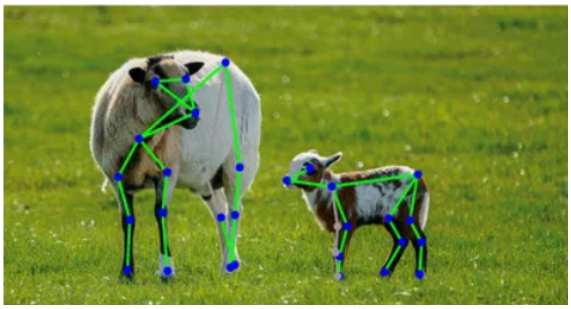

Figure 3.3 : Exemple de détection de points clés sur un objet

3.3. Estimation de la longueur des poissons

L'estimation de la longueur des poissons s'articule principalement autour de plusieurs approches, dépendant des conditions d'acquisition : stéréovision sous-marine ou analyse d'images en environnement aérien contrôlé.

En milieu sous-marin, par exemple, dans l'étude menée par Muñoz et al. [13], les auteurs ont combiné la détection par YOLO et la segmentation par Mask R-CNN, puis converti les coordonnées de points clés museau et queue en distances euclidiennes 3D via un système stéréoscopique calibré. Garcia et al. [12] ont proposé une méthode similaire, en affinant les contours segmentés et en extrayant le squelette du poisson à partir d'images stéréo, avant de modéliser sa courbure avec un polynôme ajusté par RANSAC. Ces deux approches, bien que techniquement avancées, souffrent de l'absence de mesures de référence et d'une complexité opérationnelle élevée.

Une autre méthode, qui ne repose pas sur la segmentation, a été proposée par Yu et al. [14]. Ils ont construit un ensemble de données annotées appelé Fish-Keypoints et ont utilisé un réseau neuronal profond (DNN) pour détecter les poissons ainsi que leurs points clés. Grâce à un système de caméras binoculaires, ils ont reconstruit un système de coordonnées 3D permettant de mesurer les points situés à la tête et à la queue des poissons, facilitant ainsi le calcul de leur longueur.

En environnement contrôlé, Álvarez-Ellacuría et al. [6] ont appliqué Mask R-CNN pour estimer automatiquement la longueur du merlu européen à partir d'images de caisses de poisson. Leur méthode consistait d'abord à segmenter la tête du poisson, puis à mesurer sa longueur en pixels. Une première régression statistique permettait de convertir cette mesure en centimètres, tandis qu'une seconde estimait la longueur totale. Les prédictions ont été comparées aux longueurs annotées pour évaluer la performance.

4. Présentation et analyse des données

L'objectif de la présente étude est d'obtenir automatiquement la longueur des poissons à partir d'images de caisses de poissons prises dans les ports. Dans cette section, nous présentons les jeux de données ainsi que leur analyse.

4.1 Données

Les données principales utilisées dans ce projet proviennent du portail en ligne de la criée de La Rochelle. Elles comprennent des photographies de caisses de poissons, prises automatiquement par un dispositif intégré au système de vente, et des métadonnées commerciales, associées à chaque lot.

4.1.1. Base de données des débarquements quotidiens

Lors du débarquement, chaque caisse est déposée sur un poste de pesée équipé d'un dispositif automatisé combinant une balance, une caméra positionnée en hauteur et un terminal d'enregistrement. Ce système déclenche automatiquement la pesée, la prise de vue et l'enregistrement des informations commerciales.

L'ensemble des données est ensuite horodaté et stocké dans une base de données centrale, utilisée pour la gestion des ventes et la traçabilité des produits.

Le tableau suivant présente une description synthétique des principales informations associées à chaque caisse.

Donnée	Description
Vente	Numéro de vente associé à l'enregistrement la caisse
lot	Identifiant unique du lot commercial
Espèce	Nom commercial de l'espèce
Poids	Poids total de la caisse
Date	Date d'enregistrement
Bateau	Nom du navire ayant fourni la caisse
Catégorie	Code de tri commercial attribué par la criée, reflétant une classe de taille.

Figure 4.1 : Données descriptives associées aux caisses de poissons

Les images ci-dessous illustrent les caisses de débarquement et d'enregistrement à la criée.

Figure 4.2 : Débarquement de poissons à la criée et poste de pesée

4.1.2. Images des caisses de poissons

Chaque image représente une caisse homogène, contenant exclusivement des individus appartenant à une même espèce et à une même catégorie commerciale de taille. Les poissons y sont disposés sans ordre strict. Ils peuvent être placés de manière espacée ou dense, parfois en une seule couche, parfois avec un léger empilement. Dans plusieurs cas, certains individus sont partiellement occultés par d'autres ou par la paroi de la caisse. Il est fréquent que seules les têtes soient visibles, tandis que les queues sont souvent masquées.

Certains clichés présentent des éléments perturbateurs tels que des protections plastiques transparentes recouvrant les poissons probablement pour des raisons d'hygiène et de conservation qui peuvent générer des reflets masquant certains détails des poissons. Par ailleurs, des artefacts visuels peuvent apparaître : flou lié au mouvement lors du passage sous la caméra, déformation ou variation d'éclairage.

