

# CHALLENGE MIASHS

IDENTIFICATION DES COLLEMOBOLES

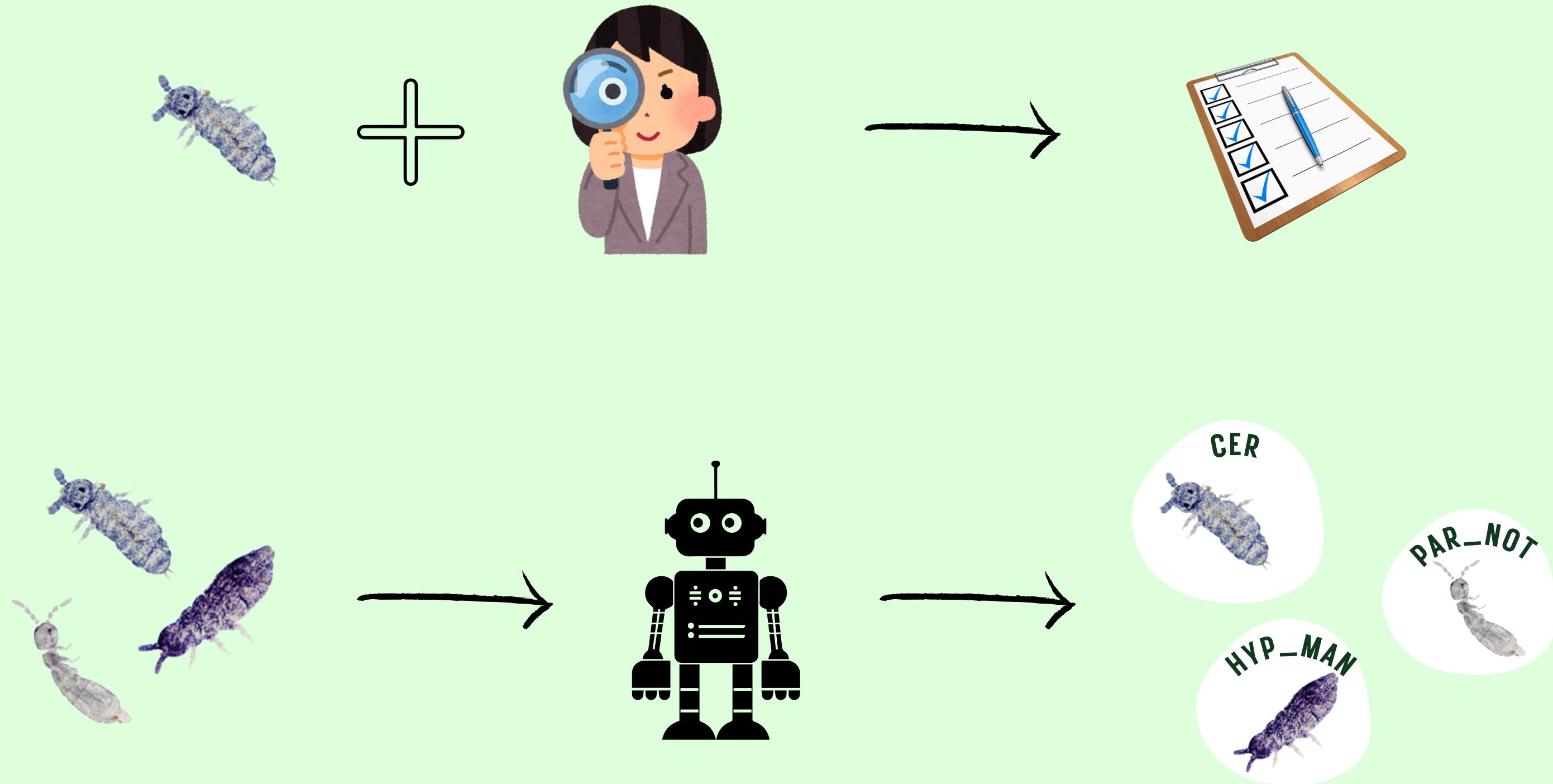
Groupe n°1 - Thomas Ayrivié, Brice Grivet, Karim Tchere, Fehizoro Randriamisanta

-----  
Master 2 MIASHS  
-----

Université de Montpellier Paul-Valery



# Contexte



# Les objectifs

01

## LABELLISATION

Classer les espèces de Collemboles en 9 classes avec F1-macro comme métrique.

02

## CORRECTION DES DONNÉES

Résoudre les désaccords entre experts dans les annotations d'entraînement.

03

## EXPLICABILITÉ

Comprendre les critères de décision des modèles.

# Sommaire



- 1 Exploration des données
- 2 Approche méthodologique
- 3 Analyse des résultats
- 4 Explicabilité du modèle
- 5 Propositions d'améliorations
- 6 Conclusion



# EXPLORATION DES DONNÉES

# Exploration des données

## Données d'entraînement:

- 1 117 images ( $3072 \times 2048$ )
- Un fichier txt par image:

5_5_5_5	Ecopic	2018	0.66	0.47	0.30	0.17
7_7_7_7	Ecopic	2018	0.89	0.53	0.17	0.55

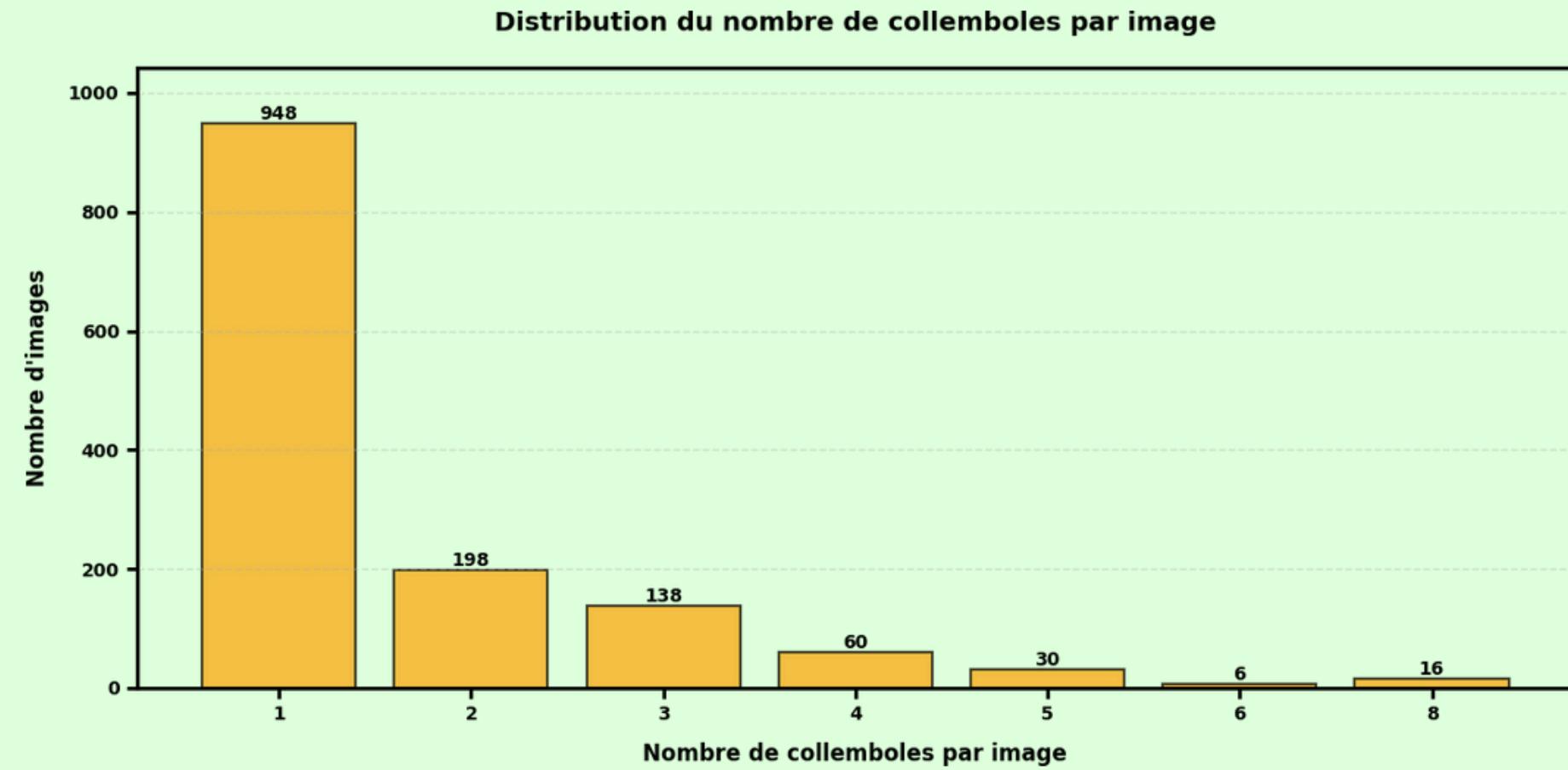
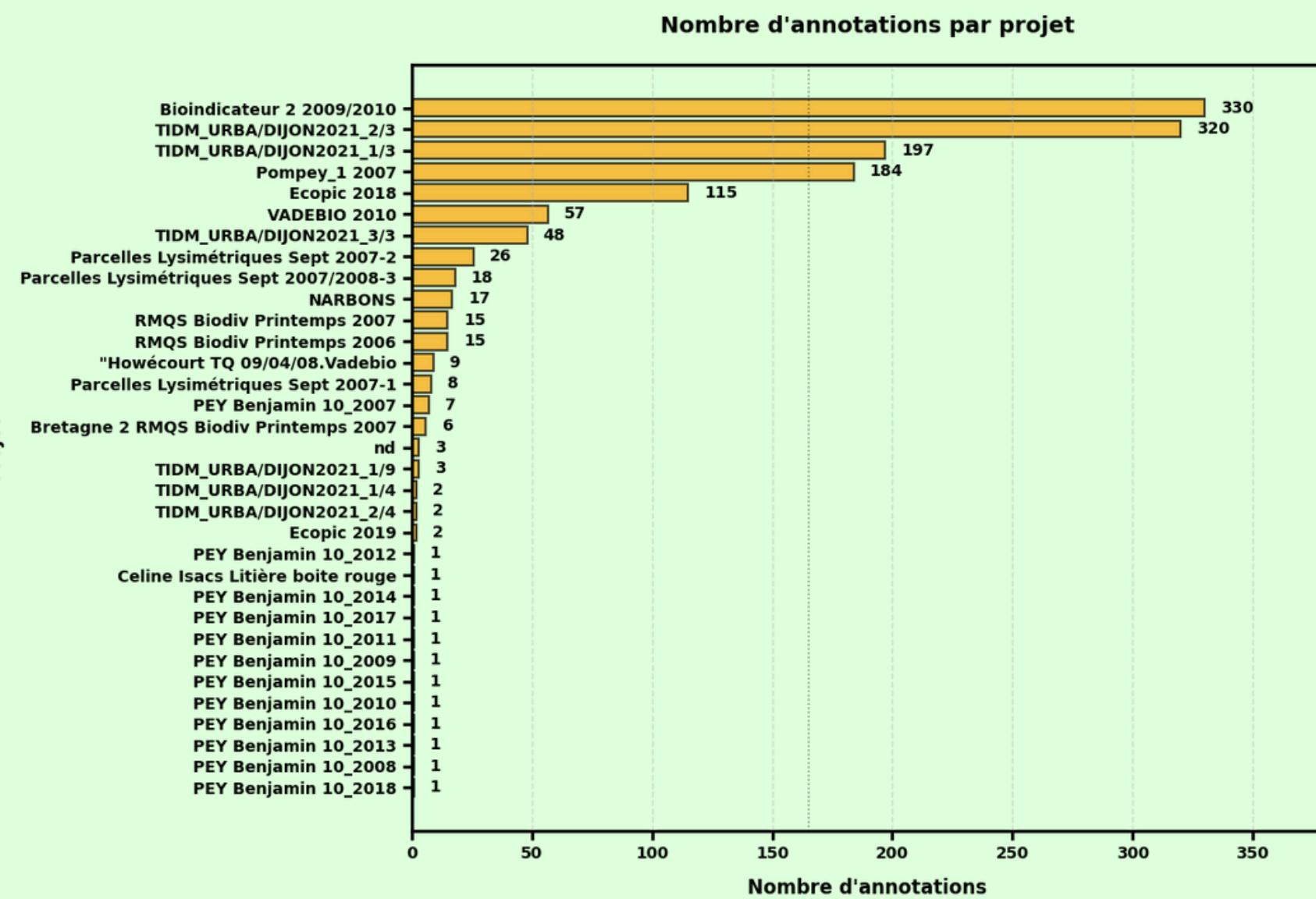
## Données de test:

