



# Estimation automatisée sur vidéos de la biodiversité des poissons coralliens

Sébastien VILLON

Directeurs :

Marc CHAUMONT

David MOUILLOT

Encadrants :

Thomas CLAVERIE

Gérard SUBSOL

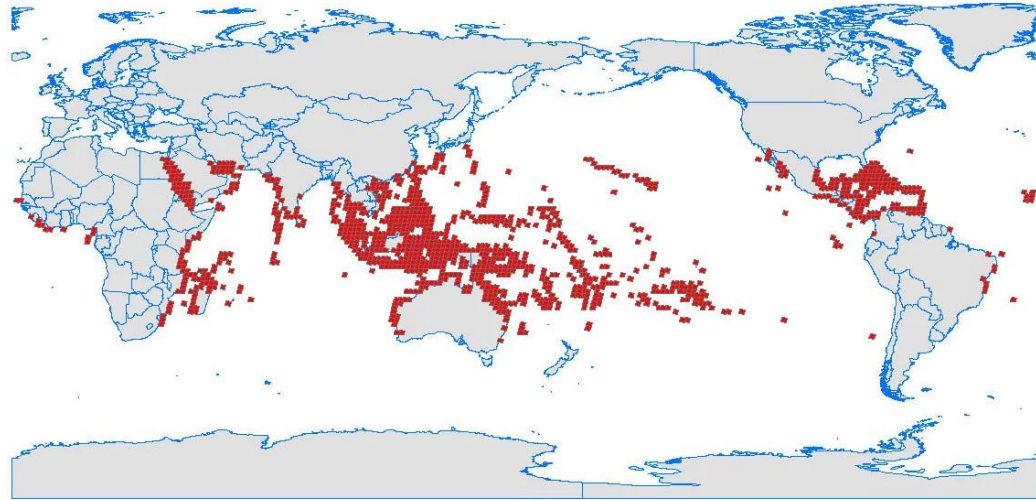
Sébastien VILLÉGER

Emily S. DARLING



# Les récifs coralliens

0,1% de la surface océanique



25% des espèces de poissons marins

# Pressions affectant les communautés de poissons

## Pressions locales



*Photo : E. Maire*



## Pressions globales

*Photo : The Ocean Agency / XL Catlin Seaview Survey / R. Vevers*

# Intérêt des communautés de poissons coralliens

## Rôles et Services des poissons coralliens

Rôle écologique  
(herbivorie, bio-érosion,  
recyclage des nutriments)

Rôle économique  
(Tourisme, pêche)

Source de vitamines,  
minéraux, protéines

*Brandl, S. J., et coll. "The hidden half: ecology and evolution of cryptobenthic fishes on coral reefs." Biological Reviews (2018)*

*Harris, D. L., et coll. "Coral reef structural complexity provides important coastal protection from waves under rising sea levels." Science advances (2018)*

*Moberg, F., and Carl F. "Ecological goods and services of coral reef ecosystems." Ecological economics (1999)*

**Menaces** importantes sur les poissons coralliens

Mesurer la biodiversité des poissons à **haute fréquence temporelle** et **large échelle spatiale**

→ **Recenser** les poissons



# Recensement d'une communauté de poissons

**Détecter et identifier**  
tous les individus



Deux tâches complexes

## Comptage visuel en plongée



Photos : B. Holbrook



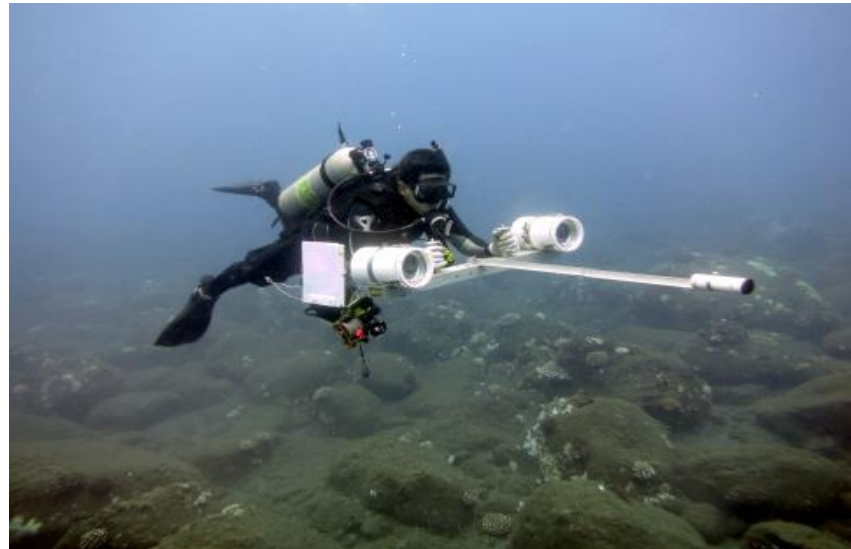
Photo : Fisheries and Oceans Canada

- **Expertise** taxonomique rare
- **Limitations** physiologiques (et légales!)
- **Faible rendement** (20-50 min = 500 m<sup>2</sup>)