Les images sont prises en vue zénithale, sans système de calibration connu. La distance entre la caisse et la caméra peut varier légèrement d'une scène à l'autre, ce qui introduit une variabilité supplémentaire.

5. Méthodologie

Cette section présente la méthodologie mise en œuvre dans le cadre de ce travail.

5.1 Estimation de la longueur des poissons

Estimer la taille individuelle des poissons à partir d'images 2D, sans information de profondeur, constitue une tâche complexe en vision par ordinateur. Dans ce travail, la démarche méthodologique a été guidée par une série de questions structurantes visant à répondre aux défis posés par la variabilité des scènes et la qualité d'observation des individus.

1. Quelle approche de segmentation automatique permet d'extraire avec précision les contours des poissons visibles dans leur intégralité ?
2. Comment extraire une estimation fiable de la longueur du poisson, en tenant compte d'éventuelles courbures corporelles ?
3. Comment assurer une conversion robuste des longueurs exprimées en pixels vers des unités métriques ? Existe-t-il, au sein des images, des objets de référence stables persistants sur toute la banque d'images disponibles ?
4. Dans les cas d'occultation partielle, de chevauchement ou de mauvaise qualité d'image, une estimation indirecte reste-t-elle envisageable ?
5. Comment articuler automatiquement les deux approches d'estimation de la longueur, en concevant une stratégie de décision dynamique adaptée à la configuration visuelle de chaque caisse, de manière à optimiser le taux de mesure tout en maintenant un niveau de fiabilité acceptable ?

Afin de répondre à ces enjeux, deux stratégies principales ont été mises en œuvre : une estimation directe fondée sur la silhouette complète des poissons, et une estimation indirecte exploitant des repères anatomiques partiels. Chacune de ces approches a nécessité des choix techniques spécifiques, détaillés dans les sections suivantes.

5.1.1. Estimation directe par segmentation d'instance et extraction du squelette

Compte tenu des méthodes analysées dans le chapitre 3, notamment les approches supervisées comme YOLO ou Mask R-CNN, nous avons orienté notre choix vers l'utilisation de Grounded SAM. Cette solution combine Grounding DINO,qui prédit avec précision l'emplacement des poissons en générant automatiquement des boîtes englobantes à partir de requêtes textuelles, et SAM, qui génère les masques de segmentation correspondants aux poissons détectés. Cette combinaison nous permet de segmenter les

poissons sans recourir à un jeu de données annoté manuellement, ce qui permet de contourner une contrainte importante en termes de temps.

À partir des masques générés, une procédure de traitement géométrique a été mise en œuvre pour estimer la longueur des poissons. Elle repose sur les étapes suivantes :

- Extraction du squelette : chaque masque binaire est transformé en une représentation simplifiée de l'axe médian du poisson, par squelettisation.

- Régularisation de l'axe : l'axe extrait est ensuite lissé à l'aide d'un ajustement polynômial, limité aux contours du masque, afin de corriger les irrégularités liées à la pixellisation.

- Estimation de la longueur : la longueur du poisson est calculée comme la somme des distances euclidiennes entre les points successifs de l'axe ajusté.

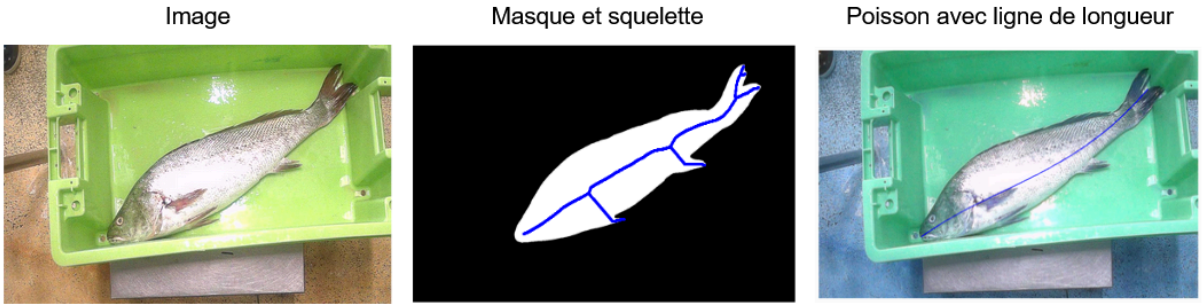

Figure 5.1 : Visualisation de la mesure de la longueur du poisson en pixels

Cette méthode fournit une mesure en pixels de la longueur apparente du poisson.

5.1.2. Estimation indirecte par détection de la tête

Dans notre cas, nous cherchons à mesurer les poissons, même lorsqu'ils sont partiellement visibles.
Un modèle YOLO11-Pose a été entraîné pour détecter quatre points anatomiques caractéristiques sur la tête des poissons : l'œil, la bouche annotée par deux points distincts lorsqu'elle est ouverte et la nageoire pectorale. Ces points ont été sélectionnés en raison de leur visibilité fréquente sur les images et de leur pertinence morphologique.

L'annotation a été réalisée manuellement sur 300 images, à l'aide d'un outil développé en interne dédié à la labellisation d'images pour entraîner des réseaux de neurones profonds. Chaque tête a été annotée à la fois par une boîte englobante et par les quatre points clés, dans un ordre strictement défini afin d'assurer la cohérence des données d'apprentissage.