- 1 344 imagettes (taille variable)
- Vérité terrain (ex: test ADN)



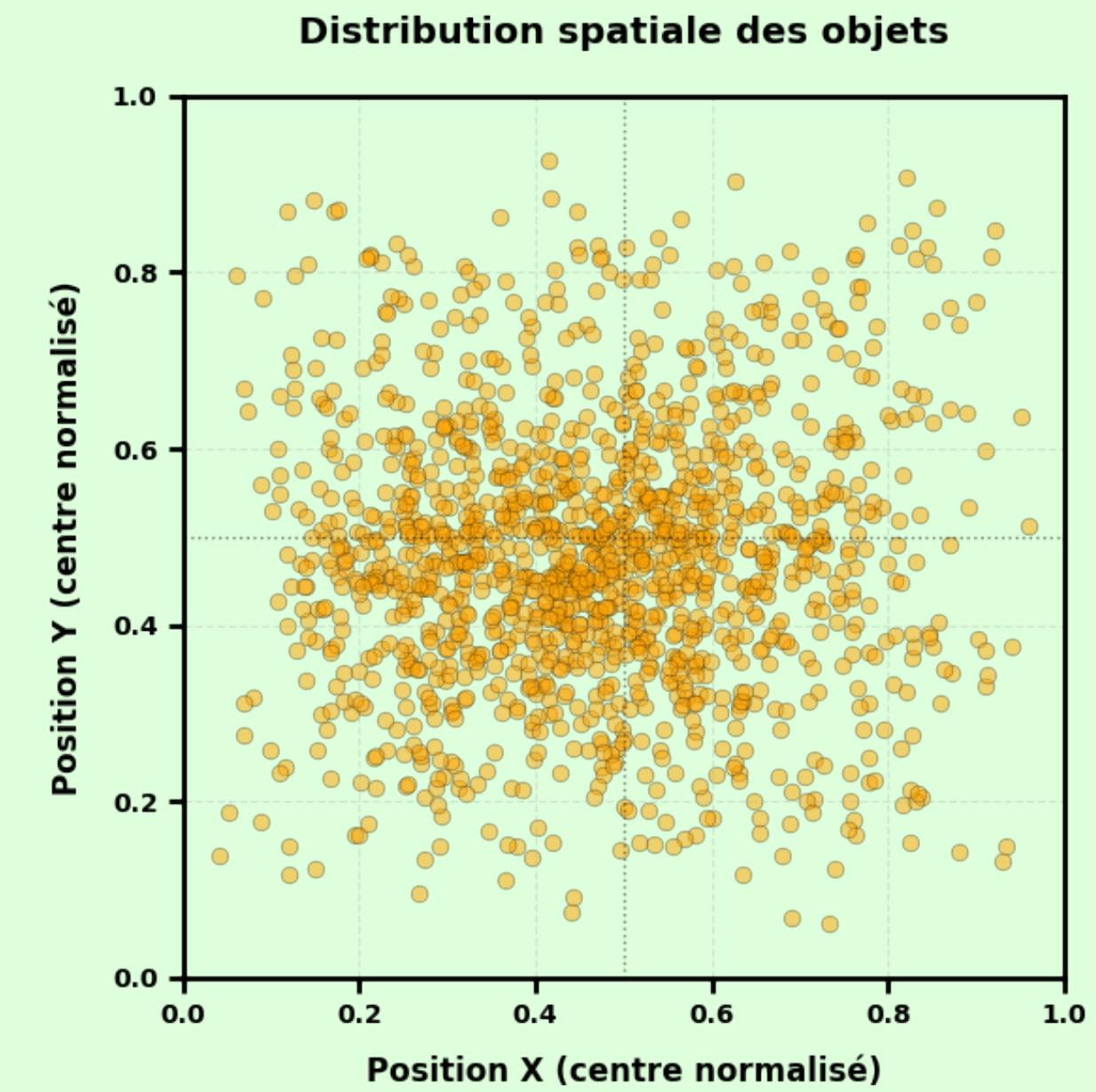
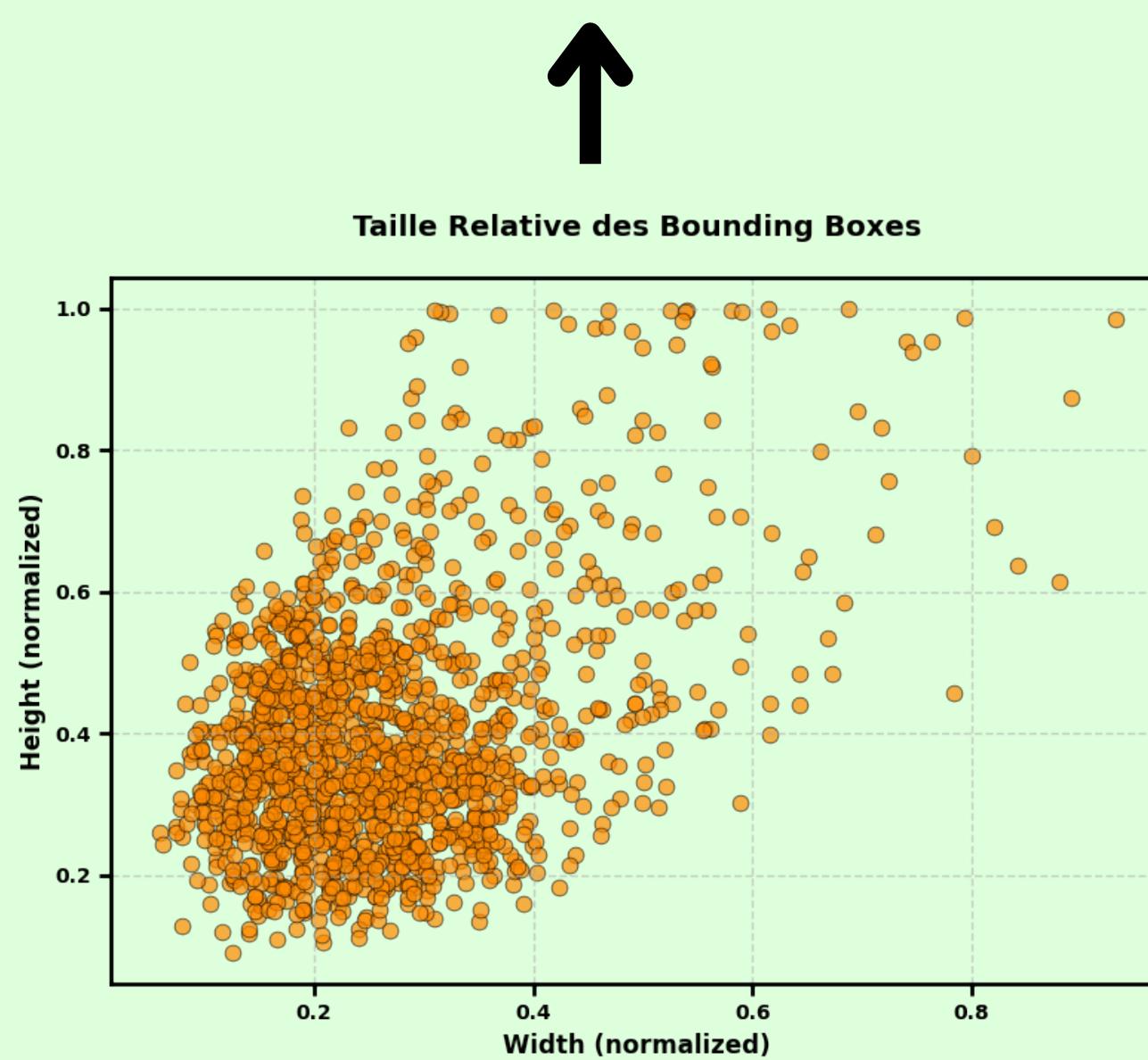
# Exploration des données

