

RAPPORT DE TER

TAHARASTE Rayan, GUILHEM Adrian, INGABIRE Alvin,
NIYONKURU Berline, ZEROUAL Walid



Département MIASHS, UFR 6 Informatique, Mathématique et Statistique
Université Paul Valéry, Montpellier 3

14 mai 2025

Déclaration de non-plagiat

Nous déclarons que ce rapport est le fruit de notre seul travail, à part lorsque cela est indiqué explicitement.

Nous acceptons que la personne évaluant ce rapport puisse, pour les besoins de cette évaluation :

- la reproduire et en fournir une copie à un autre membre de l'université ; et/ou,
- en communiquer une copie à un service en ligne de détection de plagiat (qui pourra en retenir une copie pour les besoins d'évaluation future).

Nous certifions que nous avons lu et compris les règles ci-dessus. En signant cette déclaration, nous acceptons ce qui précède.

Signature : GUILHEM Adrian

Date : 14/05/2025

Signature : NIYONKURU Berline

Date : 14/05/2025

Signature : INGABIRE Alvin

Date : 14/05/2025

Signature : ZEROUAL Walid

Date : 14/05/2025

Signature : TAHARASTE Rayan

Date : 14/05/2025

Remerciements

Nous tenons à exprimer notre profonde gratitude à toutes les personnes qui ont contribué de près ou de loin à la réalisation de ce travail.

Nos remerciements les plus sincères vont en premier lieu à notre encadrant pédagogique, Jérôme Pasquet, pour son accompagnement précieux, ses conseils avisés et sa disponibilité tout au long de ce projet. Son expertise et son regard critique ont été d'une aide inestimable pour mener à bien cette recherche.

Nous adressons également nos vifs remerciements à Julien Renoult, commanditaire de ce projet, pour sa confiance, son soutien constant et ses retours constructifs qui ont grandement enrichi notre travail.

Une pensée particulière va à l'ensemble de l'équipe pédagogique du département MIASHS pour son soutien indéfectible, la qualité de son enseignement et les ressources mises à notre disposition tout au long de notre formation.

Enfin, nous ne saurions oublier nos proches (famille, amis, collègues) pour leur soutien moral, leurs encouragements et leur patience durant cette aventure intellectuelle.

14/05/2025.

Résumé

Ce rapport présente notre travail de recherche explorant les liens potentiels entre les caractéristiques physiologiques des oiseaux et leur environnement dans un contexte de sélection sexuelle. À travers l'analyse de milliers d'images issues de la plateforme iNaturalist, combinée à des techniques de computer vision et d'analyse statistique, nous avons cherché à valider ou infirmer l'hypothèse selon laquelle les oiseaux pourraient développer des traits physiques s'appropriant des caractéristiques environnementales à des fins de séduction.

Table des matières

Table des matières	4
1 Introduction	5
1.1 Définition claire de la question de recherche et mise en place des outils de travail collaboratif	5
2 Planification du projet, des rôles et des tâches	6
2.1 Planification du projet	6
2.2 Tâches	6
3 Acquisition, nettoyage, mise en forme des données et Analyse Exploratoire	8
3.1 Acquisition des données	8
3.2 Nettoyage et mise en forme des données	8
3.3 Analyse exploratoire des données	8
4 Méthodologie	10
4.1 Pipeline, résultats et interprétation	10
4.1.1 Extraction d'oiseaux	10
4.1.2 Détection de traces humaines :	13
4.1.3 Détection de zones floues :	16
4.1.4 Extraction Background :	19
4.1.5 Features ResNet :	20
4.1.6 Données géographiques et cartes :	20
4.1.7 Similarités	21
5 Conclusion	27
Bibliographie	28
A Code d'extraction des métadonnées	29
B Liste des packages utilisés	31

Chapitre 1

Introduction

1.1 Définition claire de la question de recherche et mise en place des outils de travail collaboratif

- **Question de recherche :** Existe-t-il un lien entre les caractéristiques physiologiques des oiseaux et les caractéristiques de leur environnement, autre que pour une question de camouflage ?

Le but du commanditaire est de démontrer ou réfuter sa théorie selon laquelle les oiseaux s'approprient des caractéristiques de leur environnement dans le but de séduire. Cette recherche explore l'hypothèse que l'apparence physique des oiseaux peut être influencée par des facteurs environnementaux dans un contexte de séduction, plutôt que simplement pour le camouflage.

- **Objectifs du projet :** Le projet vise à analyser si des caractéristiques physiques des oiseaux peuvent être liées à des éléments de leur environnement dans une perspective comportementale, en particulier concernant la séduction. Cette question est d'une grande importance pour mieux comprendre l'évolution des comportements et des adaptations physiques des oiseaux.

- **Outils collaboratifs :**

- **Overleaf :** Pour la rédaction collaborative du rapport en LaTeX, permettant un travail simultané avec historique des modifications.
- **GitHub :** Pour le versionnage du code avec gestion des branches.
- **Google Colab :** Pour l'exécution des notebooks Python, offrant des ressources GPU gratuites pour l'entraînement des modèles.
- **Docker :** Pour exécuter facilement les premières étapes de la pipeline.
- **Discord :** Pour la communication en temps réel avec création de salons dédiés (données, modélisation, rédaction).

- **Choix technologiques :**

- **Python :** Langage principal pour sa richesse en librairies scientifiques (Pandas, NumPy) et frameworks de deep learning (PyTorch pour YOLOv5). Pour une liste des packages utilisés, voir Figure B.1.
- **OpenCV :** Pour le prétraitement des images (détection de flou, normalisation).
- **Reproductibilité :**
- **Gestion des données :**
- **Google Drive :** En lien avec Colab pour afin de récupérer sans faute les modèles lors des entraînements de modèles sur Colab .

Chapitre 2

Planification du projet, des rôles et des tâches

2.1 Planification du projet

- Le projet ne dispose pas d'une date limite stricte pour les différentes étapes intermédiaires, mais des objectifs sont définis chaque semaine. Une fois ces objectifs complétés, nous passons aux tâches suivantes. Une réunion est organisée au moins une fois par semaine pour faire le point sur l'avancement du projet et ajuster les priorités si nécessaire. Ces réunions ont lieu avant chaque rendez-vous avec notre professeur superviseur et le commanditaire, afin de nous assurer que nous restons alignés avec les attentes et les besoins du projet.

2.2 Tâches

- **Séparation des tâches :** Les différentes étapes du projet ont été réparties en fonction des compétences et des spécialisations des membres de l'équipe. Chaque membre s'est concentré sur une tâche spécifique avant que toutes les étapes soient réunies pour l'analyse finale.
- **Annotations d'images :** Une base de données d'images a été constituée en annotant les images d'oiseaux pour entraîner un modèle YOLOv5, afin d'améliorer la détection des oiseaux dans les images.
- **Uniformisation des images :** Les images ont été uniformisées pour garantir la cohérence du dataset et assurer une meilleure qualité d'entraînement pour le modèle de détection. Cela inclut la normalisation de la taille des images et l'ajustement des couleurs.
- **Entraînement du modèle YOLOv5 :** Un modèle de détection d'objets basé sur YOLOv5 a été entraîné pour identifier et localiser les oiseaux dans les images annotées, avec une attention particulière portée à l'amélioration des performances du modèle.
- **Détection de flou en arrière-plan :** Une étape de prétraitement a été mise en

place pour détecter et traiter le flou en arrière-plan des images, afin d'améliorer la précision du modèle et éviter des erreurs de détection dues à des images de mauvaise qualité.

- **Récupération des caractéristiques environnementales** : Les caractéristiques environnementales des images ont été extraites, telles que les types de végétation, les paysages et les éléments naturels présents, pour mieux comprendre l'influence de l'environnement sur les oiseaux.
- **Récupération des caractéristiques physiologiques des oiseaux** : Les informations physiologiques, telles que la taille, la couleur et la forme des oiseaux, ont été extraites des images pour analyser comment ces caractéristiques sont liées à leur environnement.
- **Extraction des coordonnées des photos et des données environnementales** : Les coordonnées géographiques des photos ont été utilisées pour récupérer les informations climatiques, les écosystèmes, les écorégions et les biomes selon différentes classifications associées à chaque image. Ensuite, les habitats spécifiques à chaque espèce ont été identifiés. Toutes ces données sont visibles sur des cartes avec le fond associé.
- **Calcul des similarités entre les caractéristiques des oiseaux et des arrière-plans** : A l'aide des caractéristiques que l'on a extraits des oiseaux mais aussi des arrière-plans, on calcule les similarités entre les oiseaux et les arrière-plans de chaque espèce afin de déterminer les interactions entre ceux-ci.