Les coordonnées de chaque point clé ont été extraites pour chaque tête détectée. À partir de ces coordonnées, les distances euclidiennes entre les points ont été calculées en pixels, permettant de reconstituer les trois mesures morphométriques sur l'ensemble des individus. Ces mesures en pixels ont ensuite été converties en millimètres à l'aide du facteur de conversion obtenu par la méthode décrite en section 5.1.4, avant d'être utilisées comme variables explicatives dans les modèles statistiques présentés ci-après.

5.1.3. Modélisation statistique de la longueur à partir des repères morphométriques

L'estimation de la longueur totale Lt du poisson repose sur l'établissement de relations morphométriques entre cette dernière et des mesures anatomiques localisées sur la tête.
Trois mesures ont été étudiées :

- La longueur pré-orbitaire Lpo : définie comme la distance entre l'extrémité de la bouche et le bord antérieur de l'œil.

- La longueur pré-pectorale Lpp : correspondant à la distance entre l'extrémité de la bouche et point d'insertion de la nageoire pectorale.

- La longueur orbito-pectorale Lop : correspondant à la distance entre le bord postérieur de l'œil et le point de l'insertion de la nageoire pectorale.

Un total de 178 poissons a été mesuré manuellement en conditions réelles, dans le but de déterminer laquelle de ces trois distances offre la meilleure capacité prédictive pour estimer la Lt.

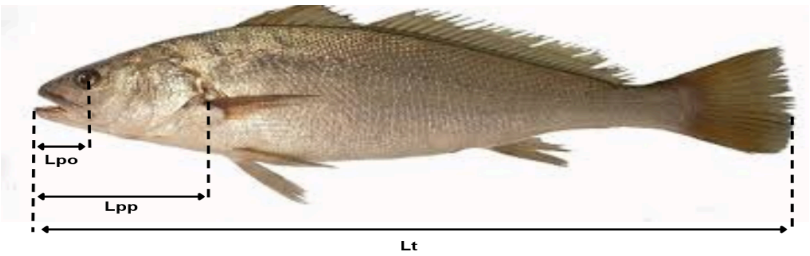

Figure 5.2 : Les mesures morphométriques utilisées

Pour la mesure pré-pectorale, un jeu de données complémentaire a permis d'atteindre un effectif de 231 individus.

Trois modèles linéaires ont été ajustés afin d'évaluer la capacité explicative de chaque mesure céphalique à prédire la longueur totale.

5.1.4 Conversion des longueurs en valeurs métriques

Pour rendre exploitables les longueurs estimées en pixels à partir des images, une conversion des distances exprimées en pixels vers des valeurs métriques a été mise en place. Cette étape repose sur la présence, dans l'image, d'un objet de dimensions connues servant de référence.

- **Utilisation de la caisse**

Dans un premier temps, la caisse contenant les poissons a été retenue comme repère. Ses dimensions sont standardisées à 80 centimètres de longueur sur 44 de largeur. Grâce à une requête textuelle de type *box*, le masque de la caisse est automatiquement extrait à l'aide de Grounded-SAM, puis analysé pour détecter les coins. Selon la visibilité partielle ou totale de la caisse dans l'image, la dimension la mieux représentée longueur ou largeur a été sélectionnée. La distance en pixels détectée a ensuite été mesurée et rapportée à la dimension physique correspondante, conduisant au calcul d'un facteur pixels/cm.

Cependant, cette méthode présente certaines limites observées sur un ensemble d'images. La caisse n'est pas toujours entièrement visible dans le champ de la caméra.

- **Détection des poignées de caisse**

Afin de surmonter ces difficultés, une méthode alternative a été mise en œuvre. Elle consiste à détecter les trous situés autour des poignées de la caisse. Ces poignées sont visuellement saillantes, leurs dimensions sont constantes, avec une largeur réelle de 14 centimètres, et elles sont souvent bien visibles dans le champ de la caméra.

Un modèle YOLO11-Pose a été entraîné sur 300 images annotées pour repérer précisément ces repères visuels. Pour chaque détection, la distance entre les deux trous a été mesurée en pixels, puis comparée à leur écartement réel, afin d'obtenir un facteur de conversion spécifique à l'image.

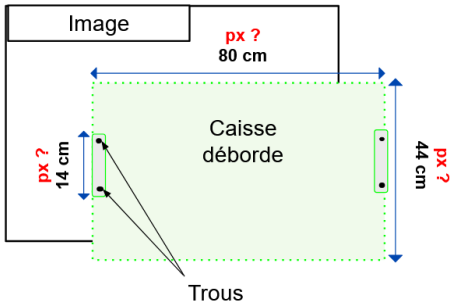

Figure 5.2 : Méthodes de référence pour la conversion pixel/cm

5.1.5 Constitution des données d'évaluation

Afin d'évaluer les performances des deux approches d'estimation de la longueur totale des poissons , une campagne de mesure manuelle a été menée par notre équipe à la criée de La Rochelle.Cette phase terrain, conduite selon un protocole standardisé, a permis de mesurer un total de 178 individus.