→ **Impossible à grande échelle**

# Méthodes actuelles de recensement

Vers des systèmes  
d'acquisition vidéo



*Photo : A. Gray*



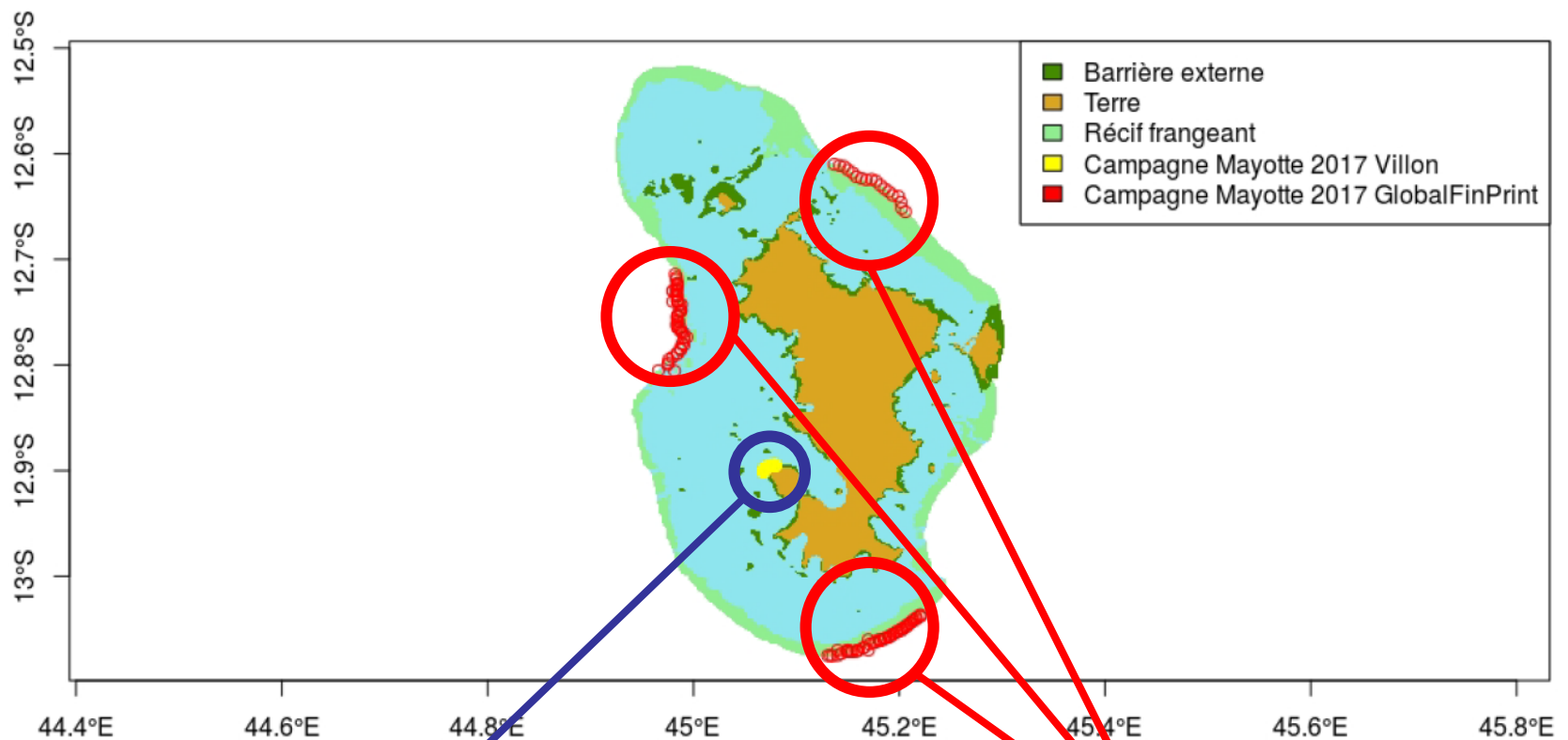
*Crédit : LIRMM*



*Photo : J. Chalifour*



# Méthodes actuelles de recensement

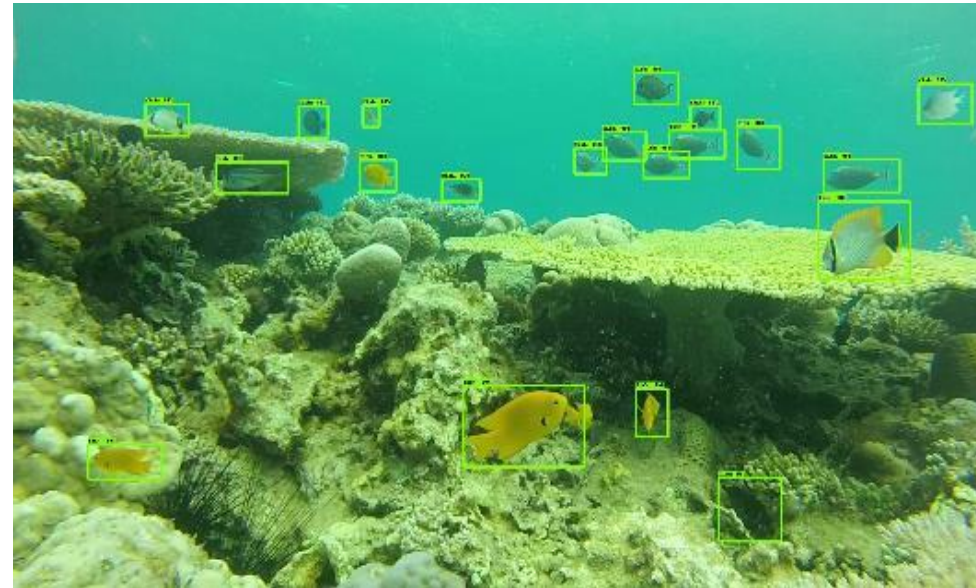


1 site  
3 jours  
90h de vidéos

3 sites  
5 jours  
500h de vidéos

# Traitement automatique de vidéos

Recenser = Localiser tous les individus et les **identifier** dans une image



*Pomacentrus sulfureus*



*Chaetodon trifascialis*



*Naso elegans*

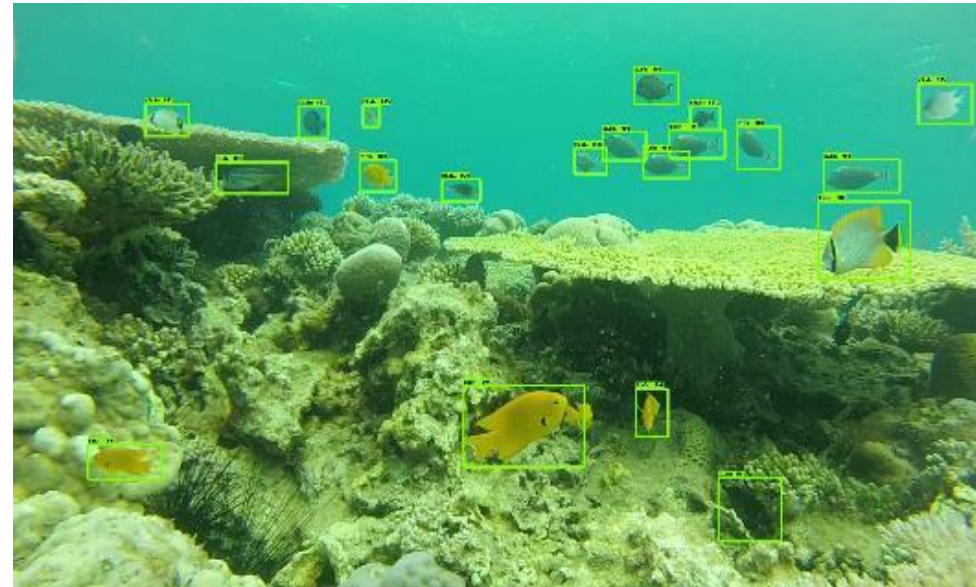
- **Quelle méthode choisir pour le recensement automatique?**

Comment améliorer la base d'entraînement?