Projet



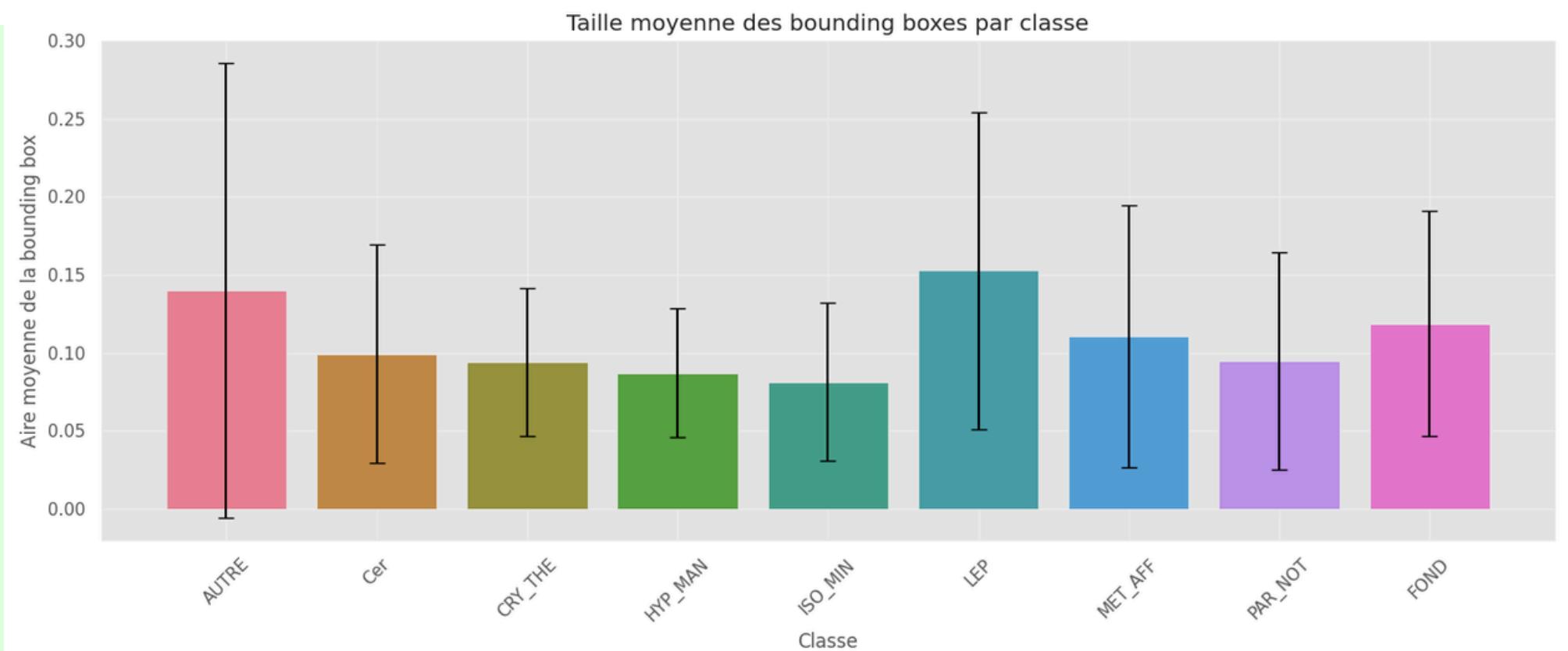
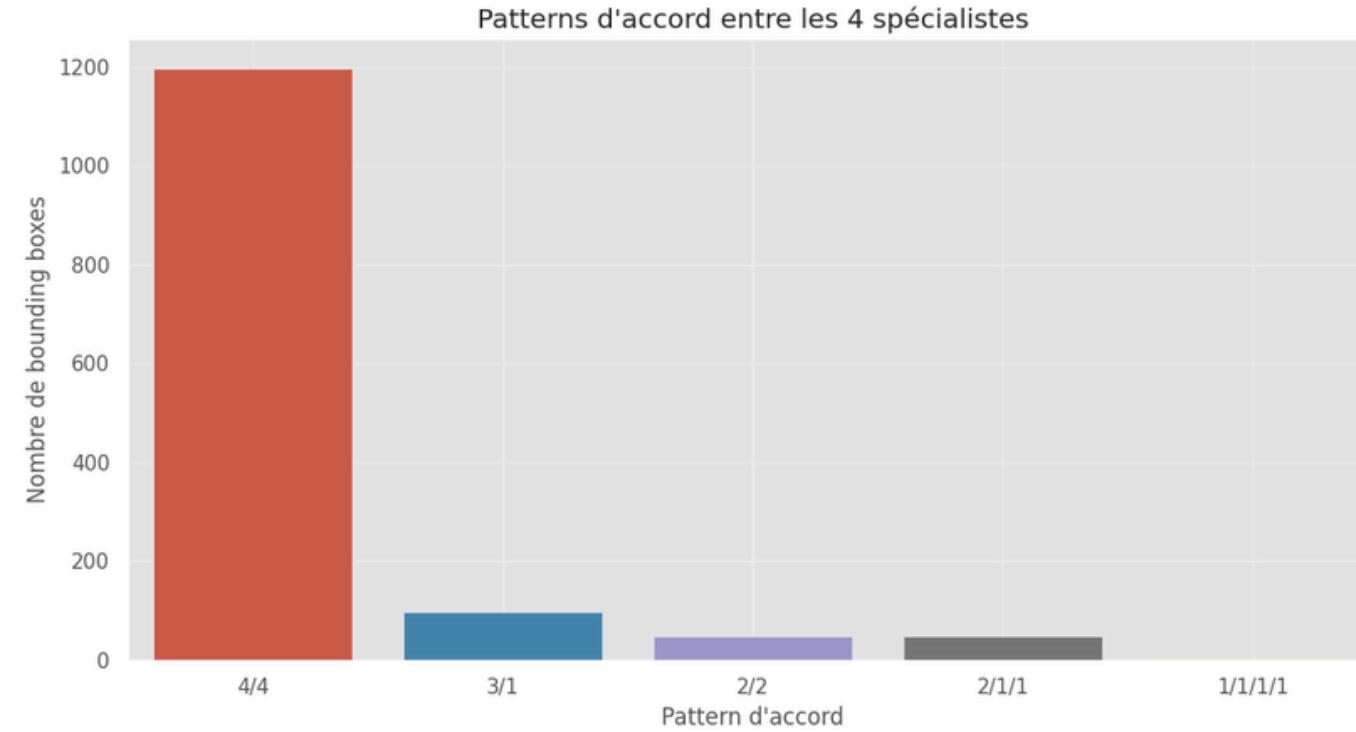
# Exploration des données

Objets détectés  
généralement de  
petites tailles.



Distribution relativement  
aléatoire sur l'image.

# Exploration des données

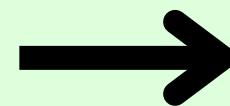
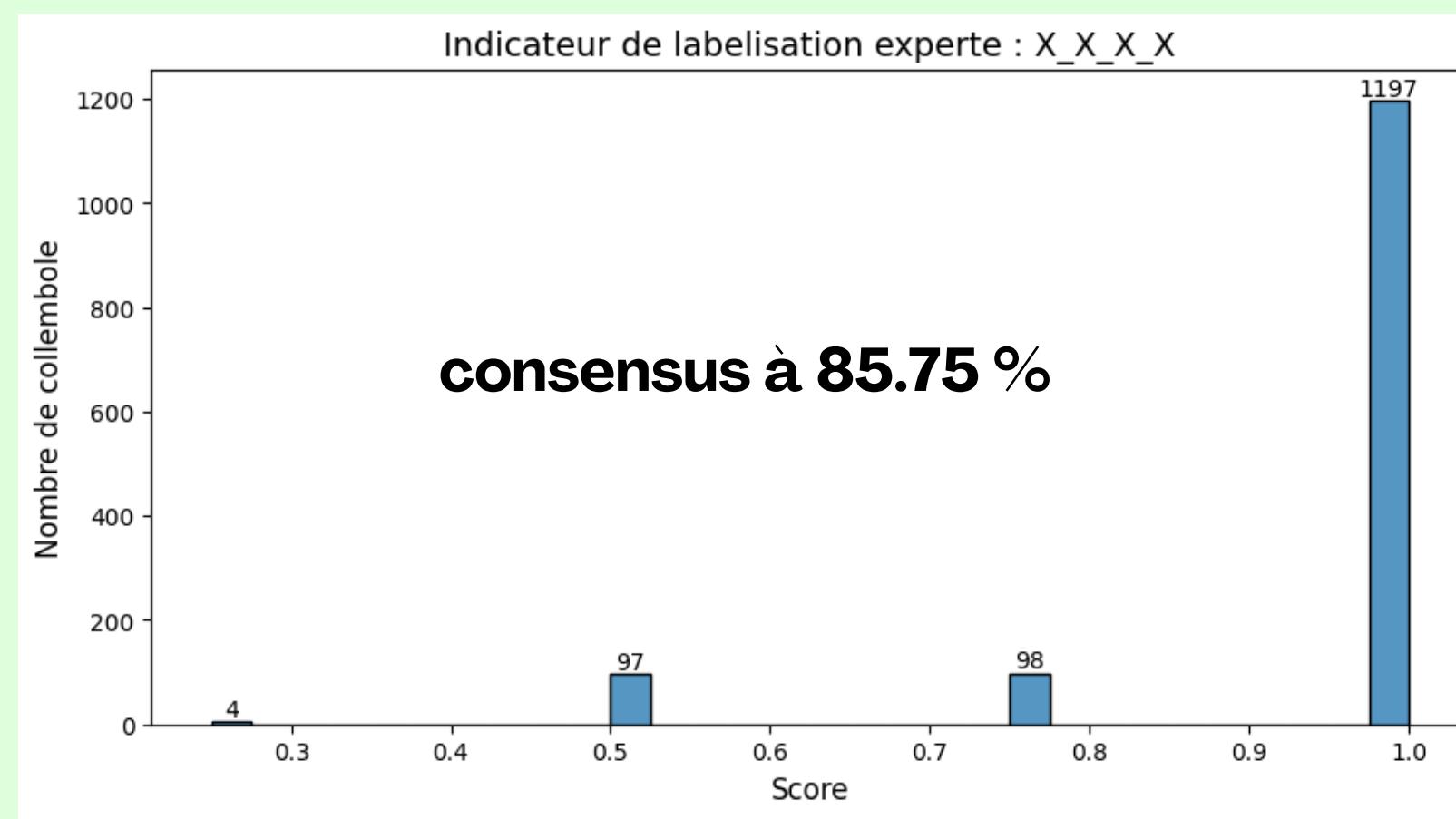