Pour chaque poisson, quatre mesures ont été enregistrées sur le terrain : la longueur totale, mesurée en centimètres à l'aide d'un mètre gradué, et les trois mesures céphaliques pré-orbitaire, pré-pectorale,et orbito-pectorale, relevées en millimètres à l'aide d'un pied à coulisse. Ces poissons ont été répartis dans 12 caisses distinctes, chacune ayant été photographiée. Les 12 images obtenues constituent la base d'évaluation visuelle utilisée pour tester les prédictions automatiques.

L'évaluation a ensuite été réalisée en deux volets distincts :

- **Approche morphométrique**

Dans cette approche, les longueurs totales sont estimées à partir des mesures céphaliques détectées automatiquement, puis injectées dans les relations allométriques établies.
Cependant, toutes les têtes détectées n'étant pas valides certaines boîtes regroupent plusieurs têtes ou présentent des erreurs de détection, une sélection manuelle des têtes correctes a été effectuée en amont, de manière à garantir la fiabilité des données d'entrée. Ainsi, l'évaluation de cette approche repose uniquement sur les têtes correctement détectées et aux individus identifiés grâce à l'étiquetage manuel réalisé lors de la campagne de terrain.

- **Approche squelettique**

L'estimation fondée sur la silhouette corporelle nécessite que les poissons soient entièrement visibles sur les images. Pour identifier automatiquement ces cas, un pipeline de filtrage géométrique a été mis en œuvre. Il repose sur la détection de la tête et de la nageoire caudale, la segmentation du corps à l'aide de Grounding DINO et SAM, puis l'extraction de l'axe central du poisson.
Seuls les poissons dont le squelette commence dans la boîte de la tête, se termine dans celle de la nageoire caudale, et présentent un alignement cohérent entre les trois centres, ont été retenus.

Ce processus a permis d'extraire un sous-ensemble de 14 individus entièrement visibles, qui ont été utilisés pour évaluer l'estimation de la longueur à partir du squelette.

Figure 5.3 : Exemple de sélection des poissons entièrement visible sur l'image

6. Résultats et analyse

L'objectif de ce chapitre est d'analyser les résultats issus de la méthodologie décrite au chapitre précédent, en comparant les performances des différentes approches mises en œuvre pour la reconstruction des données et la prédiction de la longueur des poissons. Cette analyse est essentielle pour évaluer l'efficacité de chaque méthode, tout en identifiant leurs points forts et leurs limites, afin d'éclairer le processus de décision.

6.1 Analyse comparative des méthodes d'estimation du facteur de conversion

L'efficacité de l'estimation des longueurs repose directement sur la précision du facteur de conversion entre pixels et centimètres. Pour évaluer sa qualité, une première phase de collecte de données « à la main » a été réalisé, un ensemble de 342 images a été annoté manuellement dans ImageJ pour obtenir un facteur de conversion de référence pour chaque image.

Les deux approches développées ont été testées sur cet ensemble. La méthode fondée sur la détection des poignées fournit une couverture complète de l'ensemble du jeu de données. À l'inverse, la méthode basée sur la détection de caisse ne génère un facteur exploitable que dans environ 64 % des cas, en raison de masques incomplets ou de cadrages partiels. Parmi ces cas détectés, environ 11 % présentent des valeurs incohérentes, traduisant une instabilité du processus de segmentation.

Pour permettre une comparaison rigoureuse, l'analyse des différents facteurs de conversion a été restreinte à un sous-ensemble de 184 images sur lesquelles les deux méthodes produisent des résultats exploitables. L'erreur moyenne a été calculée, reflétant l'écart entre le facteur estimé avec l'outil d'IA facteur de conversion observé.

Une erreur moyenne de 0,00149 px/cm pour la méthode par caisse correspond à environ 0,12 cm d'écart sur une longueur de 80 cm. Cette précision est amplement suffisante pour des mesures morphométriques comprises entre 30 et 80 cm. La méthode par poignées entraîne un écart d'environ 0,27 cm sur 80 cm, ce qui reste acceptable dans ce contexte.

En résumé, la méthode par caisse est plus précise mais moins souvent applicable. La méthode par poignées est moins précise mais plus robuste.

6.2. Corrélation entre les mesures céphalique et la longueur totale

6.3. Évaluation des performances de la chaîne de traitement

Les performances du système sont évaluées selon deux aspects complémentaires. Premièrement, la capacité du modèle à détecter correctement les têtes de poissons dans les images. Deuxièmement, la fiabilité des mesures morphométriques extraites automatiquement à partir de ces têtes détectées.

6.3.1. Évaluation de la détection automatique des têtes

Afin d'évaluer les performances du modèle de détection des têtes de poissons, un sous-échantillon de 12 images a été analysé manuellement. Pour chaque image, le nombre réel de poissons visibles a été comparé au nombre de têtes détectées automatiquement, en distinguant les détections valides des faux positifs.

Sur cet échantillon, le taux de détection moyen est de 71,97 %, avec une performance maximale de 100 % sur certaines images et un minimum de 0 % dans des cas de fort chevauchement ou de faible qualité visuelle. Le taux moyen de faux positifs s'élève à 15,28 %, ce qui reflète une proportion modérée de détections erronées, souvent dues à des regroupements de têtes dans des zones densément peuplées.