- Le traitement automatique est-il aussi efficace que l'humain pour l'identification d'espèces de poissons?
- Comment gérer les mauvaises classifications du traitement automatique
- Vers un outil opérationnel de recensement automatique?

# Identification de poissons dans des images

## Localisation manuelle



## Identification automatique



Algorithmme

*[Pomacentrus sulfureus]*



# Identification automatique d'images

Phase d'apprentissage

Base d'apprentissage:



*Chaetodon trifasciatus*



*Zebrasoma scopas*

**Algorithme**  
Optimisation  
des paramètres

Définition des classes :

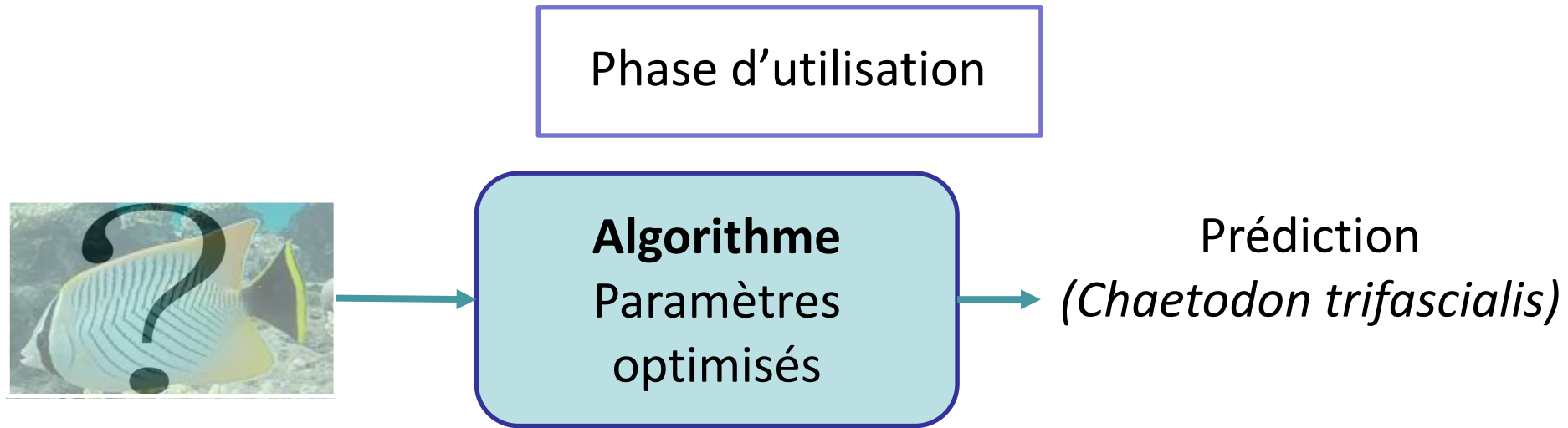
*Pomacentrus sulfureus*

*Chaetodon trifasciatus*

*Zebrasoma scopas*

.  
. .  
.