Chapitre 3

Acquisition, nettoyage, mise en forme des données et Analyse Exploratoire

3.1 Acquisition des données

Les images sont récupérées grâce à l'API de iNaturalist. Vous pouvez consulter la documentation officielle de l'API iNaturalist pour plus de détails à l'adresse suivante : <https://api.inaturalist.org/>.

3.2 Nettoyage et mise en forme des données

On récupère toutes les images sous leur format original et on enlève avec un modèle yolov5 toutes celles qui comportent des traces humaines.

3.3 Analyse exploratoire des données

La base de données nous intéressant est celle contenant toutes les images d'oiseaux de inaturalist qui contient 29.822.668 d'images au moment où on écrit ce rapport. De par la quantité de données et leur nature on ne peut pas vraiment faire d'analyse exploratoire des données générale et très pertinente.

On a dans un premier temps travaillé sur un échantillon réduit de cette base de données, afin de pouvoir développer notre pipeline assez rapidement. Nous exécuterons la pipeline avec un plus grand nombre de photos vers la fin du projet afin d'avoir des résultats.

Nous avons mis en place la possibilité de projeter sur des cartes les points correspondant aux différentes photos que l'on a téléchargé. Il est possible de projeter les points sur les fonds de carte des climats, des écosystèmes et des écorégions. Sur chaque point on pourra aussi retrouver d'autres informations en plus de celles-ci, comme l'habitat ou les biomes selon les classifications de Leemans, . Celà permettra de visualiser la distribution géographique des photos que l'on a, possiblement espèce par espèce, afin de s'assurer que l'on a assez de diversité dans nos données, et aussi d'expliquer certains résultats.

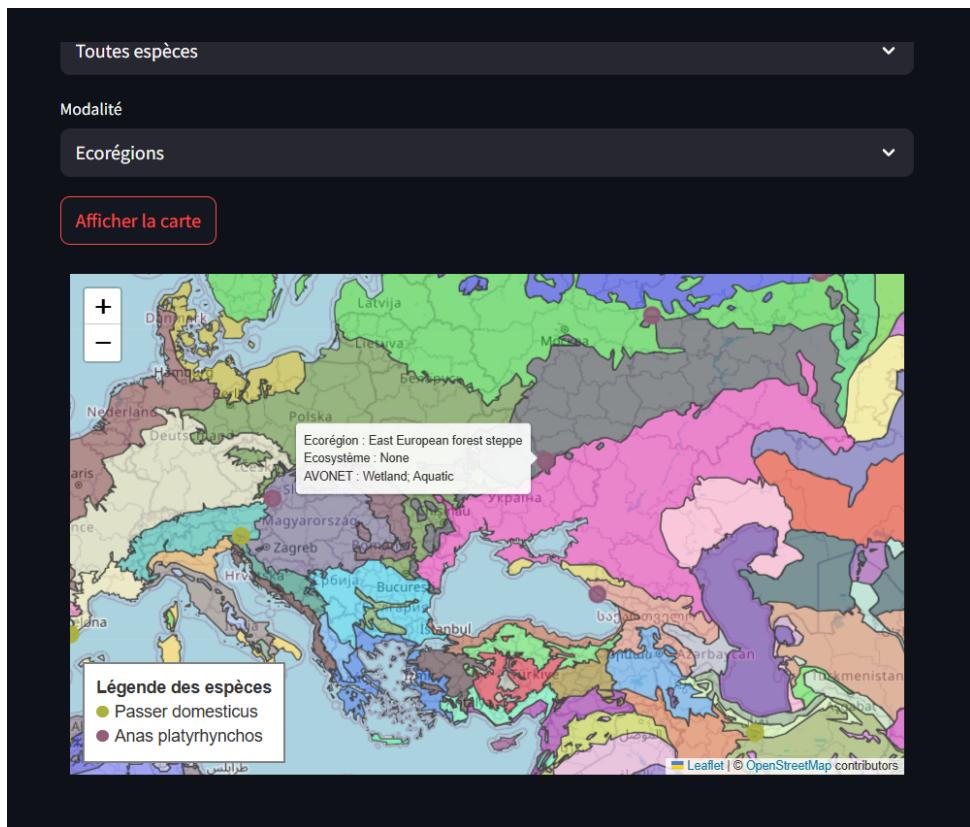


FIGURE 3.1 – Carte des écorégions générées par l’application

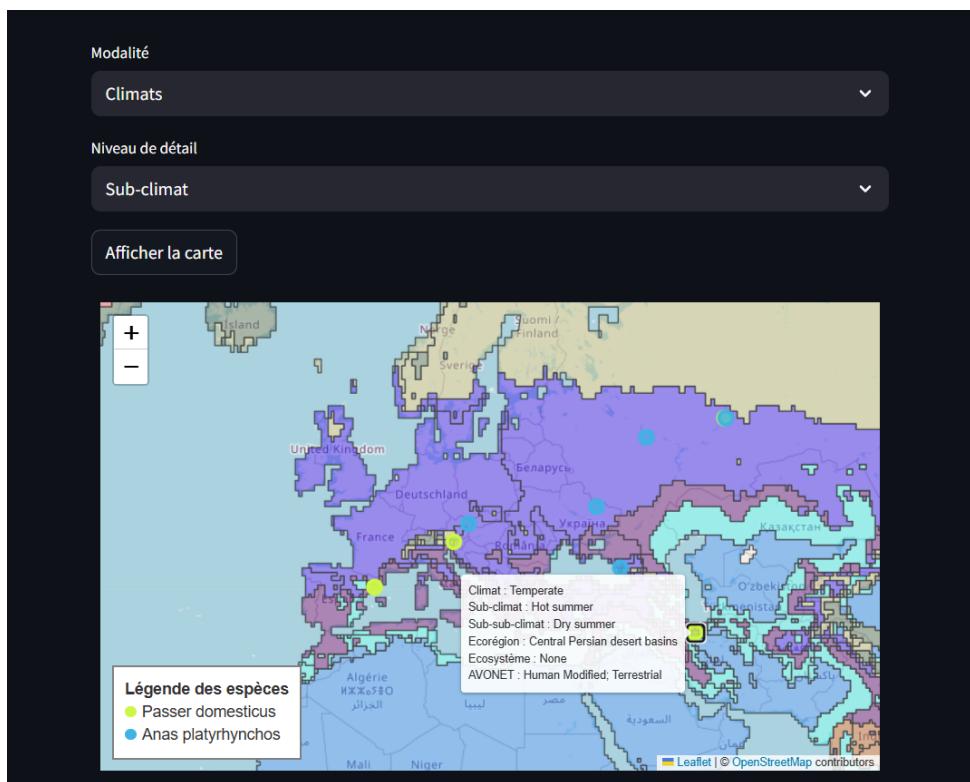


FIGURE 3.2 – Exemple d’une carte des sous-climats générée par l’application

Chapitre 4

Méthodologie

4.1 Pipeline, résultats et interprétation

Mise en place d'une pipeline : Le travail sur les features des oiseaux et les backgrounds a été facilité par la pipeline mise en place grâce à Docker. Il suffira d'indiquer le nombre d'espèces et le nombre d'images par espèce qu'on veut traiter et Docker se chargera d'exécuter séquentiellement les tâches qui vont être expliquées dans la suite.

4.1.1 Extraction d'oiseaux

Extraction d'oiseaux

Pour détecter et extraire efficacement les oiseaux des images, nous avons entraîné un modèle YOLOv5x spécifiquement dédié à cette tâche. Nous avons choisi YOLOv5x pour son excellent compromis entre précision et vitesse, malgré sa consommation accrue en ressources. Cette approche nous a permis d'obtenir des performances supérieures à un modèle générique.