# APPROCHE MÉTHODOLOGIQUE

# Prétraitement

- Découpage des bounding boxes
- Extraction de nouveaux “fonds” (classe 8)
- Calcul d'un indicateur d'appartenance de classe

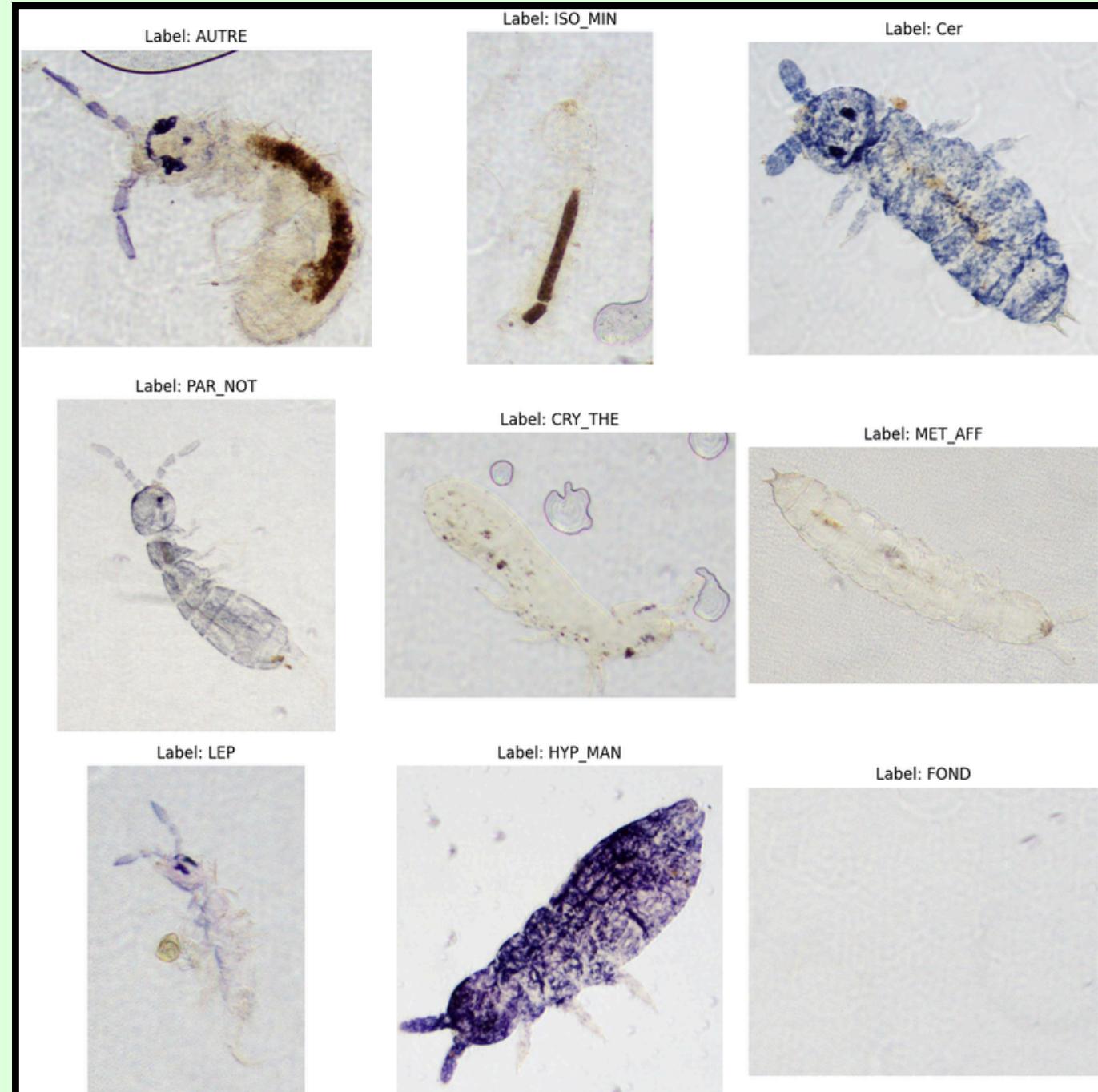


Seuil fixé à 1  
Base d'entraînement : 1197  
images.

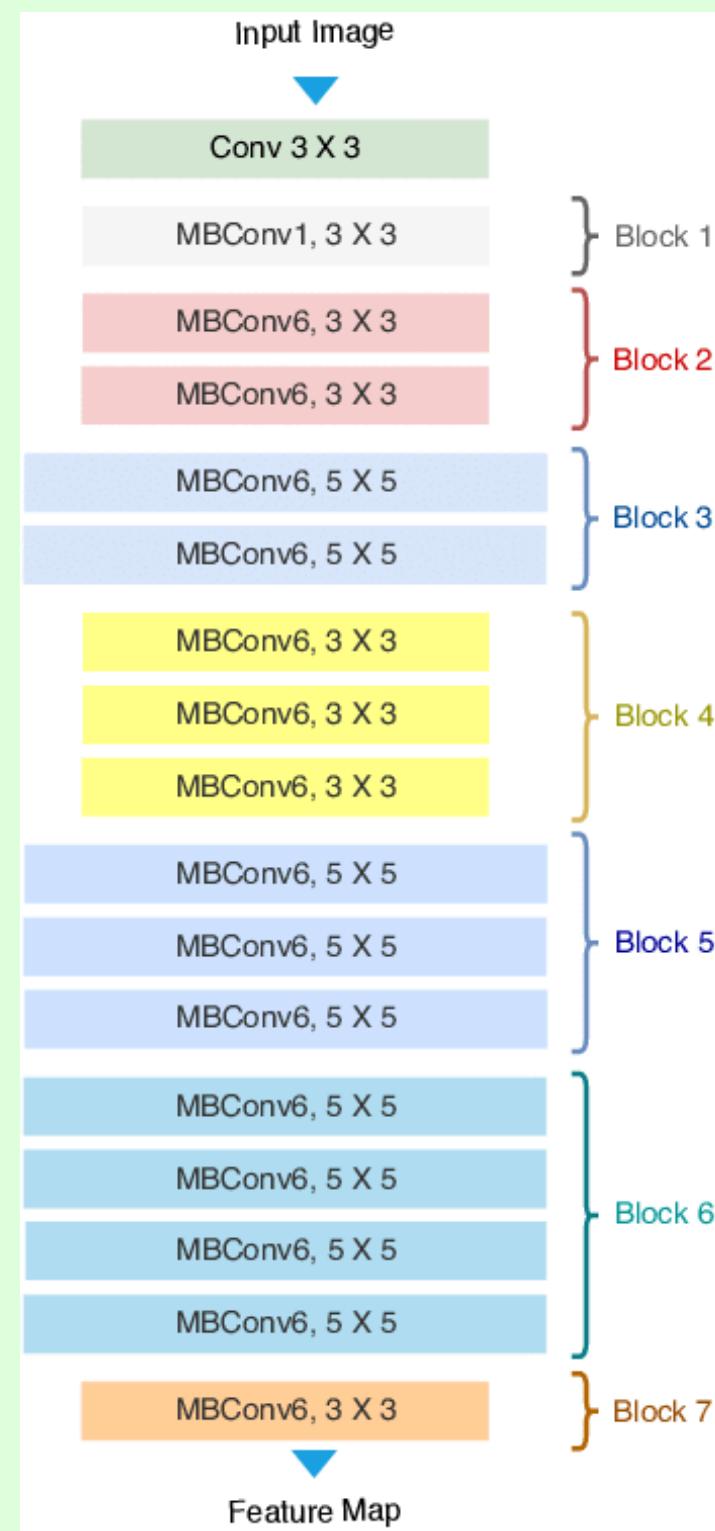
# Augmentation

- Rotation aléatoire jusqu'à 360° avec remplissage par réflexion.
- Recadrer aléatoirement une zone de l'image, en conservant entre 80% et 100% de la taille originale.
- Retournement horizontal (HorizontalFlip) avec une probabilité de 50%.
- Retournement vertical (VerticalFlip) avec une probabilité de 30%.
- Modification aléatoire de la luminosité, du contraste et de la teinte.
- Application d'un flou gaussien.

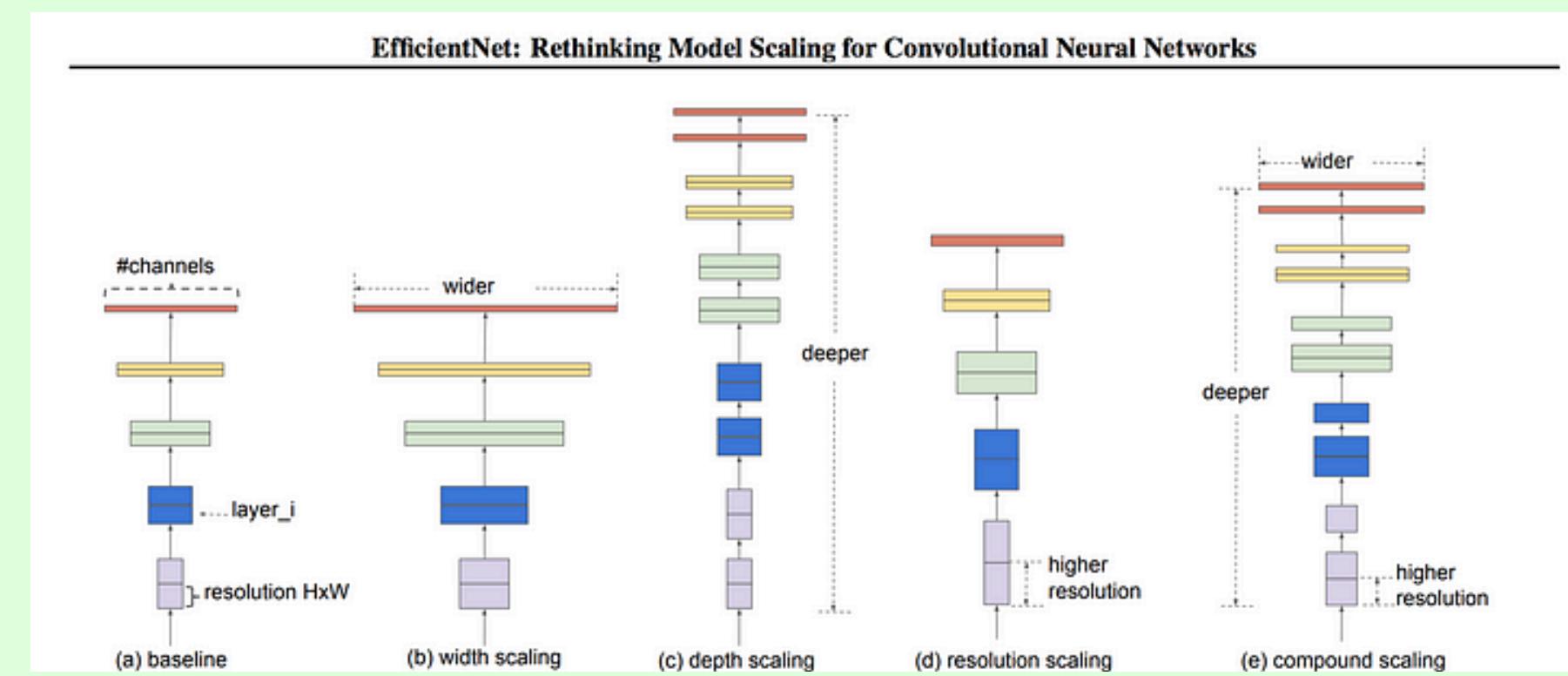
→ Application d'un poids inversement proportionnel à la fréquence de chaque classe par la suite.