# Identification automatique d'images



# Identification automatique d'images

## Machine Learning



Algorithme d'extraction  
de caractéristiques  
prédéfinies

Algorithme  
de  
Classification

Résultats

## Deep Learning

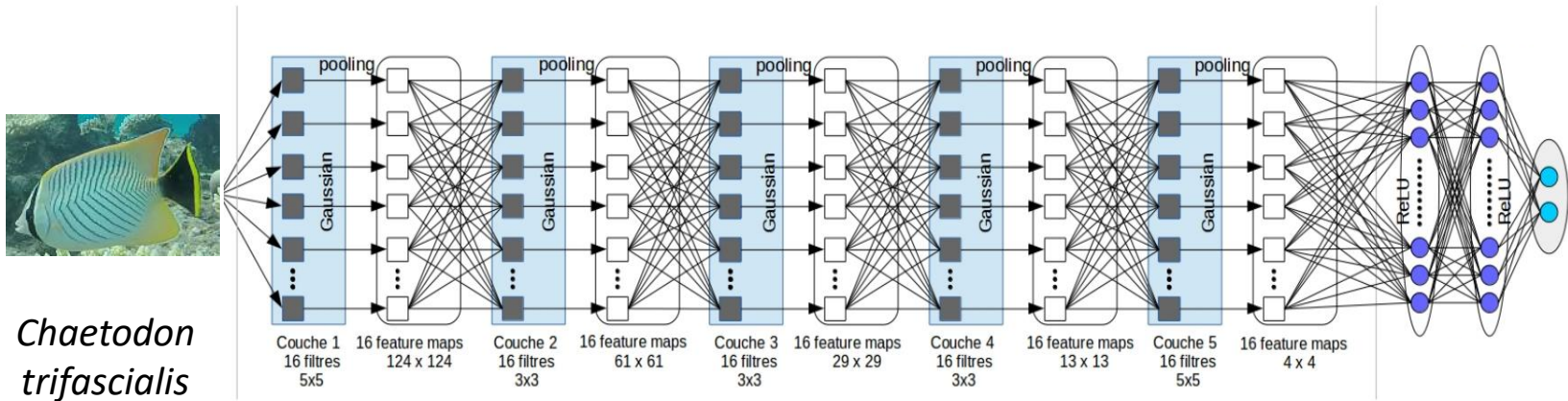


Algorithme  
(extraction + classification)

Résultats

# Identification automatique d'images

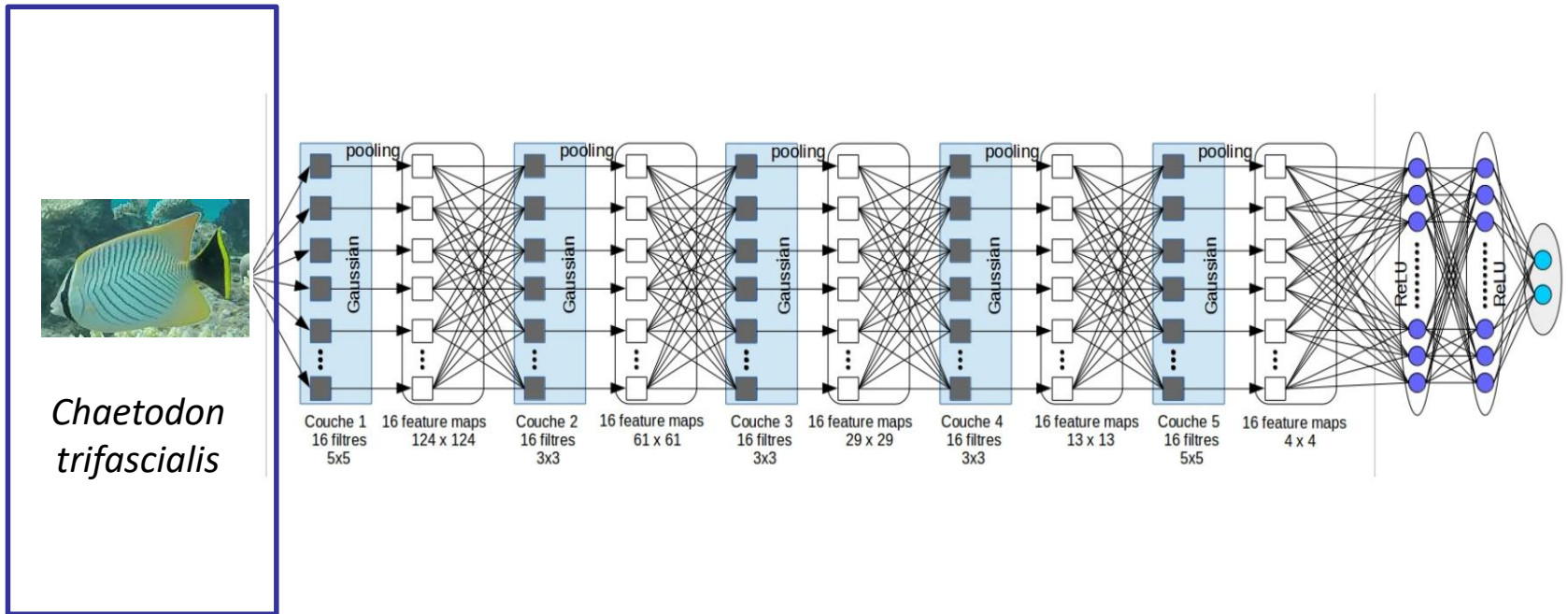
## Deep Learning en traitement d'images : Les réseaux de neurones convolutifs (CNN)