— Phases d'entraînement :

1. **Entraînement initial** : 2000 images avec annotations manuelles
2. **Fine-tuning 1** : 4888 images d'entraînement / 1858 validation
3. **Fine-tuning 2** : 3657 images d'entraînement / 1214 validation
4. **Évaluation finale** : 615 images de test

— Configuration technique :

```
# Phase 1 - Premier entraînement
!python train.py --img 640 --batch 16 --epochs 100
--data data.yaml --weights yolov5x.pt

# Phase 2 - Fine-tuning
!python train.py --img 960 --batch 8 --epochs 50
--data data.yaml --weights runs/train/exp1/weights/best.pt

# Phase 3 - Fine-tuning 2
!python train.py --img 960 --batch 8 --epochs 100
--data data.yaml --weights runs/train/exp2/weights/best.pt
```

```
# Phase 4 - Finalisation
!python train.py --img 960 --batch 8 --epochs 150
--data data.yaml --weights runs/train/exp3/weights/best.pt
```

- Optimisation des paramètres d'entraînement :
 - Augmentation progressive de la résolution (640px → 960px) :
 - Problème initial : La taille native de 640px limitait la détection des petits oiseaux.
 - Solution : Augmentation graduelle pour :
 - Améliorer la détection des détails fins
 - Réduire les faux positifs sur fonds complexes
 - Permettre une meilleure généralisation
 - Compromis : Coût mémoire multiplié par 2.25 ($960^2/640^2$) mais nécessaire pour la qualité
 - Réduction de la taille des batchs (16 → 8) :
Les Batchs de 16 causaient une saturation de la mémoire du GPU rendant impossible de train un modèle sur collab donc on a reduit à 8.
 - Difficultés rencontrées :
 - Limitations matérielles : Quota GPU Colab
 - Déconnexions imprévues (perte de progression)
 - Temps d'entraînement élevé
 - Solutions implémentées :
 - Stratégie de sauvegarde incrémentielle (Google Drive)
 - Scripts de reprise automatique (checkpoints)
 - Optimisation mémoire : mixed-precision training
 - Parallelisation des prédictions (batch processing)

Performances

- Modèles évalués :
 - Modèle A : Premier entraînement (YOLOv5s)
 - Modèle B : Fine-tuning initial
 - Modèle C : Fine-tuning avancé
 - Modèle D : Modèle finalisé
- Métriques clés :

Métrique	Modèle A	Modèle B	Modèle C	Modèle D
mAP@0.5	0.83	0.922	0.991	0.988
Précision	0.779	0.957	0.986	1
Rappel	0.96	1	0.99	0.98
F1-score	0.79	0.98	0.98	0.98
Seuil optimal (F1)	0.285	0.276	0.164	0.148

TABLE 4.1 – Comparaison des performances entre les différentes itérations du modèle

— Comparaison visuelle des performances :

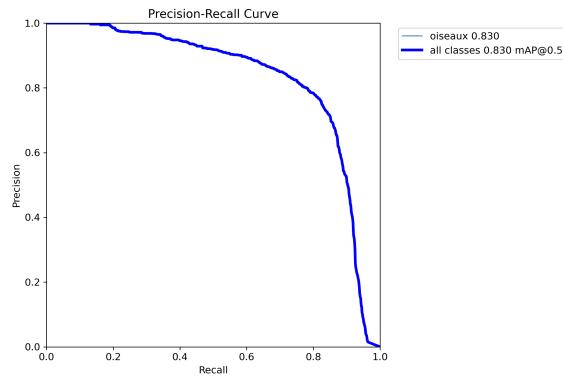


FIGURE 4.1 – Courbe Precision-Recall du modèle A (moins performant)

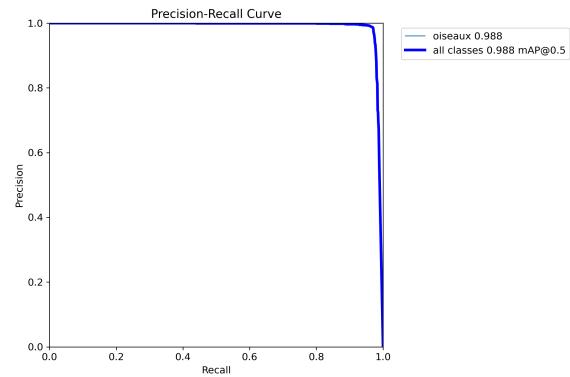


FIGURE 4.2 – Courbe Precision-Recall du modèle D (meilleur)

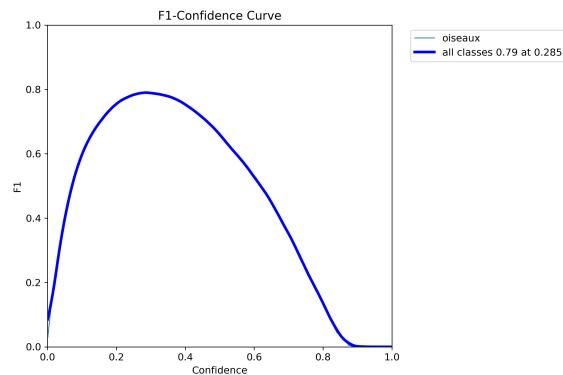


FIGURE 4.3 – F1-score du modèle A (moins performant)

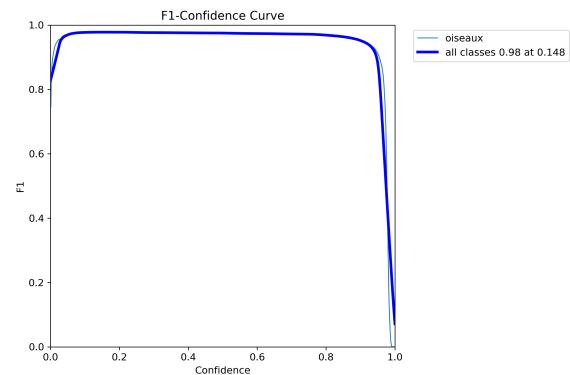


FIGURE 4.4 – F1-score du modèle D (meilleur)

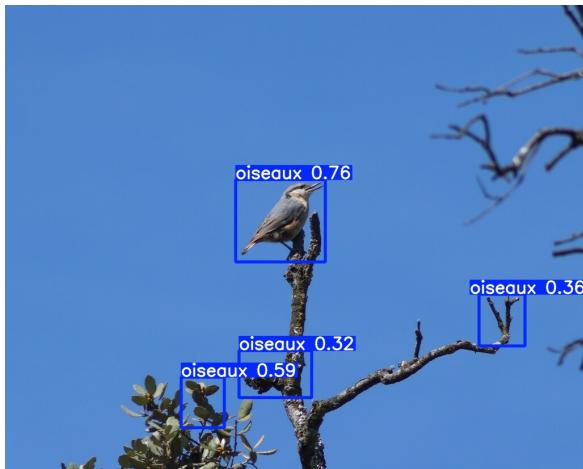


FIGURE 4.5 – Prédictions du modèle A (moins performant)



FIGURE 4.6 – Prédictions du modèle D (meilleur)

- **Principales observations :**
 - Progression significative des performances entre le Modèle A et D
 - Le Modèle D atteint une précision parfaite (1.0)
 - Rappel maintenu à un niveau élevé (>0.98) tout au long de l'amélioration
 - mAP@0.5 proche de 1 pour les modèles finaux
 - Le modèle A et D détecte tout deux l'oiseau mais le D a une bien meilleure valeur de confiance et le A condond des feuilles et morceaux de branches avec des oiseaux.
- **Avantages du modèle final :**
 - Détection excellente des petits oiseaux
 - Réduction drastique des faux positifs
 - Robustesse aux conditions variables (éclairage, orientation)

4.1.2 Détection de traces humaines :

Dans le cadre du projet, on a entraîné un réseau de neurones basé sur le modèle YOLOv11, une version plus récente et plus performante proposée par Ultralytics, afin de détecter la présence de traces humaines sur des images.

On a d'abord constitué une base de données en téléchargeant des images via une API, en utilisant des mots-clés comme "voiture", "humain", "mains", etc. Après le téléchargement, on a réalisé une annotation manuelle en indiquant, pour chaque image, la présence ou l'absence d'êtres humains.