# EfficientNet



- Famille de modèles de réseaux de neurones convolutifs (pré-entraîné sur ImageNet)
- Utilisation d'une approche d'échelle composite pour équilibrer profondeur, largeur et résolution, avec des couches MBConv et l'optimisation Squeeze-and-Excitation.
- EfficientNet comprend huit modèles, de EfficientNet-BO à EfficientNet-B7
- **EfficientNet-B5 :** offre un excellent compromis entre précision élevée et faible coût computationnel





# **ANALYSE DES RÉSULTATS**

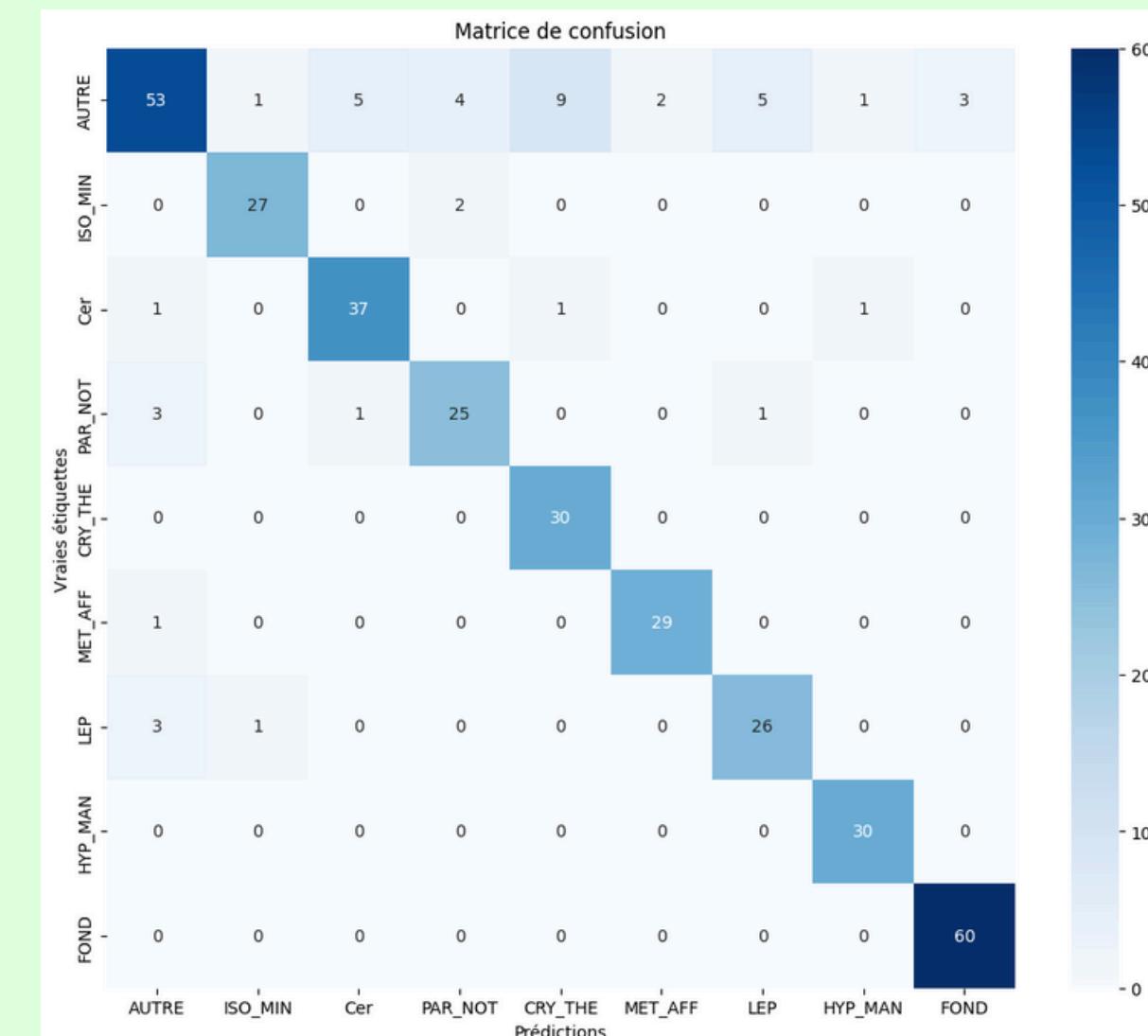
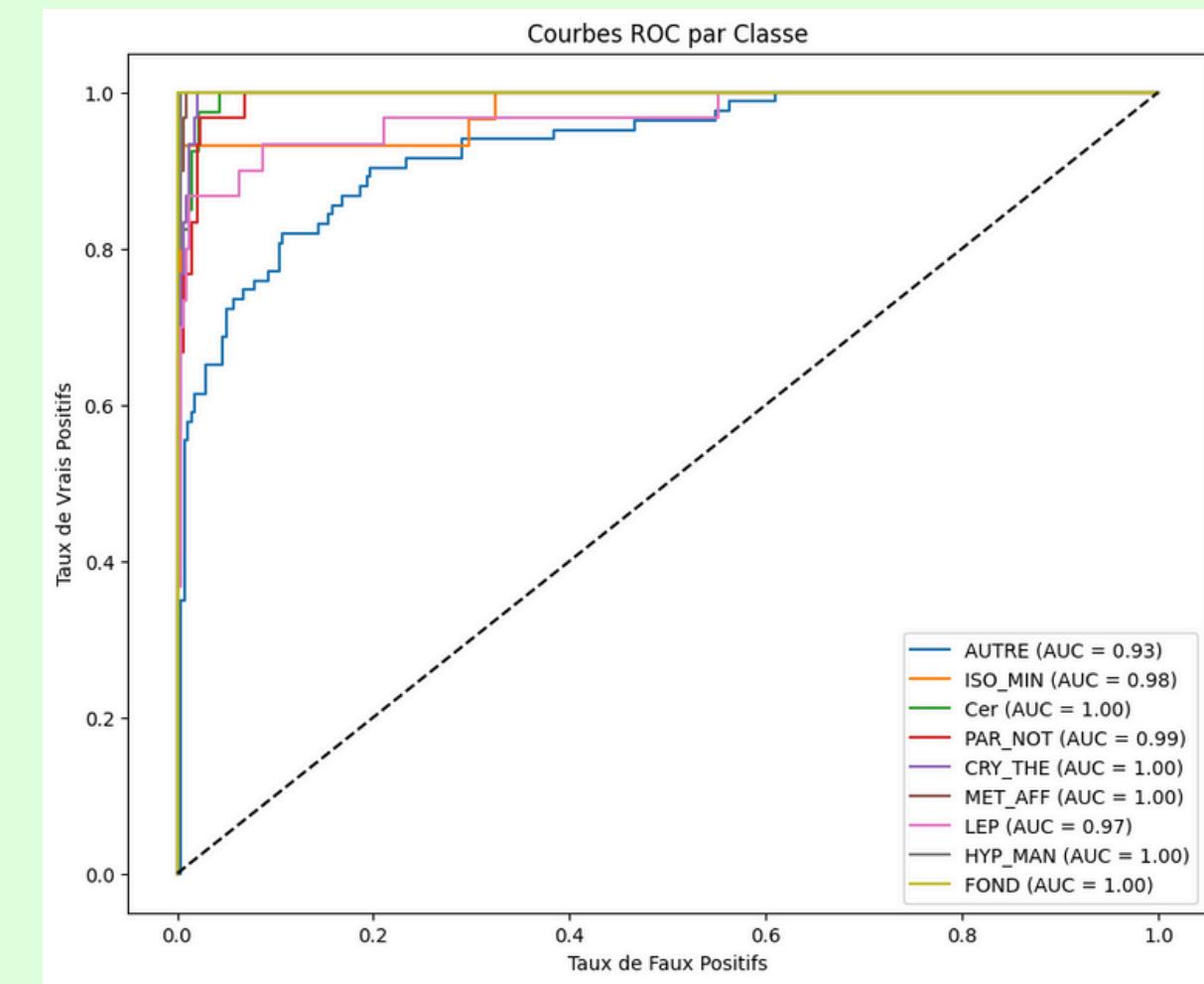
# EfficientNet-B5

## Choix d'un seuil de détection

- ◆ F1-Scores majoritairement au-dessus de 0.85, indiquant une bonne classification.
- ◆ Plusieurs classes atteignent un rappel de 1, montrant une identification précise.
- ◆ AUC proche de 1 pour de nombreuses classes, confirmant une distinction efficace.