# Identification automatique d'images

## Deep Learning en traitement d'images



*Chaetodon trifascialis*

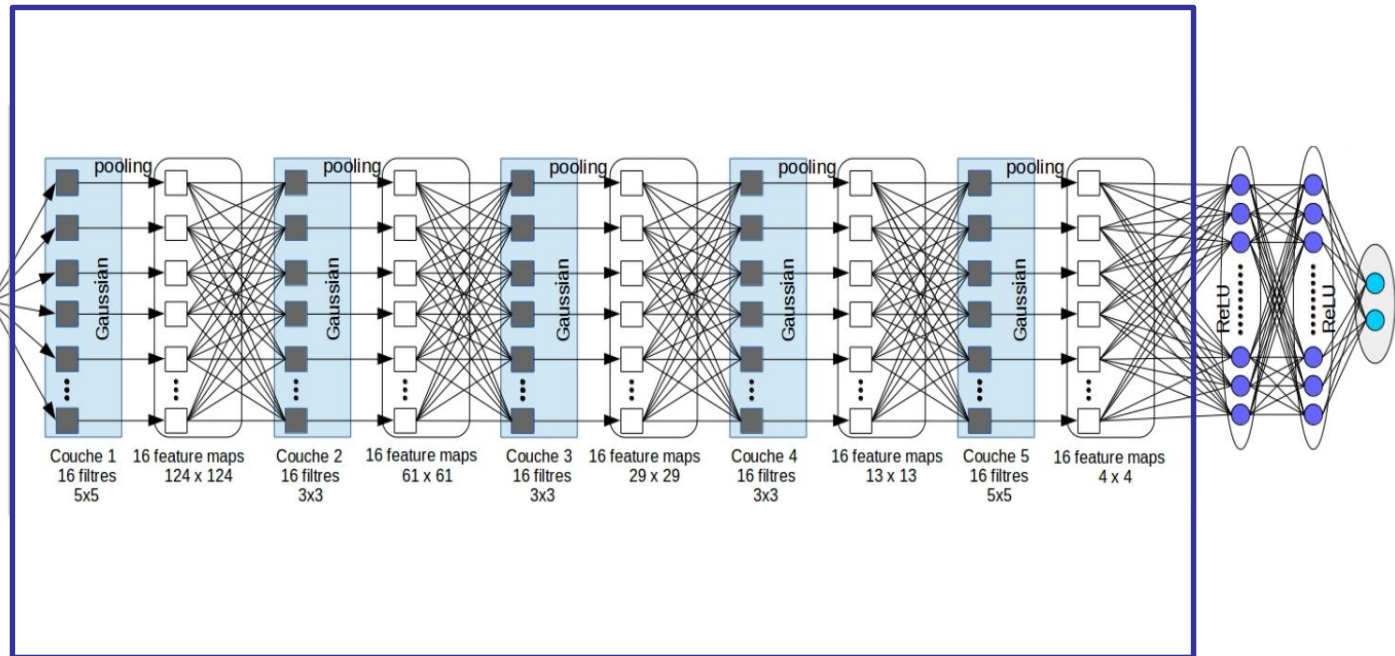
Données d'entrée

# Identification automatique d'images

## Deep Learning en traitement d'images



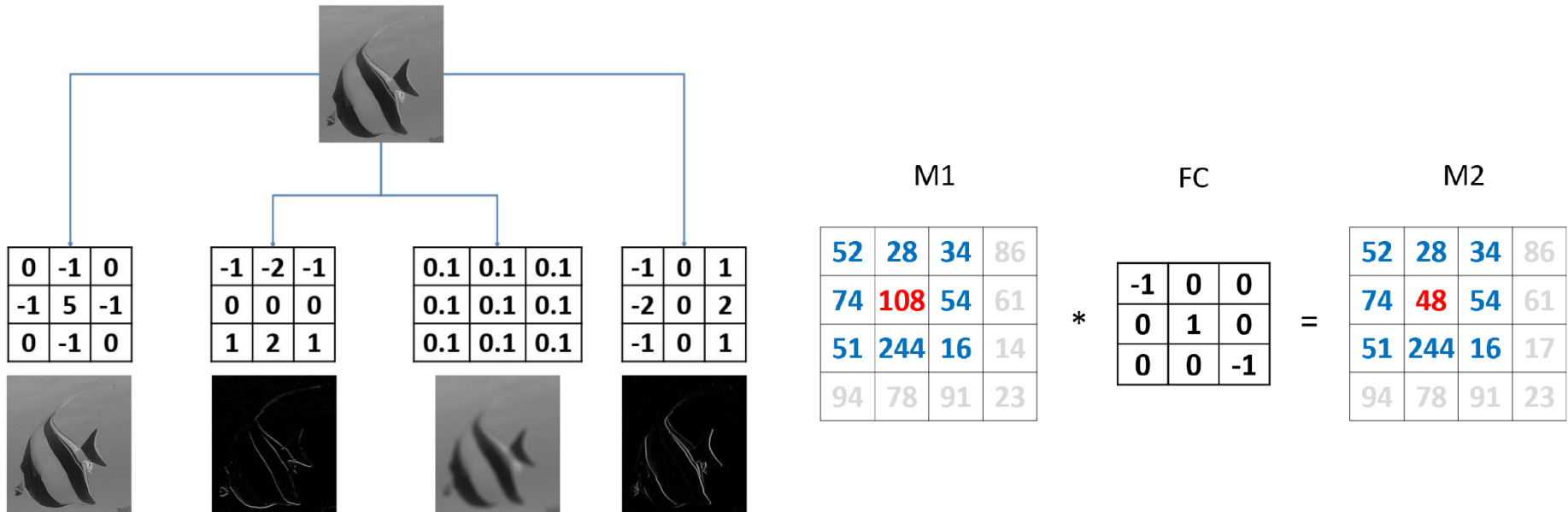