Un premier entraînement du modèle a été effectué, mais les performances initiales se sont révélées insuffisantes : le modèle confondait régulièrement des oiseaux avec des traces humaines. Pour améliorer cette différenciation, j'ai enrichi la base de données avec plus de 500 images d'oiseaux annotées spécifiquement. Un nouveau cycle d'entraînement a alors permis d'obtenir de meilleurs résultats, bien que le modèle ne soit pas encore totalement optimal.

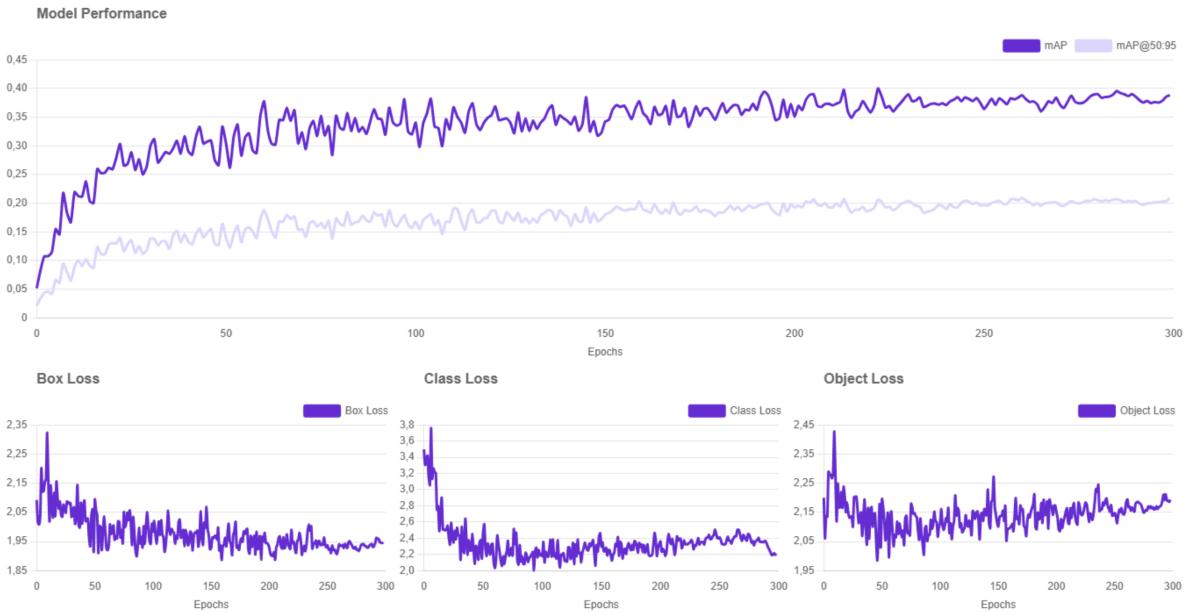


FIGURE 4.7 – Métriques YOLOv11

Pour améliorer encore les performances, il aurait été pertinent de mettre en place des techniques de data augmentation et d'utiliser un éventail plus large de mots-clés lors de la collecte d'images.