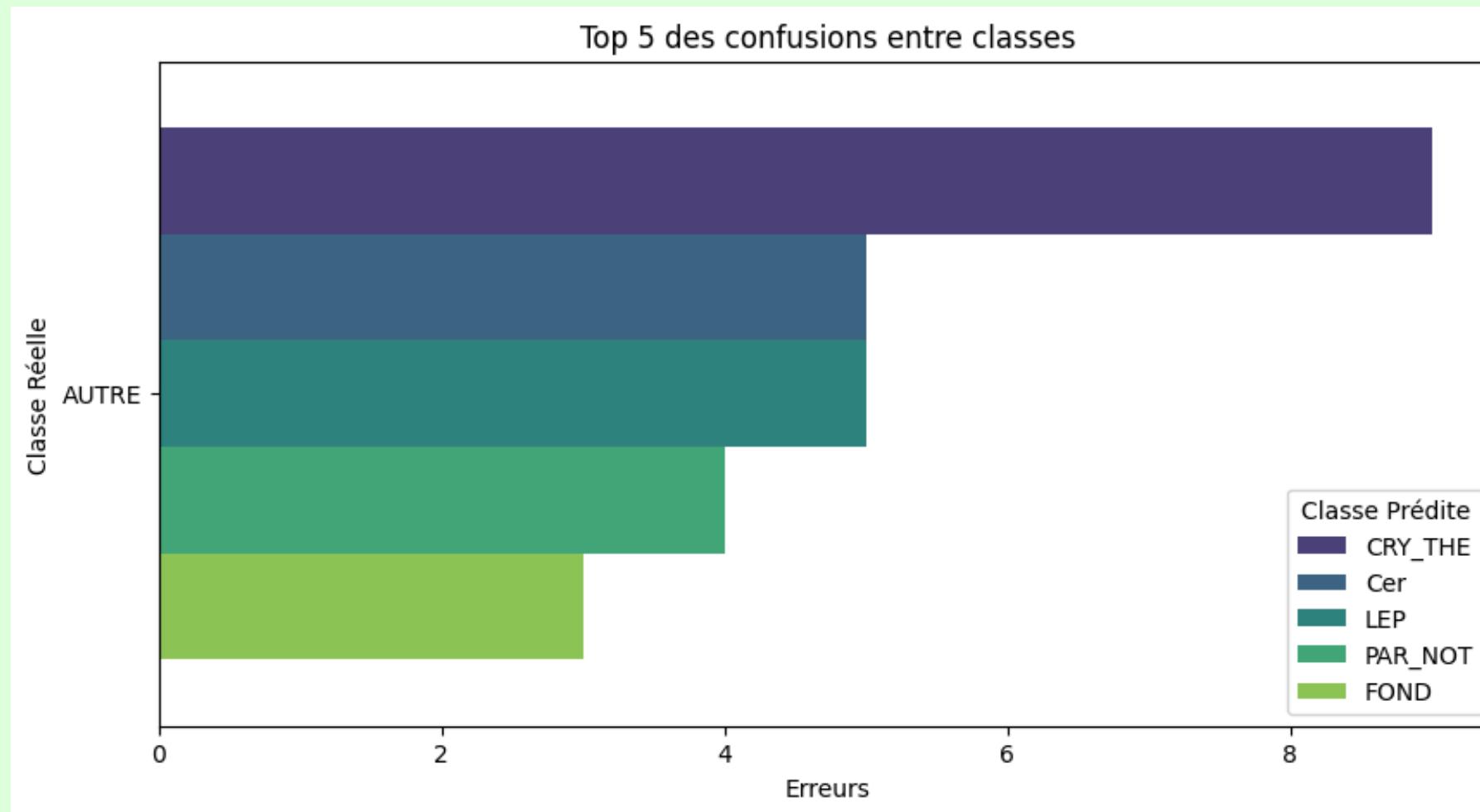
	précision	rappel	f1-score	support
AUTRE	0.87	0.64	0.74	83
ISO_MIN	0.93	0.93	0.93	29
Cer	0.86	0.93	0.89	40
PAR_NOT	0.81	0.83	0.82	30
CRY_THE	0.75	1.00	0.86	30
MET_AFF	0.94	0.97	0.95	30
LEP	0.81	0.87	0.84	30
HYP_MAN	0.94	1.00	0.97	30
FOND	0.95	1.00	0.98	60
accuracy		<b>0.88</b>		362
macro avg	0.88	0.91	0.89	362
weighted avg	0.88	0.88	0.87	362

Table 1: Performance détaillée par classe



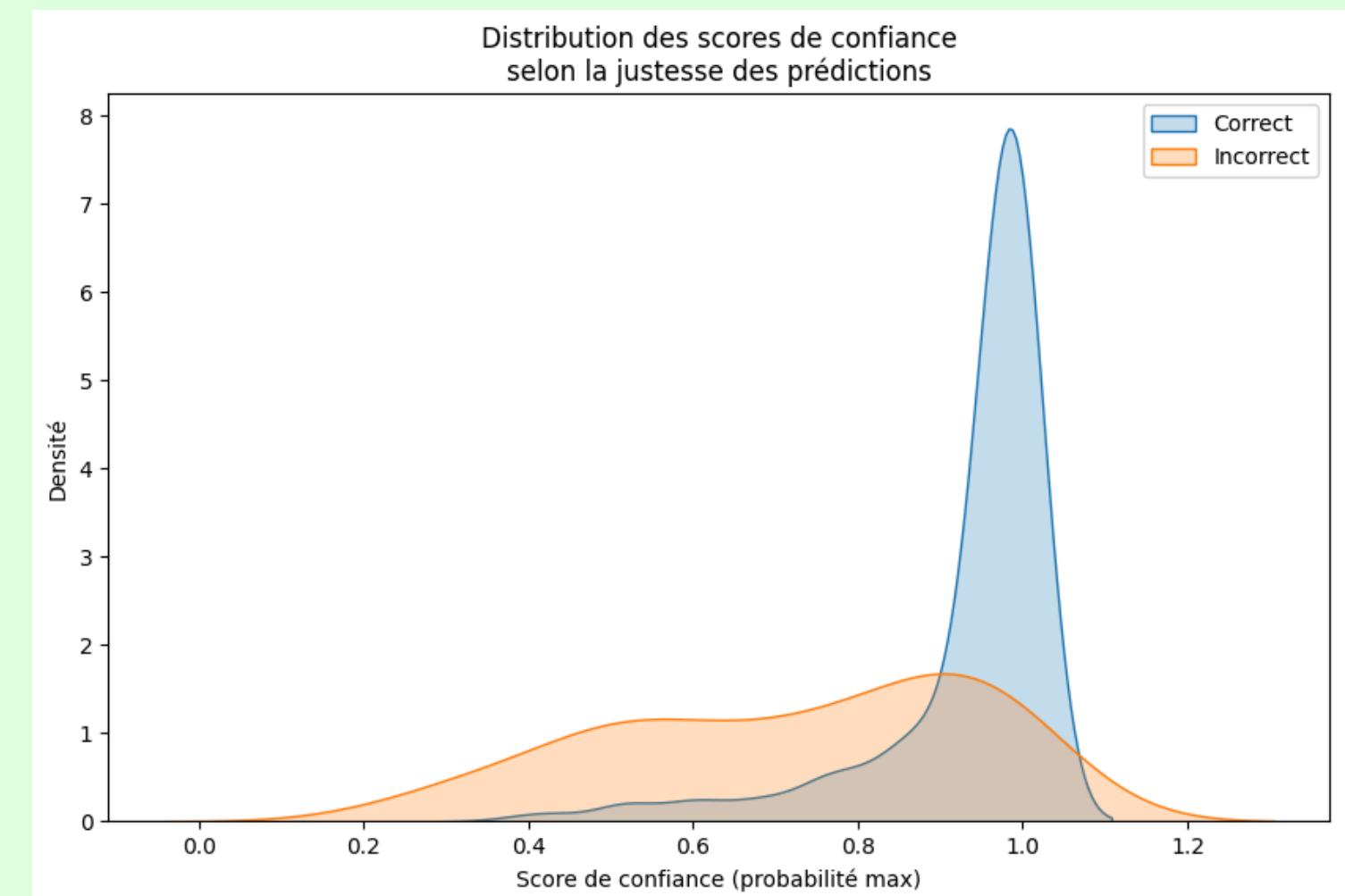
# EfficientNet-B5

## Analyse des Confusions



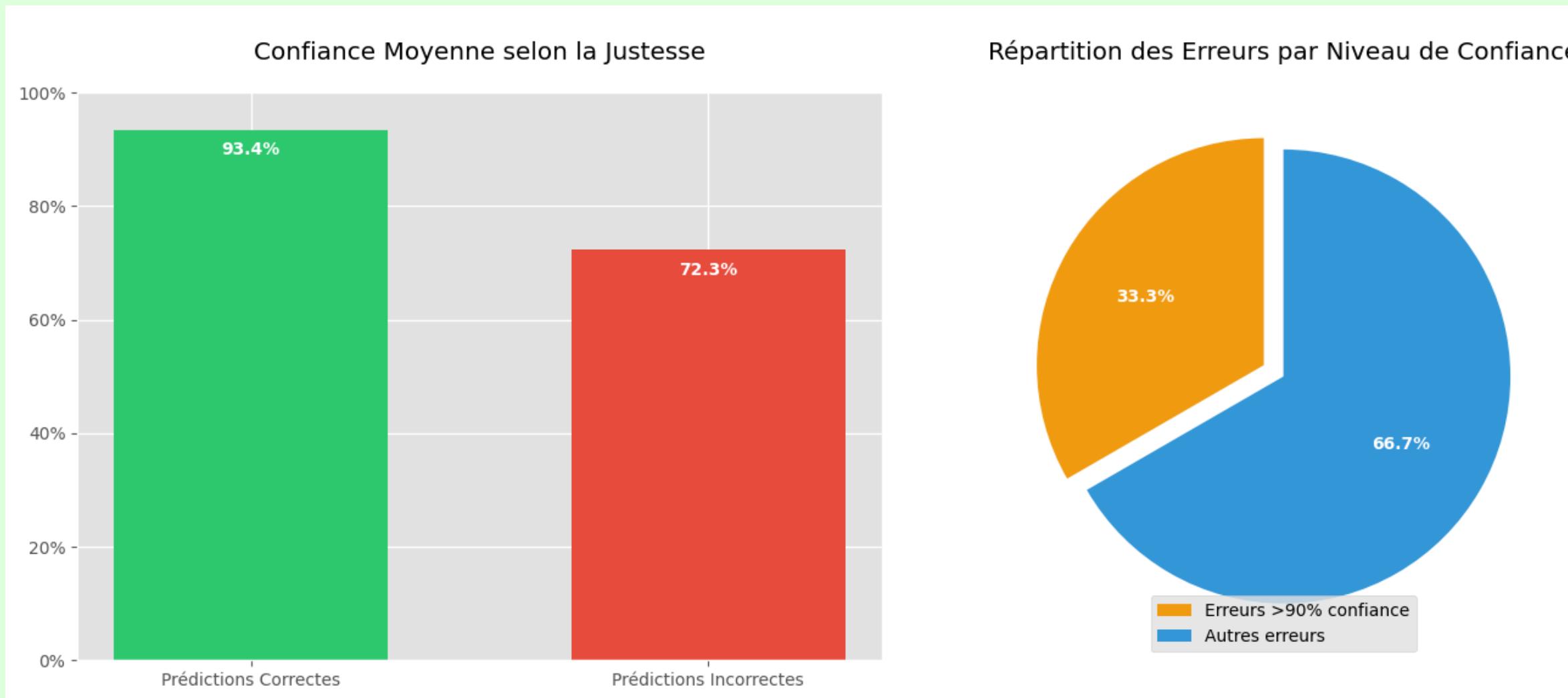
- La classe "AUTRE" est souvent confondue avec "CRY\_THE", indiquant une similarité dans les caractéristiques apprises par le modèle.