Figure 6.1 : Exemples de détection des têtes

La précision moyenne, définie comme la proportion de détections valides parmi l'ensemble des détections effectuées, est de 84,72 %. Ces résultats montrent que le système est globalement fiable pour extraire les mesures céphaliques lorsque les têtes sont bien visibles et séparées.

À ce stade, seules les têtes validées manuellement ont été conservées pour les analyses de morphométrie, afin de garantir la qualité des données. Toutefois, dans une perspective d'automatisation complète, il sera nécessaire d'intégrer un mécanisme de filtrage automatique pour exclure les faux positifs sans intervention humaine.

6.3.2 Évaluation de la stabilité et de la précision des mesures céphaliques

Les erreurs d'estimation ont été analysées sur 71 individus afin d'évaluer la fiabilité du système. Elles correspondent à l'écart entre les mesures extraites automatiquement et les valeurs de référence mesurées manuellement.

L'analyse des erreurs de détection sur les mesures céphaliques permet de quantifier à la fois leur précision et leur stabilité. Le tableau ci-dessous regroupe les principaux indicateurs calculés.

Mesure	Biais (mm)	MAE (mm)	Erreur relative (%)	CV (%)

Pré-orbitaire	-5.75	6.55	9.87	9.28
Orbito-pectorale	-5.27	7.54	22.57	23.48
Pré-pectorale	-14.75	15.89	16.37	9.69

Figure 6.2 : Indicateurs d'erreur sur les mesures céphaliques

Dans le contexte de cette étude, où les mesures céphaliques varient en moyenne entre 30 mm et 100 mm, un seuil de 10 à 15 % de coefficient de variation est jugé raisonnable pour distinguer les mesures stables de celles trop incertaines. Ce seuil permet de poser une limite pratique au-delà de laquelle l'erreur devient trop variable. Il est donc défini ici de manière empirique à partir des statistiques descriptives, en lien avec l'objectif de fiabilité globale du système. La figure ci-dessous illustre graphiquement la variabilité des erreurs de détection par type de mesure.

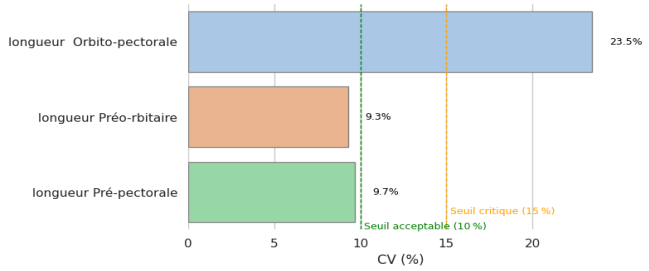

Figure 6.3 : Variabilité des erreurs de détection selon la mesure céphalique

Compte tenu des seuils définis, on constate que : les mesures pré-orbitaire et pré-pectorale, affichent un CV inférieur à 10 %, suggérant une erreur plus stable.

6.4. Analyse des performances des estimations de longueur totale

Afin d'évaluer dans quelle mesure les mesures céphaliques peuvent prédire la longueur totale, trois modèles de régression linéaire ont été ajustés à partir des données récoltées lors de la campagne de mesure. Chaque modèle établit une relation entre une mesure céphalique et la longueur totale des individus.

Les résultats mettent en évidence une relation particulièrement forte entre la mesure pré-pectorale et la longueur totale : cette variable explique à elle seule plus de 95 % de la variabilité observée. La mesure orbito-pectorale présente également un pouvoir prédictif élevé, avoisinant 87 %, tandis que la mesure pré-orbitaire montre une relation plus modérée, avec environ 77 % de variabilité expliquée.

Ces résultats suggèrent que la longueur pré-pectorale constitue un indicateur morphométrique particulièrement pertinent pour estimer la longueur totale du poisson dans des conditions d'acquisition visuelle.

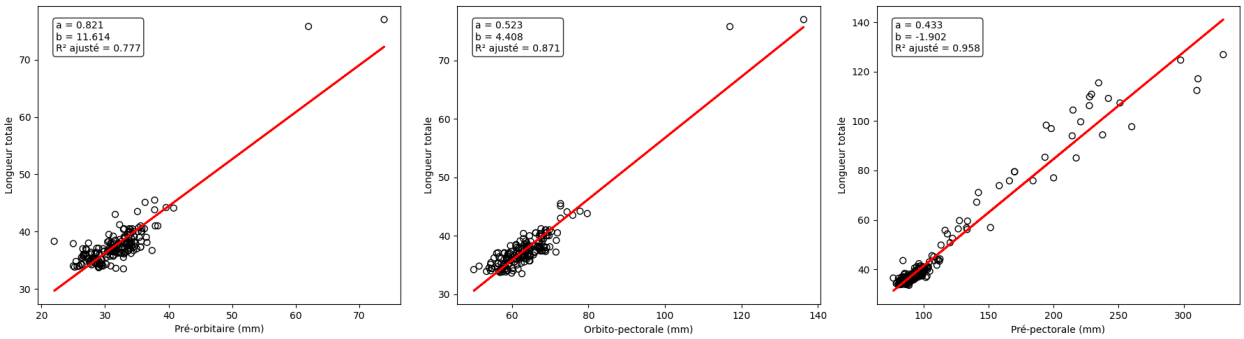

Figure 6.4 – Modélisation de la longueur totale à partir repères morphométriques

6.4.1. Évaluation de l'estimation de la longueur totale à partir du squelette

L'estimation de la longueur totale fondée sur la silhouette du corps du poisson, obtenue via l'extraction de l'axe squelettique présente une sous-estimation systématique, indépendamment du degré du modèle polynomial utilisé. Le modèle de degré 3 offre une performance légèrement supérieure, mais l'erreur reste notable, avec un biais négatif marqué.