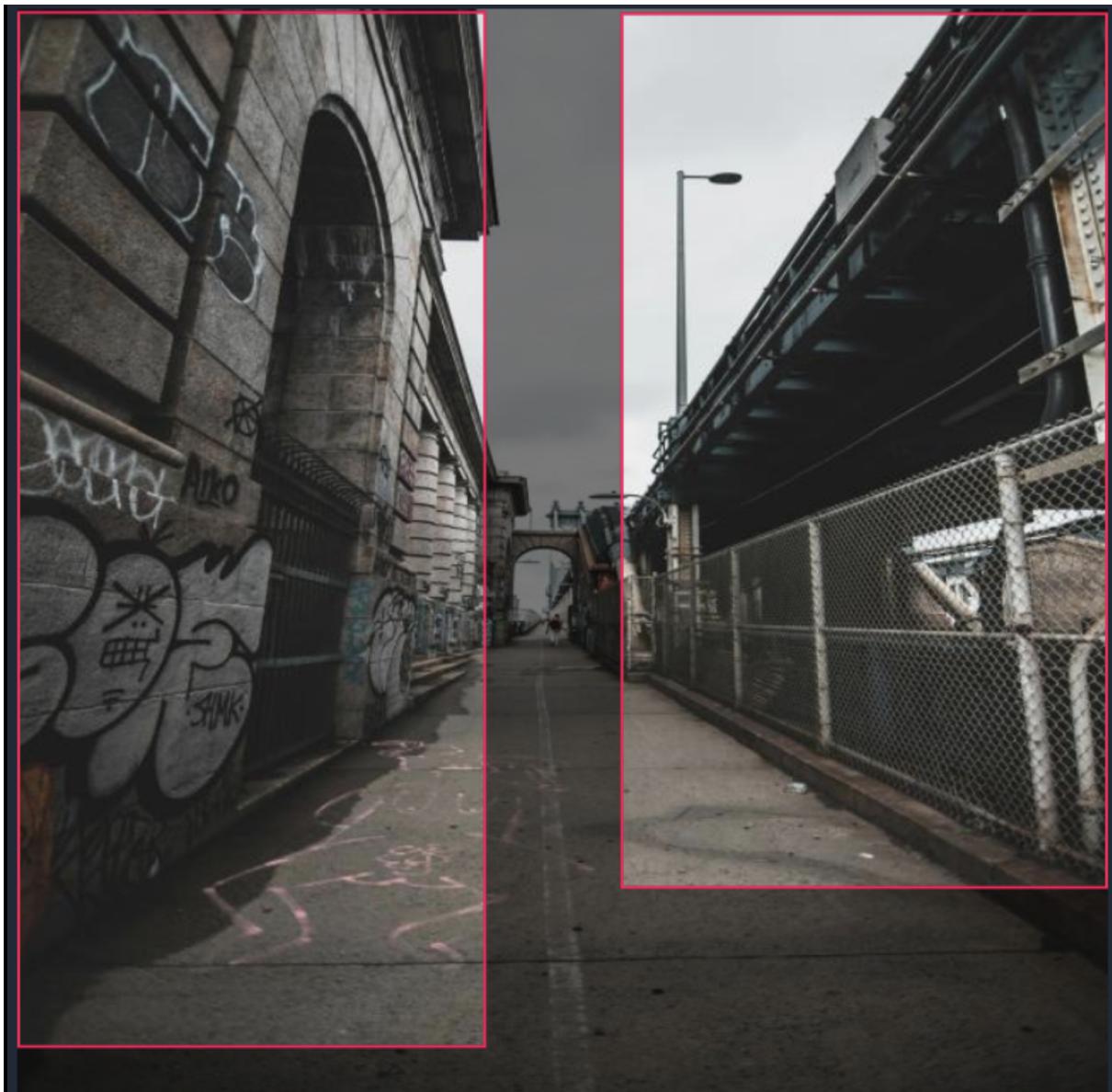


FIGURE 4.8 – Annotation traces humaines

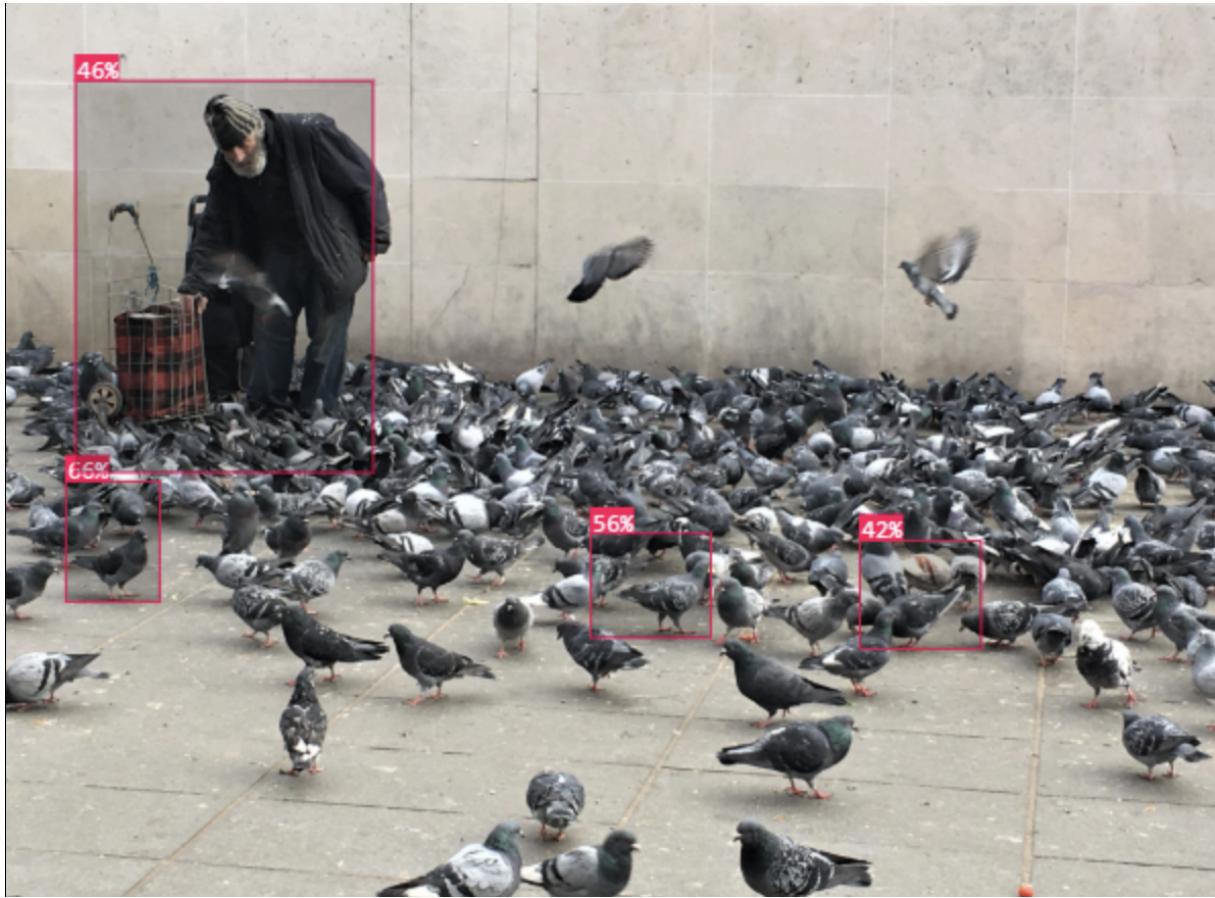


FIGURE 4.9 – Détection traces humaines

4.1.3 Détection de zones floues :

Nous avons initialement exploré des méthodes classiques pour détecter les contours, en nous appuyant sur des filtres comme Sobel, Laplace ou des analyses de fréquence via la transformée de Fourier. Bien que ces approches soient relativement simples à mettre en œuvre, leurs limites apparaissent rapidement dès que les images présentent des arrière-plans uniformes — comme le ciel, la mer ou des murs unis — ou lorsqu'elles manquent clairement de texture. Dans ces cas, les détecteurs génèrent trop de faux positifs, ou au contraire, ne parviennent pas à repérer des zones floues peu contrastées comme des petites zones floues au sein d'une image globalement nette.

Pour pallier ces difficultés, nous avons opté pour une approche par apprentissage profond, plus robuste et capable de généraliser sur des cas variés.

À partir d'une base de données d'images claires , nous avons appliqué un flou gaussien sur une zone rectangulaire aléatoire de 250×250 pixels. Le noyau de flou (de taille impaire comprise entre 9 et 29) était tiré au hasard pour générer une diversité de niveaux de flou. Les images modifiées ont ensuite été annotées manuellement sur Roboflow, pour détecter la zone floutée. Nous avons fait une cross validation.

Pour l'entraînement, nous avons utilisé le modèle pré-entraîné YOLOv5. Le lancement de l'entraînement s'est fait avec une image d'entrée redimensionnée à 640×640 , un batch size de 16 (soit environ 47 itérations par époque pour 747 images), et un total de 100

époques.

Les résultats obtenus sont très satisfaisants : après 100 époques, notre détecteur atteint une précision de 92,8 % et un rappel de 88,9 % sur le jeu de validation. Nous avons pu visualiser l'évolution des pertes (box_loss, obj_loss, cls_loss) et des métriques (précision, rappel, mAP).

Pour améliorer cette détection sur des images d'oiseaux, il sera essentiel d'annoter précisément les zones floues de vraies photos d'oiseaux (par exemple arrières-plans volontairement floutés). Le modèle pourra ainsi mieux apprendre.

Au-delà du flou gaussien, il conviendra de générer et d'annoter des flous de mouvement (traînées, flou directionnel) ainsi que d'éventuelles pixelisations, afin de permettre au réseau de reconnaître plus finement différentes natures de flou.

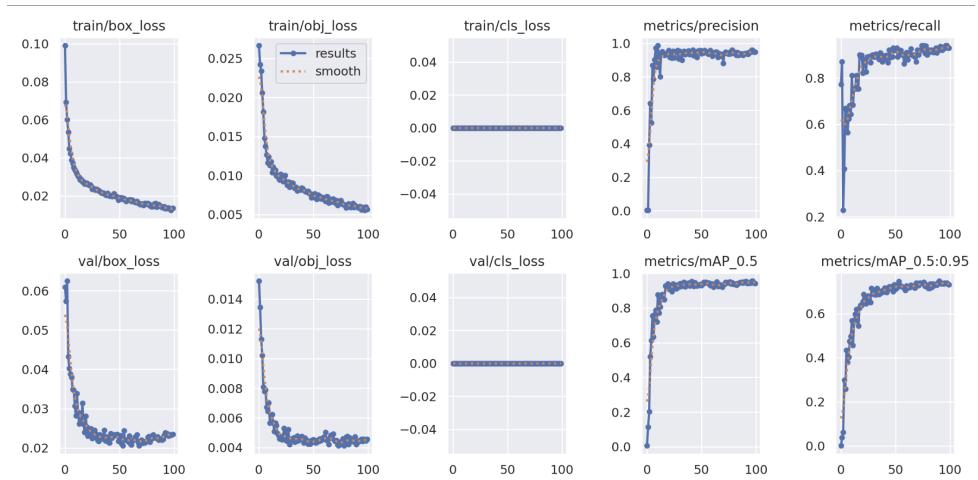


FIGURE 4.10 – Métriques



FIGURE 4.11 – Exemple de détection de zones floues

4.1.4 Extraction Background :

Dans le cadre de ce projet, l'un des objectifs importants était d'extraire automatiquement les régions de fonds (backgrounds) qui se trouvent dans les images d'oiseaux, en excluant ces derniers. Cette phase nous donnera, plus tard, la possibilité d'analyser les caractéristiques de l'environnement en comparaison avec ceux des oiseaux . À cet effet, une stratégie a été mise en œuvre en vue d'identifier les régions intéressantes de chaque image en éliminant à la rigueur les régions où se trouve l'oiseau.

La première étape de cette procédure repose sur l'utilisation des fichiers .txt de YOLOv5 associés à chaque image. Ces fichiers contiennent les coordonnées des boîtes englobantes (bounding boxes) indiquant l'emplacement de l'oiseau dans l'image. À partir de ces données, nous générerons un masque permanent (permanent_mask) dans lequel chaque boîte est élargie d'une marge de 15 %. Cette marge est nécessaire pour éviter que les régions extraites ne contiennent une portion de l'oiseau, ce qui fausserait l'analyse du fond.

En plus, un masque de visite (visited_mask) est utilisé pour se souvenir des régions déjà visitées, afin d'éviter d'extraire deux fois la même zone. La mise en œuvre de l'extraction repose elle-même sur un processus d'exploration guidé :

- Des points au hasard sont choisis uniquement dans les régions qui ne sont pas masquées, c'est-à-dire, ne sont pas occupées par l'oiseau (masque permanent), ni déjà visitées (masque de visite). Dans chaque point tiré, on effectue une exten-

sion de la zone en les quatre directions principales (haut, bas, gauche, droite) jusqu'à atteindre un masque permanent ou le bord de l'image. Pour qu'une zone soit conservée, elle doit répondre à certains critères : elle doit être suffisamment large (au moins 1 % de l'image) et être remplie. L'extraction stoppe dès qu'il y en a quatre zones invalides consécutives ou dès qu'on a plus de zones non masquées. Chaque zone valide est enregistrée sous forme d'image indépendante.