*Chaetodon trifascialis*



Extraction automatique de caractéristiques

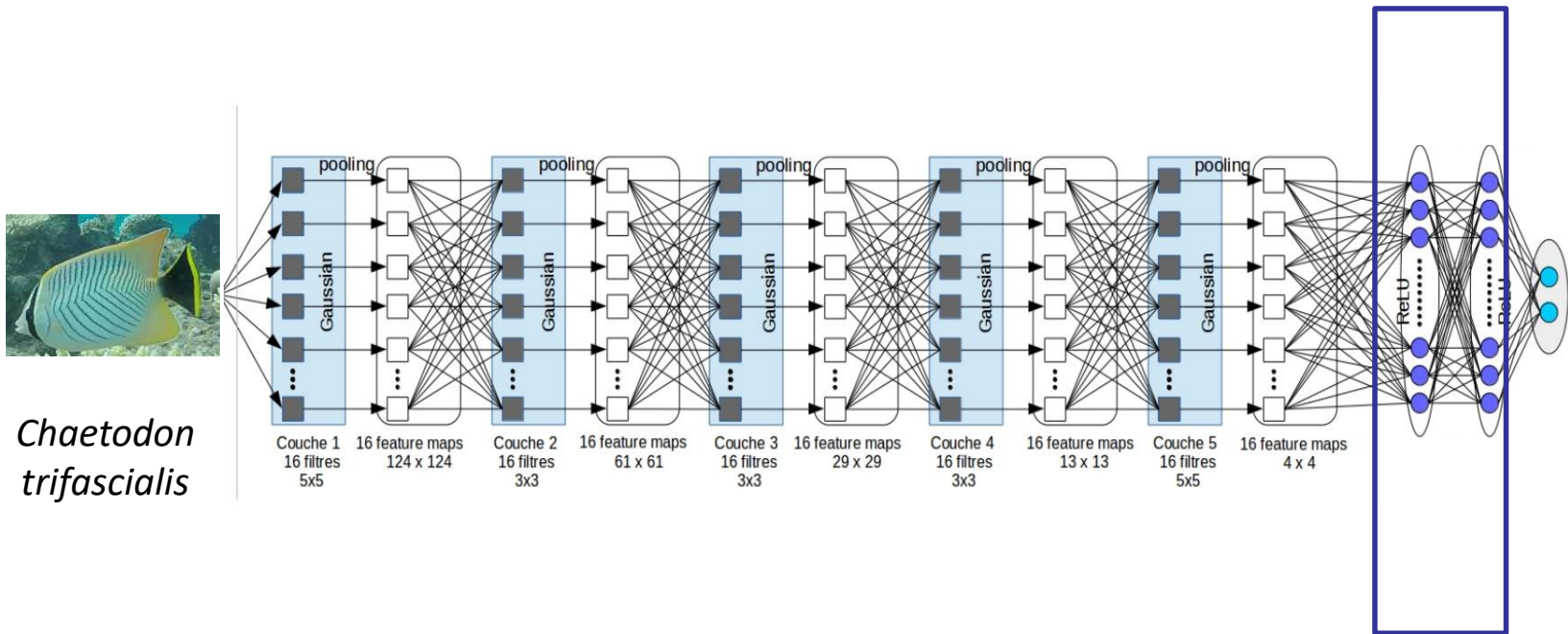
# Identification automatique d'images

## Deep Learning en traitement d'images : La convolution



# Identification automatique d'images

## Deep Learning en traitement d'images

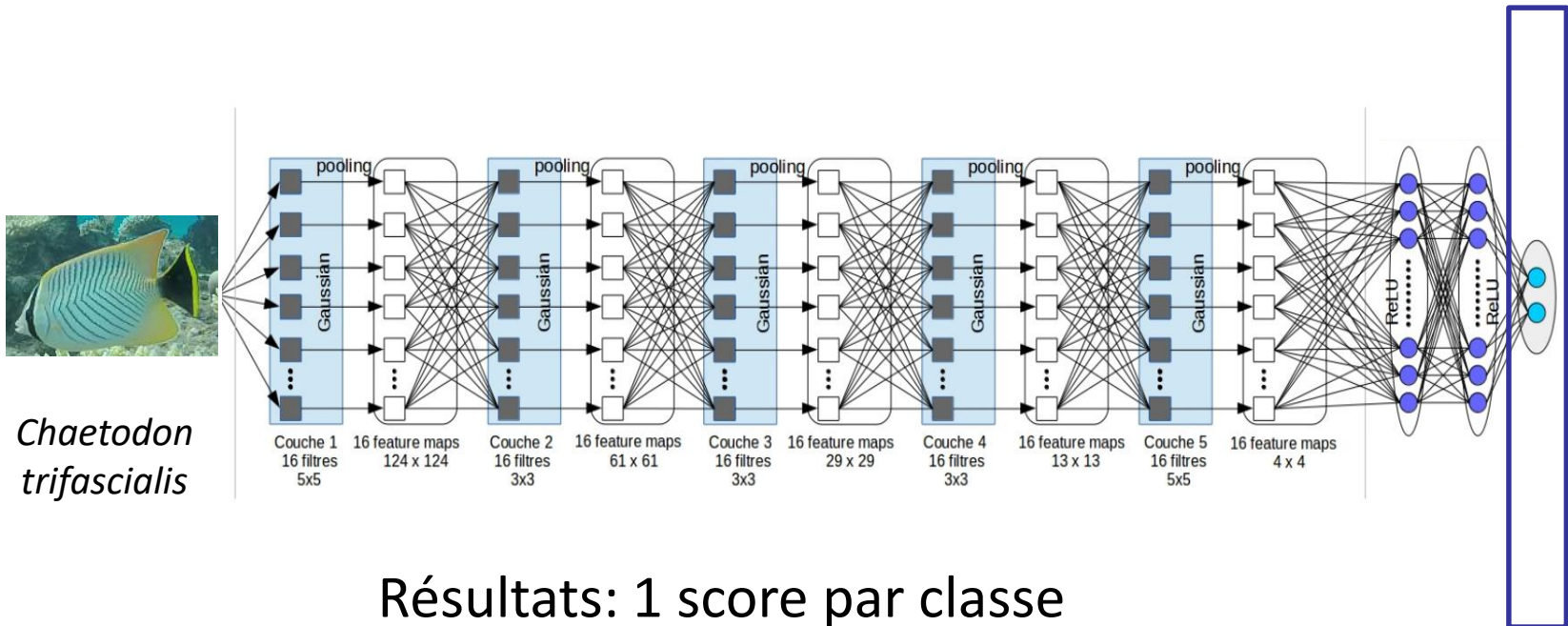