## Scores de Confiance



- Les prédictions correctes ont généralement des scores de confiance plus élevés, centrés autour de 1.
- Les prédictions incorrectes montrent une distribution plus large et des scores de confiance plus bas.

# EfficientNet-B5



- Le modèle est **21.1 % plus confiant quand il a raison**
- **1 erreur sur 3 survient avec une confiance > 90 %**
- **Paradoxe de confiance : haute confiance n'assure pas toujours la justesse !!!**

# Benchmarking

- **Performance Supérieure :** EfficientNet-B5 offre des scores de précision élevés dans la majorité des classes.
- **Architecture Avancée :** Bénéficie d'une architecture plus complexe par rapport à B3 et B4.
- **Capacités d'Apprentissage :** Meilleure capacité à capturer des caractéristiques subtiles et complexes.
- **Résultats Théoriques et Pratiques :** Les performances théoriques supérieures se reflètent dans les résultats pratiques du benchmarking.

Modèle	AUTRE	ISO_MIN	Cer	PAR_NOT	CRY_THE	MET_AFF	LEP	HYP_MAN	FOND	Moyenne
EfficientNet-B5	0.87	0.93	0.86	0.81	0.75	0.94	0.81	0.94	0.95	0.88
EfficientNet-B4	0.84	0.91	0.85	0.80	0.72	0.91	0.79	0.93	0.92	0.85
EfficientNet-B3	0.57	0.61	0.58	0.59	0.47	0.68	0.65	0.67	0.70	0.61
ViT	0.30	0.35	0.28	0.25	0.20	0.32	0.27	0.29	0.31	0.29

Table 1: Comparaison des précisions des modèles EfficientNet-B5, B4, B3 et ViT par classe.

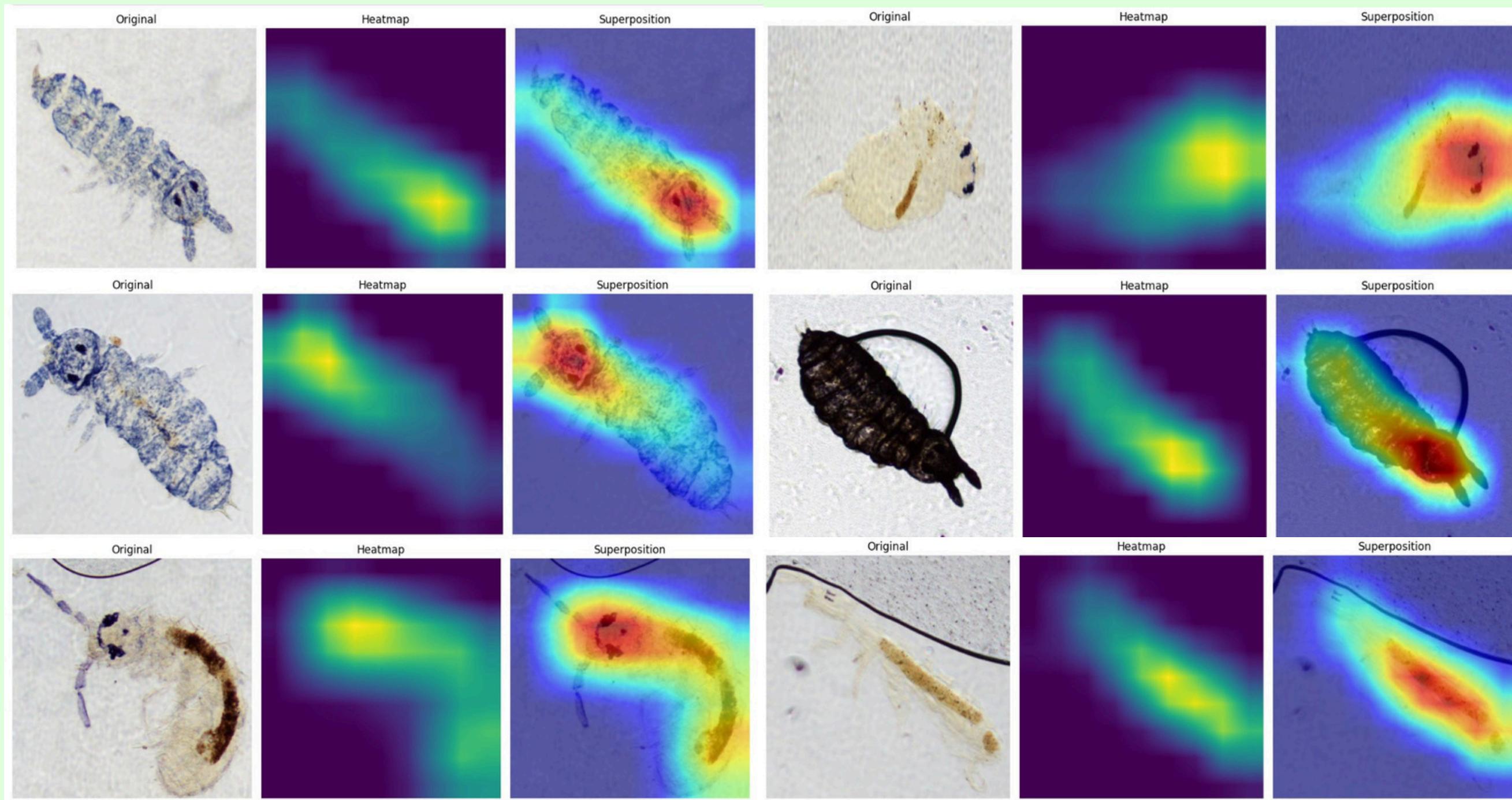


# **EXPLICABILITÉ DU MODÈLE**

# EfficientNet-B5

## Grad-CAM:

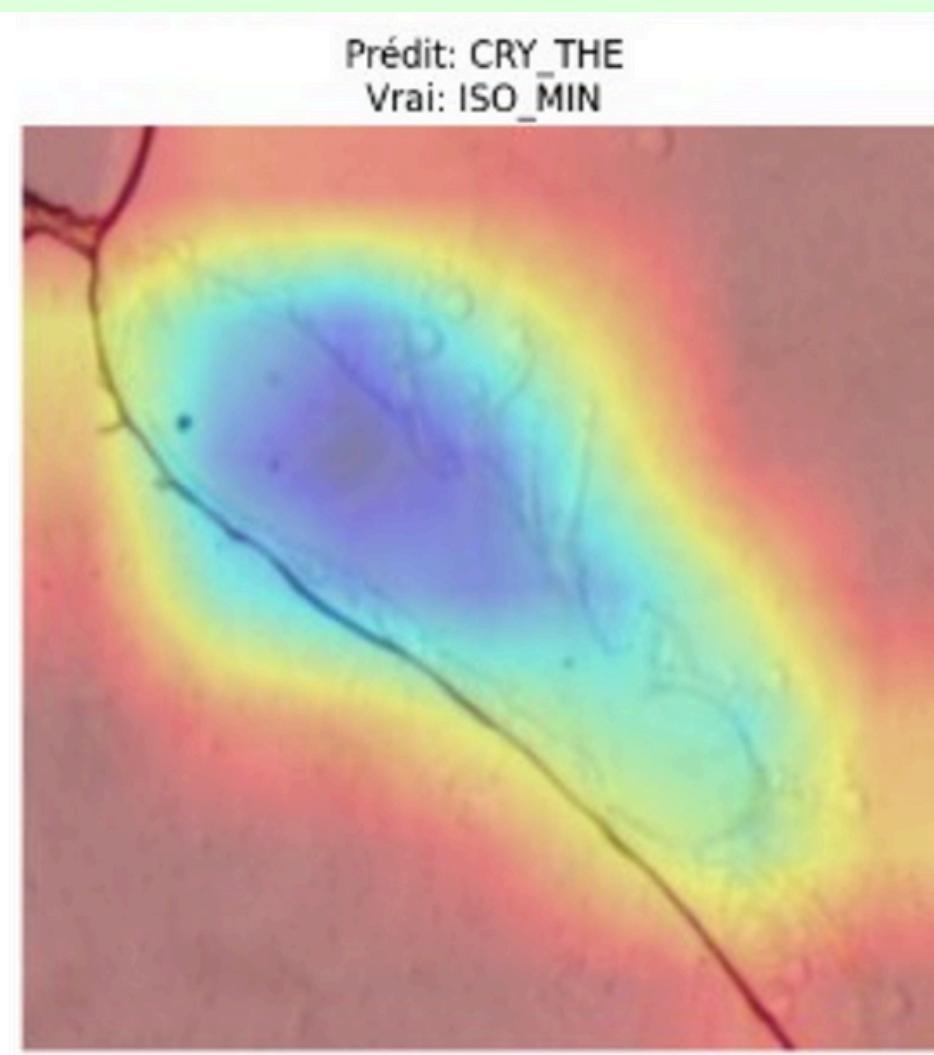
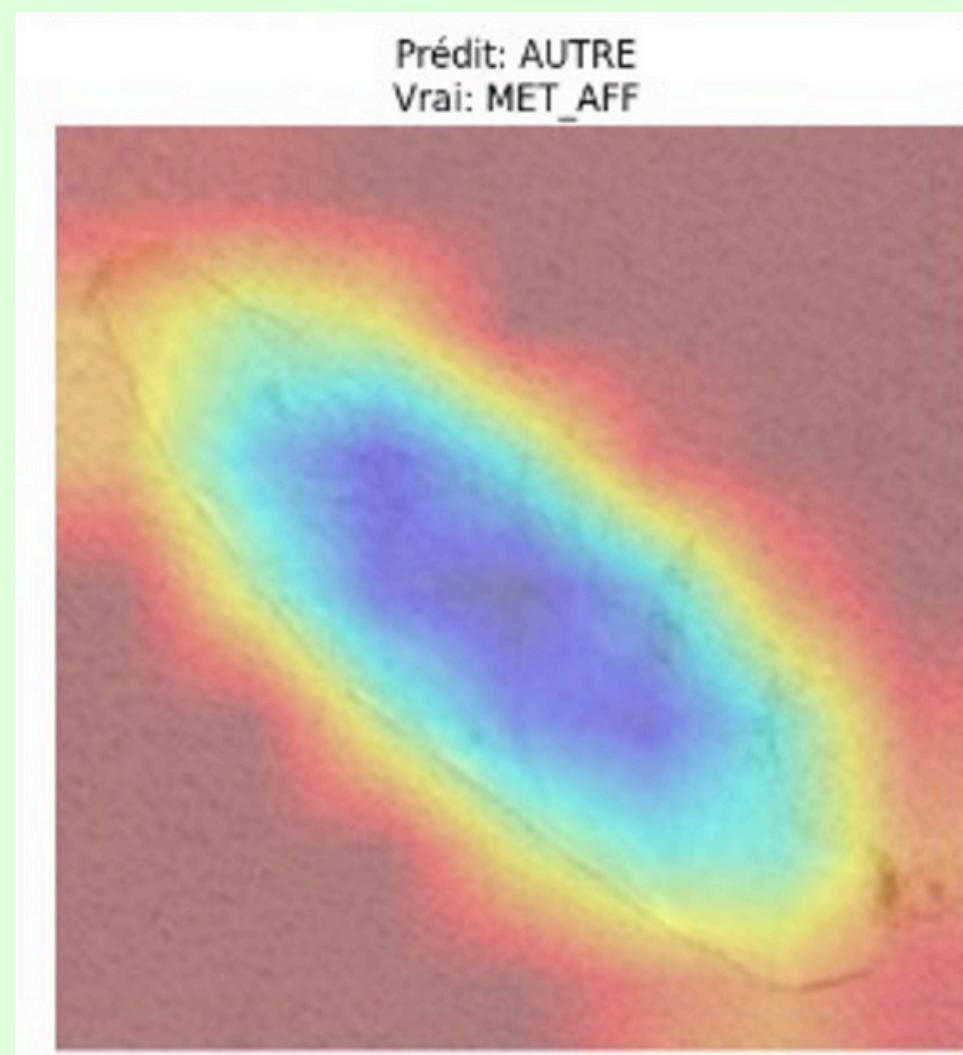
**Rouge vif = zone importante ; Jaune = influence modérée ; vert = faible ; bleu foncé = sans importance**