D'autres approches ont été envisagées, telles que l'utilisation d'un modèle de segmentation YOLOv8 couplé à une technique d'inpainting pour remplir les zones où se trouve l'oiseau, ou encore un masquage brut avec un carré noir. Toutefois, ces alternatives présentent des limites : l'inpainting peut générer des zones de fond artificielles peu fiables pour une analyse objective, tandis que le masquage brut va biaiser l'extraction des features sur l'environnement. En comparaison, la méthode que nous avons retenue, bien qu'un peu plus coûteuse en calcul, permet de produire des backgrounds naturels, sans altérations artificielles, garantissant une meilleure qualité d'analyse visuelle par des descripteurs.

4.1.5 Features ResNet :

Au premier semestre, des features HOG combinées avec HSV avaient été utilisées et il a été remarqué que le vecteur qui résultait de cette extraction était en grande dimension ; de l'ordre de 24303.

Au cours du second trimestre, il nous a été suggéré d'utiliser à la place des features ResNet-50 pré-entraîné sur la base ImageNet qui sont de plus petite dimension ; 2048.

La couche fully connected est enlevée car elle fournirait des features transformées dans le but de la classification de taille 2048, ce qui n'est pas pertinent car on ne veut pas faire de prédiction.

L'avant dernière couche du réseau quant à elle a des features transformées spécialement pour la classification mais ce qu'on veut faire c'est plutôt l'extraction des features (patterns) qui vont servir dans les analyses de similarité sur les oiseaux et les backgrounds.

Il faut donc s'assurer que l'on a fait l'extraction de features ResNet de la dernière couche de convolution et y appliquer un average pooling (un maxpooling ou global pooling) pour garder les features saillantes des images.

Le choix fait pour le nombre d'espèces a été de 200 avec 10 images chacune. Ils sont téléchargés avec leur metadata, la détection par YOLOv5 (pour de meilleure résultat doit être basée sur un modèle performant), extraction des features ResNet des oiseaux et des backgrounds, les analyses sur les features extraites (distribution, KMeans,...).

4.1.6 Données géographiques et cartes :

Afin de mieux comprendre les différentes caractéristiques des oiseaux, nous avons récupéré les coordonnées géographiques des photos lors de leur téléchargement. Ces coordon-

données permettent de connaître les caractéristiques de l'environnement de chaque oiseau, car les informations extraites de l'arrière-plan des images ne donnent qu'une partie de la vision. Ces données viennent donc enrichir notre analyse de l'influence du contexte sur les attributs des oiseaux.

Dans ce but, nous avons sélectionné plusieurs types de couches géographiques :

- **Climats** : classification de Köppen–Geiger (format **shapefile**).
- **Écorégions** : classification du WWF (format **shapefile**).
- **Écosystèmes** : classification IUCN niveau 3 (format **raster**).
- **Biomes** : classifications de Leemans, Higgins, Friedl et Olson (format **raster**).
- **Habitats** : classification AVONET (format Excel).

Nous avons ajouté à l'application la possibilité d'extraire automatiquement ces informations à partir d'une liste de points fournie par l'utilisateur. Le résultat se présente sous la forme d'un tableau récapitulatif exportable, et peut également être visualisé sur un fond de carte choisi selon la modalité souhaitée.

Spécie	Latitude	Longitude	Climat	Sub-climat	Sub-sub-climat	Ecorégion	Ecosystème	AVONET	Leemans
<i>Passer domesticus</i>	35.7096	51.4218	Temperate	Hot summer	Dry summer	Central Persian desert basins	TF1.7 Boreal	Human Modified; Terrestrial	Cool temperate steppe
<i>Passer domesticus</i>	46.367	14.1081	Temperate	Warm summer	Without dry season	Pannonic mixed forests	TF1.7 Boreal	Human Modified; Terrestrial	Cool temperate wet forest
<i>Passer domesticus</i>	57.9929	56.1561	Boreal	Warm summer	Without dry season	Sarmatic mixed forests	TF1.7 Boreal	Human Modified; Terrestrial	Cool temperate moist forest
<i>Passer domesticus</i>	41.2141	1.727	Temperate	Hot summer	Dry summer	None	TF1.7 Boreal	Human Modified; Terrestrial	Warm temperate dry forest
<i>Passer domesticus</i>	41.2172	1.7285	Temperate	Hot summer	Dry summer	Northeastern Spain and Southern France Mediterranean forests	TF1.7 Boreal	Human Modified; Terrestrial	Warm temperate dry forest

FIGURE 4.12 – Exemple d'une partie d'un tableau générée par l'application

Toutefois, les traitements sur couches raster se sont révélés très lents, ce qui nuit à l'ergonomie globale de l'application. Pour pallier ce problème, ces extractions raster sont désormais optionnelles, afin de garantir des temps de réponse acceptables et une navigation fluide pour la majorité des fonctionnalités. Ce fut aussi un frein dans la programmation de l'extraction de ces données, car il fallait toujours attendre plusieurs dizaines de minutes après avoir lancé l'exécution du code pour voir une erreur. L'extraction via le dockerfile est aussi très lente, et il faudrait donc trouver des moyens d'optimiser ces opérations afin qu'elles soient faisables dans un temps acceptable.

Ces visualisations sont intéressantes pour repérer des tendances générales comme par exemple l'absence de couverture d'une grande étendue géographique, mais ne seront pas forcément claires avec des milliers de photos de la même espèce. Il faudrait pouvoir "regrouper" certains points selon le nombre de points présents sur une zone, selon le niveau de zoom.

4.1.7 Similarités

Le principal objectif de notre projet est de détecter les similarités entre les caractéristiques des oiseaux et celles de leur environnement. Pour ce faire :

1. Extraction des *features* des oiseaux (comme décrit précédemment).
2. Extraction des *features* des arrière-plans pour chaque photo.
3. Moyennage des *features* d'un même arrière-plan lorsqu'une image comporte plusieurs régions distinctes.
4. Calcul de la distance euclidienne entre les vecteurs de l'oiseau et ceux de chaque arrière-plan.

Nous avons testé d'autres mesures (par exemple, similarité cosinus), mais avons finalement retenu la distance euclidienne pour sa simplicité et son efficacité suffisante.

À partir de ces distances, nous construisons :

- une table représentant la similarité des *features* de chaque oiseau vis-à-vis de son propre arrière-plan.
- une matrice de confusion exprimant la similarité entre les features de chaque espèce avec les features des arrière-plans de chaque espèce.

species_1	Aix_sponsa	Anas_platyrhynchos	Ardea_alba	Ardea_cinerea	Ardea_herodias	Astur_cooperii	Balearophus_bicolor	Branta_canadensis	Calypte Anna
Aix_sponsa	22.3213	28.3222	29.203	36.4721	28.1907	35.2582	35.066	31.1457	36.5199
Anas_platyrhynchos	28.3222	26.8693	28.8255	34.9578	25.8897	31.0711	35.2442	30.4609	36.0476
Ardea_alba	29.203	28.8255	19.3831	35.1632	28.9894	32.5299	30.3802	29.6384	34.4888
Ardea_cinerea	36.4721	34.9578	35.1632	10.8116	34.4243	29.5137	37.0407	38.8555	31.8353
Ardea_herodias	28.1907	25.8897	28.8994	34.4243	30.345	32.5854	35.7067	32.3326	36.3626
Astur_cooperii	35.2582	31.0711	32.5299	29.5137	32.5854	None	37.4902	34.4868	28.2333
Balearophus_bicolor	35.066	35.2442	30.3802	37.0407	35.7067	37.4902	None	38.2052	40.2487
Branta_canadensis	31.1457	30.4609	29.6384	38.8555	32.3326	34.4868	38.2052	24.5976	34.1668
Calypte Anna	36.5199	36.0476	34.4888	31.8353	36.3826	28.2333	40.2487	34.1668	35.7481