Classification



# Identification automatique d'images

## Deep Learning en traitement d'images



Résultats: 1 score par classe

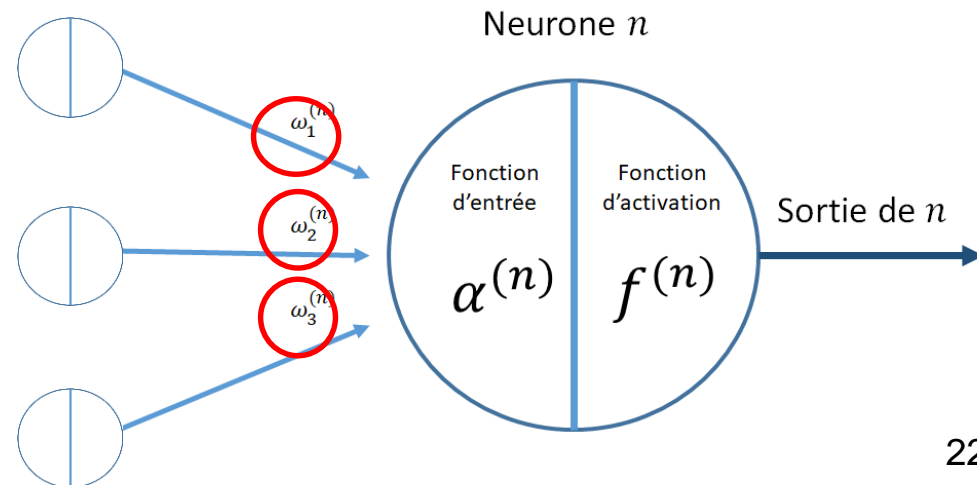
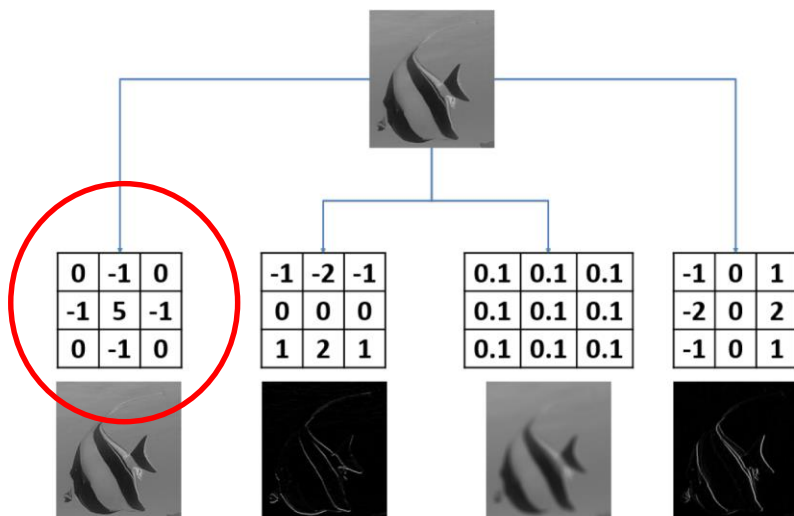
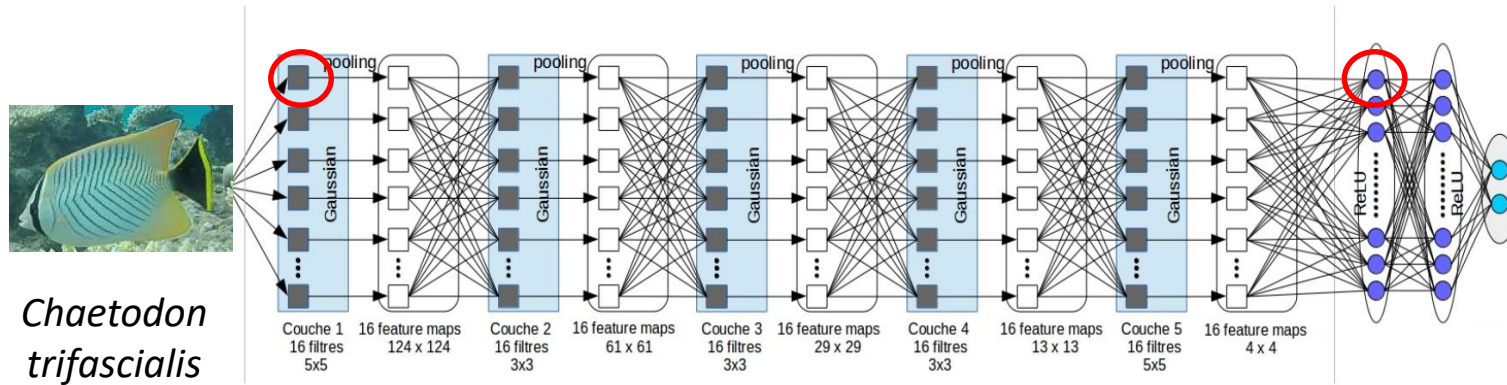
*Chaetodon trifascialis* : 0.91