Degré du polynome	Biais (cm)	MAE (cm)	Erreur relative moyenne (%)
2	-7.615532272810317	7.615532272810317	2,5

3	7.6296709016179864	7.6296709016179864	2,1

Figure 6.5 – Performance de l'estimation squelettique

Cette sous-estimation s'explique par un effet de perspective induit par la configuration spatiale dans les images.

L'évaluation des longueurs totales estimées à partir des squelettes met en évidence une erreur systématique, quel que soit le degré du polynôme utilisé. Le modèle de degré 3 présente un ajustement légèrement meilleur, mais le biais reste significatif.

6.4.2 Évaluation des longueurs estimées à partir des mesures céphaliques

Les longueurs totales ont été estimées à partir des mesures céphaliques détectées automatiquement, en les injectant dans des relations allométriques établies précédemment.

Cette analyse vise à évaluer la performance finale du système, en comparant les longueurs estimées aux longueurs mesurées manuellement sur le terrain. Trois indicateurs ont été retenus :

- **Le biais**, qui mesure la tendance à la sous- ou surestimation.
- **L'erreur absolue moyenne** , qui donne l'écart moyen entre estimation et réalité.
- **L'erreur relative moyenne (%)**, qui exprime l'erreur en pourcentage de la vraie longueur, permettant une comparaison normalisée quelles que soient les tailles.

Mesure d'entrée	Biais (cm)	MAE (cm)	Erreur relative moyenne (%)
Pré-orbitaire	-4,41	4,88	11,96
Orbito-pectorale	-2.75	3,53	8,94

Pré-pectorale	-5.46	5,55	13,99

Figure 6.6 : Indicateurs d'erreur sur l'estimation de la longueur totale à partir des mesures céphaliques

Les résultats, obtenus sur un échantillon de 71 poissons, montrent une sous-estimation systématique, plus ou moins marquée selon la mesure utilisée.

La meilleure performance est observée avec la mesure orbito-pectorale, qui affiche un biais de -2,75 cm et une erreur relative de 8,94 %, tandis que la mesure pré-pectorale, pourtant la plus prédictive en théorie, présente une erreur plus forte 13,99 % et un biais de -5,46 cm.

Ces résultats suggèrent que la stabilité de la détection influence autant que la relation allométrique elle-même, et qu'une mesure moins théoriquement prédictive mais mieux détectée peut donner de meilleurs résultats finaux.

6.5 Distribution des tailles

L'objectif final de cette étude est de reconstruire la distribution des tailles des poissons.

Chaque image analysée est associée à une catégorie commerciale, correspondant à une classe de taille. À partir des longueurs estimées automatiquement, plusieurs niveaux de distribution ont été produits.

Trois variantes de la distribution ont été générées :

- **Détection directe** : ne prend en compte que les poissons détectés automatiquement dans les images exploitables.
- **Extrapolation par caisse** : complète les tailles manquantes en estimant le nombre total de poissons à partir du poids déclaré de la caisse.
- **Extrapolation par catégorie** : permet de reconstruire la distribution globale des tailles en compensant les images non traitées, qu'il s'agisse d'une absence de repère de conversion, d'une mauvaise qualité d'image. Cette extrapolation s'appuie sur les proportions observées dans les caisses valides appartenant à la même catégorie commerciale.

6.5.1. Extrapolation des tailles

L'extrapolation repose sur une relation allométrique taille-poids, permettant d'estimer le poids moyen d'un poisson à partir de sa longueur.

En connaissant le poids total d'une caisse et la taille estimée des poissons détectés, il est possible d'estimer leur poids individuel, puis d'en déduire le nombre de poissons manquants. Ces individus manquants sont ensuite ajoutés à la distribution, en respectant la répartition des tailles observée dans la même caisse.

Cette approche assure une estimation cohérente des tailles à l'échelle globale, malgré l'exclusion de certaines images ou caisses inexploitables.