FIGURE 4.13 – Exemple d'un tableau des similarités généré par l'application

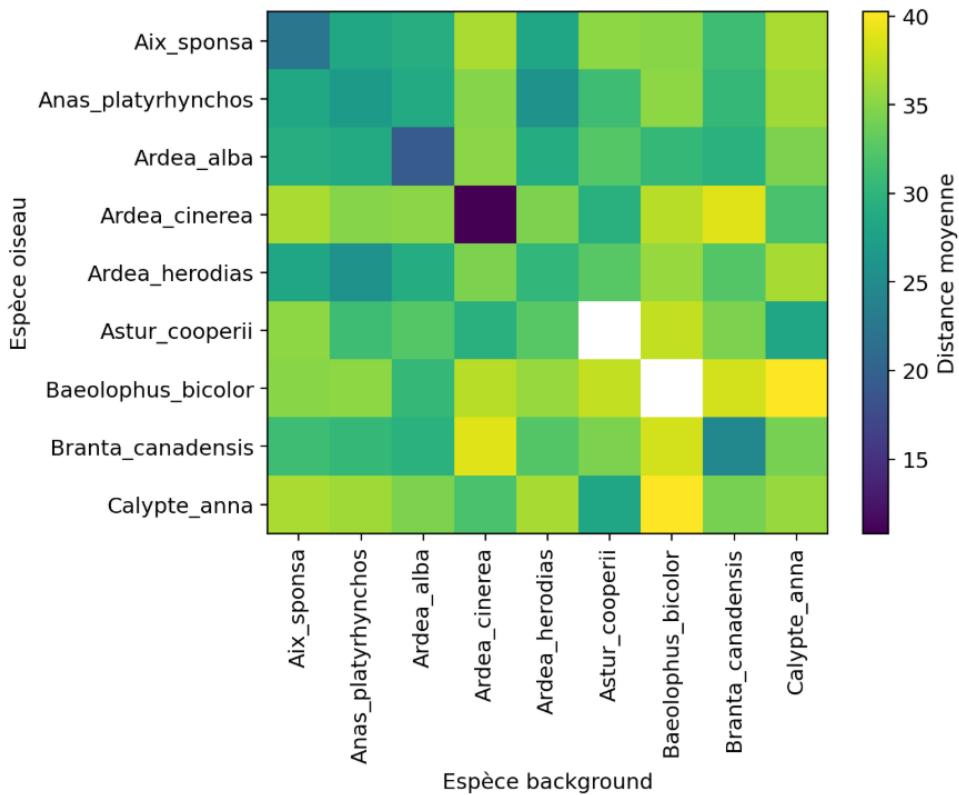


FIGURE 4.14 – Exemple d'une matrice de confusion générée par l'application

4.2.6.1 Méthodologie

La présente étude comprend deux volets complémentaires :

Analyse statistique des arrière-plans Pour évaluer la cohérence visuelle des fonds selon le contexte environnemental, chaque image de fond a été traitée par un réseau convolutionnel pré-entraîné (ResNet) afin d'en extraire un embedding de dimension 512. Deux distributions de distances euclidiennes ont alors été comparées :

- *Distances intra-environnement* : embeddings issus du même type d'habitat ;
- *Distribution inter-environnement* : bootstrap ($n \approx 1000$ échantillons sans remise) d'embeddings provenant d'habitats différents.

Analyse statistique des oiseaux On a appliqué une démarche analogue aux embeddings extraits des oiseaux :

1. *Distances intra-espèce* : individus de la même espèce.
2. *Distances inter-espèce (bootstrap)* : mille échantillons d'individus d'autres espèces.

4.2.6.2 Résultats et discussion

Les distributions intra-environnement et intra-espèce se révèlent systématiquement plus concentrées que leurs homologues inter, avec des $p < 0,05$ et des Z-scores significativement négatifs. Le score de silhouette moyen (0,3) suggère une séparation modérée des clusters, malgré des profils visuels rares. L'association cluster-climat met en évidence des tendances selon l'habitat et l'espèce.

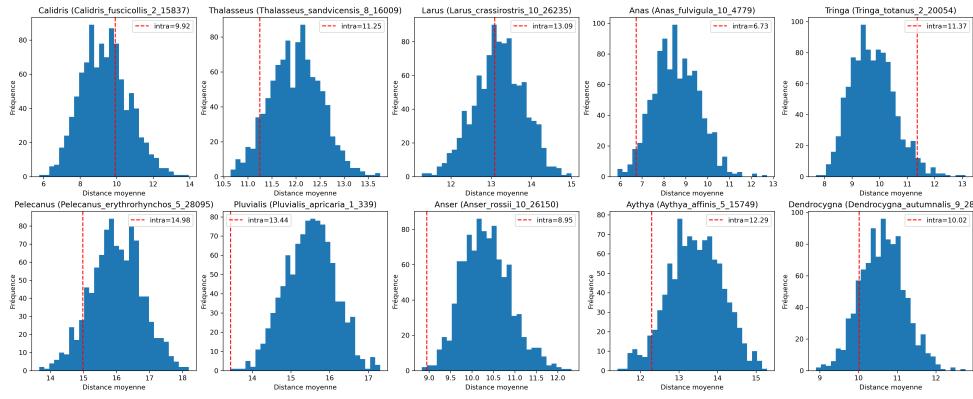


FIGURE 4.15 – Histogrammes des distances nulles (bleu) et intraclassse (rouge pointillé) pour 10 espèces d’oiseaux.

Interprétation pour les oiseaux :

Si un oiseau a des images très homogènes dans l'espace des features, sa ligne rouge se trouvera nettement « à gauche du pic bleu » ; à l'inverse, une espèce très hétérogène verra sa ligne rouge se placer « à droite » de la distribution nulle.

- **Espèces homogènes** (Anas, Anser, Pluvialis, Thalasseus...) : la faible distance intraclassse (ligne rouge très à gauche) suggère un fort regroupement des embeddings, donc une bonne similarité visuelle entre images.
- **Espèces hétérogènes** (Tringa, Larus) : la ligne rouge au centre ou à droite indique que la variation interne rivalise avec la variation inter-espèces — signal potentiel d'outliers ou de poses très diverses.
- **Cas intermédiaires** (Calidris, Aythya, Dendrocygna...) : intraclassse proche du centre de la distribution nulle, indiquant une cohésion moyenne.

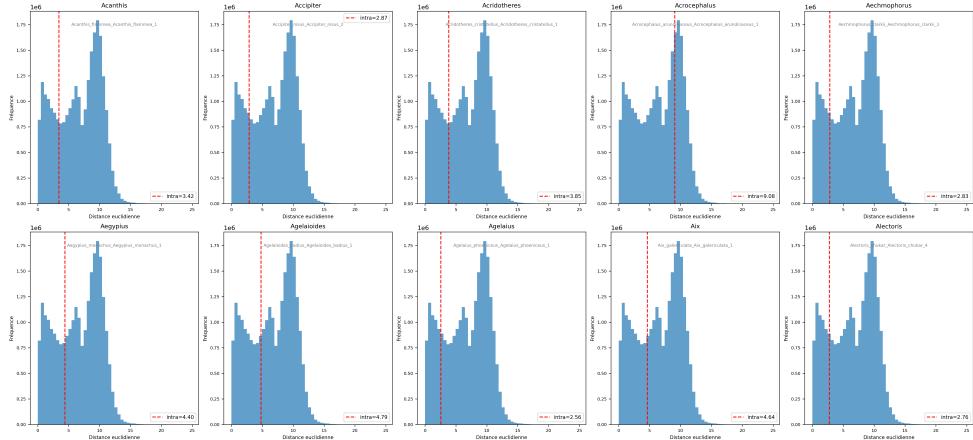


FIGURE 4.16 – Distribution nulle des distances euclidiennes entre embeddings moyens de fonds d’espèces différentes et distance intraclassse pour 10 scènes.

Interprétation de la distribution des *backgrounds*

La figure 4.16 présente pour dix scènes de fonds différents :

- en bleu, l'histogramme de la **distribution nulle** : toutes les distances euclidiennes entre embeddings moyens de fonds issus d'espèces différentes,

- en rouge pointillé, la **distance intraclasses moyenne** : la moyenne des distances entre le premier crop et les autres crops de la même image d'origine.

On observe une distribution multimodale :

1. Un premier pic vers 0–5 : scènes très similaires (par exemple, même type d'habitat ou luminosité proche).
2. Un second pic autour de 6–12 : habitats un peu plus éloignés (eau vs végétation, zones ouvertes vs boisées).
3. Une queue au-delà de 12 : scènes radicalement différentes (paysages contrastés).