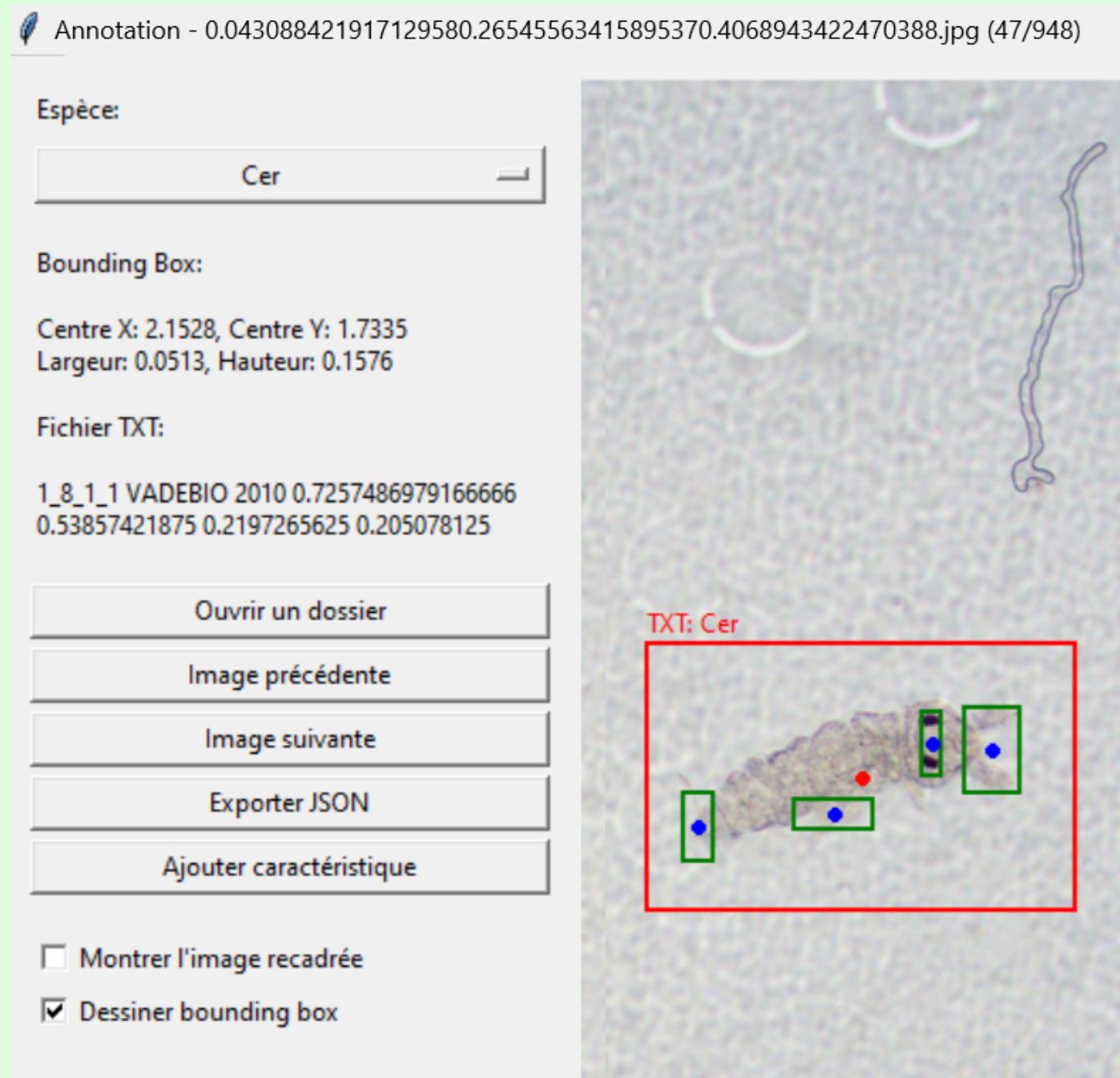
# EfficientNet-B5

## Grad-CAM:

**Rouge vif = zone importante ; Jaune = influence modérée ; vert = faible ; bleu foné = sans importance**



# L'approche par caractéristiques



# *Notre interface de notation d'images*

# L'intuition

**Les cartes de chaleur sont difficilement interprétables sur le plan statistique. Localiser les zones chaudes par rapport aux caractéristiques pourrait fournir des éléments de réponse.**

# L'annotation des caractéristiques

**Les données ont été créées à partir de 120 images qui ont été annotées sur 5 caractéristiques : les antennes, la queue, les yeux, la furca et l'appendice.**

# La généralisation aux autres images

## **Utilisation du modèle Faster R-CNN (ResNet50) adapté pour de telles tâches :**

- 1. Identification des zones (ou "propositions") où des caractéristiques pourraient se trouver.**
  - 2. Classement des propositions et ajustement des boîtes englobantes (Format YOLO+ normalisé).**

# L'approche par caractéristiques

## 2



Quelques caractéristiques prédites (sans seuil en bas à droite)

### Choix d'un seuil de détection

Pour améliorer la fiabilité de la détection, un seuil minimum de score de confiance à 0,5 est fixé, et au-dessus la valeur maximale est sélectionnée pour chaque classe le cas échéant.

### Performances de détection

Sur 1117 images, 873 présentent des antennes détectées, 302 affichent des yeux détectés, et 51 montrent la présence de la furca détectée.

- Peu d'erreurs grâce au seuil du score de confiance.
- Mais des caractéristiques non reconnues, la diversité des annotation étant limitée et non uniforme. *En particulier pour les toutes petites images.*

### Rapprochement avec les cartes de chaleur

Une fois ces étapes de prétraitement passées, la comparaison et l'élaboration de statistiques détaillées est possible.



# PERSPECTIVES D'AMÉLIORATIONS

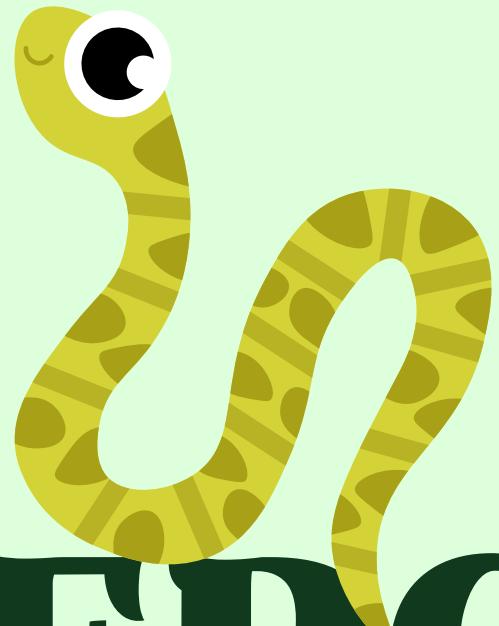
# Perspectives d'améliorations

## Analyse des performances :

- Comprendre pourquoi ViT a sous-performé par rapport à EfficientNet-B5 (taille des données ? pré-entraînement ? complexité ?).

## Exploration de modèles hybrides :

- Tester des architectures combinant CNN et Transformers (ex. ConvNeXt, Swin Transformer) pour optimiser les performances.



**MERCI**