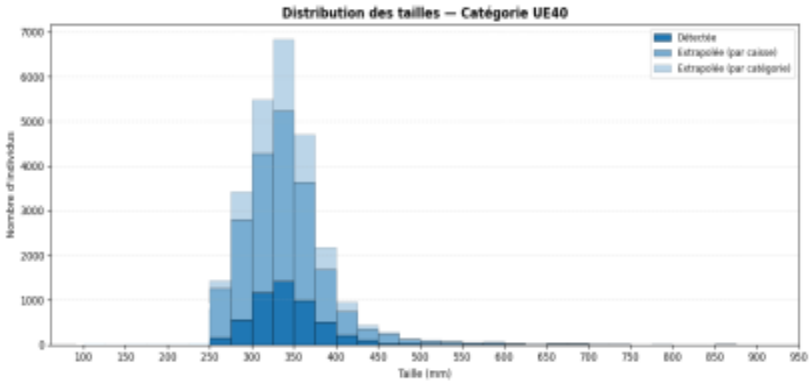

Figure 6.7 : Distribution des tailles pour une catégorie commercial

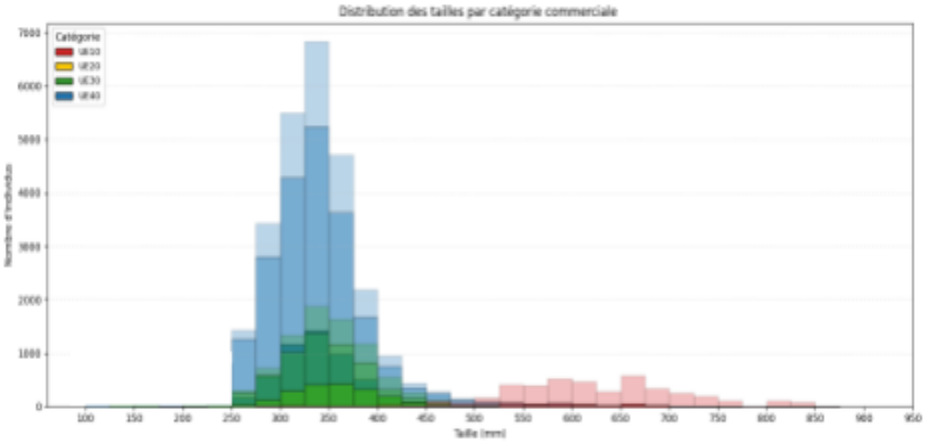

Figure 6.8 : Distribution des tailles pour toutes les catégories commerciales

La distribution obtenue à partir des données traitées automatiquement reproduit fidèlement les tendances observées sur les échantillons collectés manuellement dans des travaux antérieurs : une majorité de poissons de petite taille, et une fréquence décroissante pour les individus les plus grands. Cette cohérence générale valide la pertinence de l'approche proposée pour reconstituer des distributions réalistes à partir d'images acquises en criée.

7. Discussion générale

Ce travail a permis d'établir les fondations d'un système automatisé d'estimation de la longueur totale des poissons à partir d'images 2D complexes prises en routine, en s'appuyant sur la détection automatique de repères anatomiques. Deux approches complémentaires ont été investiguées : l'une reposant sur la géométrie du squelette, l'autre s'appuyant sur des mesures céphaliques injectées dans des relations allométriques.

Les analyses menées ont mis en évidence la forte sensibilité du pipeline à la qualité des données d'entrée, en particulier à la précision des points d'intérêt détectés. L'approche squelettique, 14 individus analysés, l'estimation présente un biais moyen de -7,6 cm. D'après les observations, cette erreur semble liée à l'effet de perspective, dû à la profondeur de champ et à la position relative des individus par rapport à la poignée l'objet de référence, situé dans un plan différent. Cela souligne l'importance de considérer la

géométrie de la scène pour améliorer la fiabilité des estimations. Une analyse exploratoire a été engagée pour quantifier cet effet, en testant la détection d'étiquettes placées à différentes profondeurs.

L'approche morphométrique, quant à elle, a permis une meilleure couverture sur l'échantillon, mais au prix d'une sous-estimation systématique des longueurs, variable selon la nature de la mesure utilisée.

Ce stage a permis de cerner les principaux verrous techniques du pipeline et d'identifier des pistes concrètes d'amélioration.

Enfin, bien que l'estimation individuelle ne soit pas encore totalement fiable, les résultats suggèrent qu'il est possible de reconstruire une distribution cohérente des tailles à l'échelle populationnelle.

8. Conclusion

Ce travail a permis de valider la faisabilité d'un système automatisé d'estimation de la longueur des poissons à partir d'images 2D, en mobilisant des repères anatomiques détectés par vision par ordinateur.. Il ouvre ainsi plusieurs pistes concrètes pour améliorer la robustesse, la précision et la généralisabilité du dispositif.

Un premier prolongement naturel consiste à approfondir l'étude des biais géométriques introduits par la perspective, particulièrement visibles dans l'approche squelettique. À ce titre, l'intégration d'informations géométriques issues de la scène, comme la focale estimée ou la profondeur relative entre le poisson et un objet de référence, pourrait constituer une avancée majeure. Des outils comme Depth_Pro pourraient être testés pour reconstituer une calibration implicite à partir d'indices visuels, permettant d'adapter dynamiquement l'échelle métrique dans chaque image. L'effet attendu serait une réduction significative du biais systématique sans complexifier le processus de capture.

Du côté des mesures morphométriques, les résultats indiquent que certaines erreurs de détection présentent un comportement relativement stable. Ces propriétés suggèrent qu'une correction statistique pourrait être envisageable pour certaines longueurs, à travers des modèles linéaires simples ou des techniques de recalage par régression. Ce type de stratégie permettrait de conserver l'intérêt d'une approche massivement automatisée tout en améliorant la qualité des estimations.