Pour neuf des dix scènes étudiées, la ligne rouge se situe dans le **premier pic**, ce qui signifie que les augmentations (crops) génèrent des embeddings *très cohérents* : leur variabilité interne est nettement plus faible que la variabilité inter-scènes.

En revanche, pour la scène correspondant à *Acrocephalus_arundinaceus*, la distance intraclasses (≈ 9 M) tombe dans le second pic. Cela suggère que ses crops couvrent deux types de contenu visuel distincts (par exemple, sol+végétation), d'où une dispersion interne comparable à la dispersion entre espèces.

Conclusion pratique : la plupart des fonds restent stables face aux crops et conservent un embedding homogène, garantissant une bonne robustesse du modèle à l'augmentation. Néanmoins, pour les scènes composites (e.g. *Acrocephalus_arundinaceus*), il conviendra peut-être de restreindre la zone de découpe ou d'ajuster la taille des crops afin de préserver la cohérence de l'embedding.

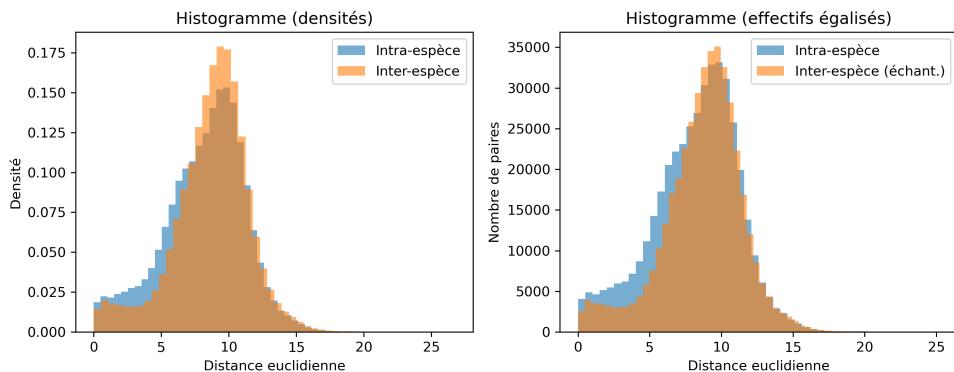


FIGURE 4.17 – Comparaison des distances euclidiennes intra-espèce vs inter-espèce entre embeddings de fonds : (a) densités ; (b) effectifs égalisés.

Distance observée : Pour chaque espèce s , on calcule

$$d_{\text{obs}}(s) = \left\| \underbrace{\bar{v}_s^{\text{bird}}}_{\text{embedding moyen de l'oiseau}} - \underbrace{\bar{v}_s^{\text{bg}}}_{\text{embedding moyen de son habitat}} \right\|.$$

Distribution nulle : On génère aléatoirement n_{null} paires croisées ($\bar{v}_A^{\text{bird}}, \bar{v}_B^{\text{bg}}$) avec $B \neq A$, et on mesure les distances $\|\bar{v}_A^{\text{bird}} - \bar{v}_B^{\text{bg}}\|$ pour construire la référence.

Si $\overline{d_{\text{obs}}}$ est *significativement* plus faible que la moyenne de la distribution nulle, alors les oiseaux présentent un mimétisme visuel de leur habitat.

2. Résultats sur les embeddings de fonds

- **Intra-espèce** (mêmes espèces, images différentes) : 429316 paires, min/max/moyenne = 0.00/25.21/8.20.
- **Inter-espèce** (espèces distinctes) : 118887663 paires, min/max/moyenne = 0.00/26.75/8.69.

a) Forme des distributions (densités) Le pic intra-espèce (6–9) est légèrement à gauche du pic inter-espèce (8–10), suggérant une homogénéité un peu plus forte au sein d'une même espèce.

b) Effectifs égalisés Après sous-échantillonnage des paires inter-espèce pour égaliser les volumes, on observe :

- Concentration des intra-espèce sous 10,
- Quelques intra-espèce très distants (> 15),
- Quelques inter-espèce très proches (< 2).

3. Conclusion Les fonds de même espèce sont un peu plus homogènes (pic à gauche), mais la variance importante et le chevauchement signifient que la distance euclidienne seule n'est pas un séparateur parfait. La différence moyenne ($\overline{d_{\text{intra}}} = 8.20$ vs $\overline{d_{\text{inter}}} = 8.69$) et le fort chevauchement indiquent que la distance euclidienne seule n'est pas suffisante pour discriminer les habitats par espèce. Pour tester l'hypothèse de mimétisme, il convient donc de comparer directement $\{d_{\text{obs}}(s)\}$ à la distribution nulle oiseau–habitat.

Chapitre 5

Conclusion

Ce projet a permis de développer une méthodologie complète pour analyser les relations entre les oiseaux et leur environnement, bien que notre hypothèse initiale sur le rôle de la sélection sexuelle n'ait pu être confirmée de manière concluante. Les avancées significatives réalisées comprennent :

- La création d'une pipeline robuste intégrant :
 - Détection précise d'oiseaux
 - Extraction fiable des arrière-plans
 - Analyse statistique des similarités
- Le développement d'outils interactifs pour :
 - Visualiser les distributions géographiques
 - Explorer les relations oiseaux-environnement
 - Analyser les clusters morphologiques

Les limitations rencontrées, notamment liées aux biais des données et aux contraintes matérielles, ouvrent des perspectives intéressantes pour des travaux ultérieurs. Les pistes d'amélioration identifiées concernent principalement :

- L'amélioration des modèles de détection (flou, traces humaines)
- L'exploitation approfondie des métadonnées géographiques
- L'amélioration des calculs de distribution
- L'optimisation des performances de l'application Streamlit notamment sur l'aspect géographique

Bien que nos résultats n'ai pas pu permettre d'établir un lien clair entre les caractéristiques des oiseaux et leur environnement dans un contexte de sélection sexuelle, la méthodologie développée constitue une base solide pour :

- De futures recherches en écologie évolutive
- L'étude d'autres taxons ou écosystèmes
- Le développement d'outils d'analyse bio-informatique

Ce travail démontre la valeur des approches computationnelles pour aborder des questions complexes en biologie évolutive, tout en soulignant l'importance d'une collaboration étroite entre informaticiens et biologistes pour interpréter correctement les résultats.

Bibliographie

- [1] iNaturalist. (2022). Documentation de l'API iNaturalist. <https://api.inaturalist.org/>
- [2] Ultralytics. (2021). YOLOv5 Documentation. <https://docs.ultralytics.com/>
- [3] Endler, J. A. (1984). Natural and sexual selection on color patterns in poeciliid fishes. *Environmental Biology of Fishes*, 9(2), 173-190.
- [4] Bradbury, J. W., & Vehrenberg, S. L. (2011). *Principles of Animal Communication*. Sinauer Associates.
- [5] Hirsch, B. T., Kays, R., Pereira, V. E., & Jansen, P. A. (2024). A synthesis of camera trap data reveals negative effects of human activity on terrestrial mammals across the globe. *Ecological Monographs*, 94(1), e1615. <https://esajournals.onlinelibrary.wiley.com/doi/epdf/10.1002/ecm.1615>

Annexe A

Code d'extraction des métadonnées

```
import requests
from PIL import Image

def get_bird_species():
    species_list = []
    page = 1

    while len(species_list) < num_species:
        url = "https://api.inaturalist.org/v1/taxa"
        params = {
            'taxon_id': 3, # ID taxonomique pour la classe "Aves"
            'rank': 'species',
            'per_page': 100, # Nombre d'espèces par page
            'page': page
        }

        response = requests.get(url, params=params)
        if response.status_code == 200:
            data = response.json()
            if 'results' in data:
                species_list.extend(data['results'])
            else:
                break
        else:
            print(f"Erreur lors de la récupération des espèces : {response.status_code}")
            break

    # Arrêter si moins de 100 résultats sont retournés (fin des résultats)
    if len(data['results']) < 100:
        break

    page += 1

return species_list[:num_species]
```

figure : Extrait du code Python pour l'extraction des données via l'API iNaturalist

Annexe B

Liste des packages utilisés

```
geopandas
pandas
shapely
rasterio
pyproj
requests
torch
torchvision
Pillow
openpyxl
folium
streamlit
numpy
opencv-python
werkzeug
scikit-learn
```

FIGURE B.1 – Capture d'écran du fichier requirement.txt