En parallèle, l'amélioration des points d'intérêt anatomiques reste une priorité. La variabilité observée dans la détection automatique, notamment liée à la qualité de pose, invite à renforcer la l'enrichissement du jeu d'entraînement par des cas plus variés angles, tailles, positions.

Dans cette dynamique d'amélioration continue, plusieurs pistes peuvent être envisagées. Le pipeline actuel, bien qu'en traîné sur une seule espèce dans un contexte maîtrisé, pourrait être adapté à d'autres espèces ou à des environnements. Cela supposerait une adaptation des relations allométriques ou l'introduction de modèles spécifiques par espèce, avec une base de données multi-source pour garantir la transférabilité du système.

Même si l'estimation individuelle peut être entachée d'un certain biais, l'analyse d'un volume important de données permet de reconstituer la distribution des tailles à l'échelle d'une population. Cette perspective est particulièrement prometteuse dans un cadre de suivi halieutique, où la tendance globale est parfois plus pertinente que la mesure isolée. Cela ouvrirait la voie à une exploitation scientifique ou réglementaire directe du système.

Bibliographie

[1] Schmidhuber, J. (2012). *Réseaux neuronaux profonds à colonnes multiples pour la classification d'images*. In : Actes de la Conférence IEEE sur la vision par ordinateur et la reconnaissance des formes (CVPR).

[2] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *Classification d'ImageNet avec des réseaux de neurones convolutifs profonds*. In : *Advances in Neural Information Processing Systems*.

[3] Ifremer. *Création de l'unité mixte de recherche DECOD – De la source à l'océan*. Communiqué de presse,17 octobre 2022 : https://www.ifremer.fr/sites/default/files/2022-11/cp_umr_decod_17102022.pdf

[4] Rajat Kumar Sinha, Ruchi Pandey et Rohan Pattnaik. *Deep learning for computer vision tasks: a review*. arXiv preprint arXiv:1804.03928, 2018.

[5] Y.-L. Boureau, J. Ponce et Y. LeCun. *A theoretical analysis of feature pooling in visual recognition*. Dans : *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, p. 111–118, 2010.

[6] Hafiz, Abdul Mueed et Bhat, Ghulam Mohiuddin. *A survey on instance segmentation: state of the art*. International Journal of Multimedia Information Retrieval, vol. 9, no 3, p. 171–189, 2020.

[7] Liang-Chieh Chen, Alexander Hermans, George Papandreou, Florian Schroff, Peng Wang et Hartwig Adam, *MaskLab : segmentation d'instances par raffinement de la détection d'objets à l'aide de caractéristiques sémantiques et directionnelles*, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 4013–4022, 2018.

[8] Amaya Álvarez-Ellacuría, Miquel Palmer, Ignacio A. Catalán et Jose-Luis Lisani, *Estimation non supervisée de la taille des poissons à partir d'images de débarquements commerciaux en utilisant l'apprentissage profond*, *ICES Journal of Marine Science*, vol. 77, n° 4, p. 1330–1339, 2020.

[9] Eko Prasetyo, Nanik Suciati et Chastine Fatichah, *Comparaison entre YOLO et Mask R-CNN pour la segmentation de la tête et de la queue des poissons*, dans *Actes de la 4e Conférence Internationale sur l'Informatique et les Sciences de Calcul (ICICoS 2020)*, p. 1–6, 2020.

[11] Pau Climent-Pérez, Alejandro Galán-Cuenca, Nahuel Emiliano García-d'Urso, Marcelo Saval-Calvo, Jorge Azorín-López et Andrés Fuster-Guilló, *Estimation automatique de la taille des poissons à partir d'images non calibrées issues de marchés, en utilisant la vision par ordinateur et l'apprentissage profond*, dans *Atelier international sur les modèles de calcul souple dans les applications industrielles et environnementales*, pp. 319–329, Springer, 2022.

[12] Rafael Garcia, Ricard Prados, Josep Quintana, Alexander Tempelaar, Nuno Gracias, Shale Rosen, Håvard Vågstøl et Kristoffer Løvall, *Segmentation automatique de poissons par apprentissage profond, avec application à la mesure de leur taille*, *ICES Journal of Marine Science*, vol. 77, no 4, p. 1354–1366, 2020.

[13] Pau Muñoz-Benavent, J. Martínez-Peiró, Gabriela Andreu-García, Vicente Puig-Pons, V. Espinosa, Isabel Pérez-Arjona, Fernando De la Gándara et A. Ortega, *Évaluation de l'impact de l'apprentissage profond sur la segmentation d'images pour la mesure automatique du thon rouge*, *Aquacultural Engineering*, vol. 99, art. 102299, 2022.

[14] Yaozhen Yu, Hao Zhang, et Fei Yuan. *Méthode de détection de points clés pour la mesure de la taille des poissons basée sur l'apprentissage profond*. IET Image Processing, vol. 17, no 14, p. 4142–4158, 2023.

[15] Yaozhen Yu, Hao Zhang, et Fei Yuan. Key point detection method for fish size measurement based on deep learning. *IET Image Processing*, vol. 17, no 14, p. 4142–4158, 2023.9.