*Pomacentrus sulfureus* : 0.08

# Identification automatique d'images

## Optimisation par rétro-propagation

*Chaetodon trifascialis* : 0.91



## Enregistrements vidéos

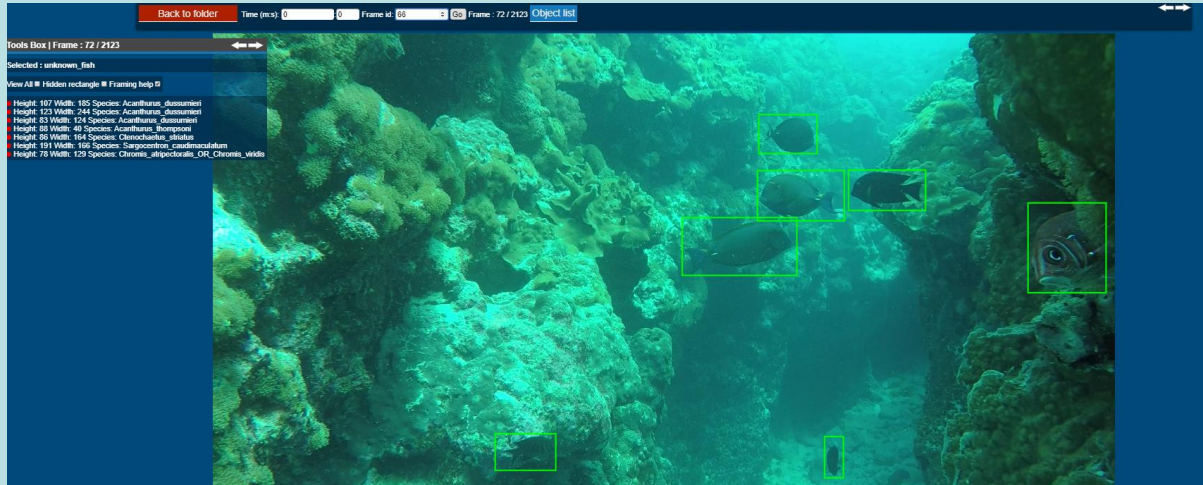
**500 heures de vidéos** réalisées entre 2015 et 2019

Vidéos **Full HD** (1920x180 pixels)

**100 sites différents** (Océan Ouest-Indien, Méditerranée, Mer Rouge, Caraïbes, Océan Pacifique)

Enregistrements vidéos (+ de 500 heures)

Développement d'une application d'annotation



C. Desgenetez (Stage DUT)



Stages  
volontaires

collabor

, suivie des meta-données

360 000 imageries

50+ volontaires



# Comparaison Machine Learning et Deep Learning



**Machine Learning**

Prédiction  
*(Chaetodon trifascialis)*



**Deep Learning**

Prédiction  
*(Chaetodon trifascialis)*

**Deux fois plus de classifications correctes**



- **Quelle méthode choisir pour le recensement automatique?**

## **Comment améliorer la base d'entraînement?**

- Le traitement automatique est-il aussi efficace que l'humain pour l'identification d'espèces de poissons?
- Comment gérer les mauvaises classifications du traitement automatique
- Peut-on envisager un outil opérationnel de recensement automatique?

# Base de données et Deep Learning

Création du modèle profond d'identification

Première base  
d'entraînement :  
20 classes  
espèces



# Base de données et Deep Learning

Création du modèle profond d'identification

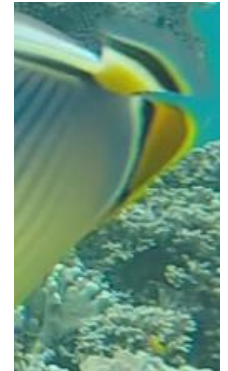
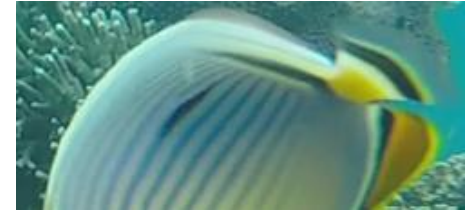
Deuxième base  
d'entraînement :  
Base 1  
+ 1 classe  
Fond Marin



# Base de données et Deep Learning

Création du modèle profond d'identification

Troisième base  
d'entraînement :  
Base 2  
+20 classes parties de  
poissons



## Métriques de performances

Vérité terrain

Prédictions



*Zanclus  
cornutus*

Vrai positif (VP)



*Chromis  
viridis*

Faux négatif (FN)



*Zanclus  
cornutus*

Faux positif (FP)



## Métriques de performances

$$\text{Rappel (A)} = \frac{VP_A}{VP_A + FN_A}$$

→ 1 quand aucun faux négatifs


$$\text{Précision (A)} = \frac{VP_A}{VP_A + FP_A}$$

→ 1 quand aucun faux positifs

$$\text{F-Mesure (A)} = 2 \cdot \frac{\text{Rappel}(A) \cdot \text{Précision}(A)}{\text{Rappel}(A) + \text{Précision}(A)}$$

# Base de données et Deep Learning

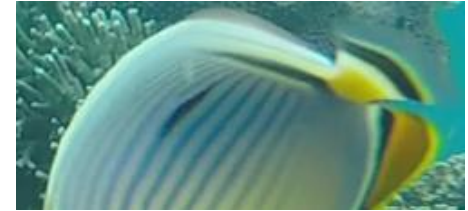
## Création du modèle profond d'identification



Meilleure:  
20 classes espèces,  
20 classes parties de  
poissons,  
1 classe fond marin



94% de Vrai Positifs



- Quelle méthode choisir pour le recensement automatique?
- **Le traitement automatique est-il aussi efficace que l'humain pour l'identification d'espèces de poissons?**

# Expertise humaine et Deep Learning

Le Deep Learning est-il aussi efficace que l'humain pour l'identification d'images de poissons?



Classification par  
des humains

Résultat de  
Classification



Algorithme  
*GoogLeNet*

Résultat de  
Classification

## Protocole pour l'identification humaine



14 personnes

20 minutes



Planchette  
d'identification  
disponible

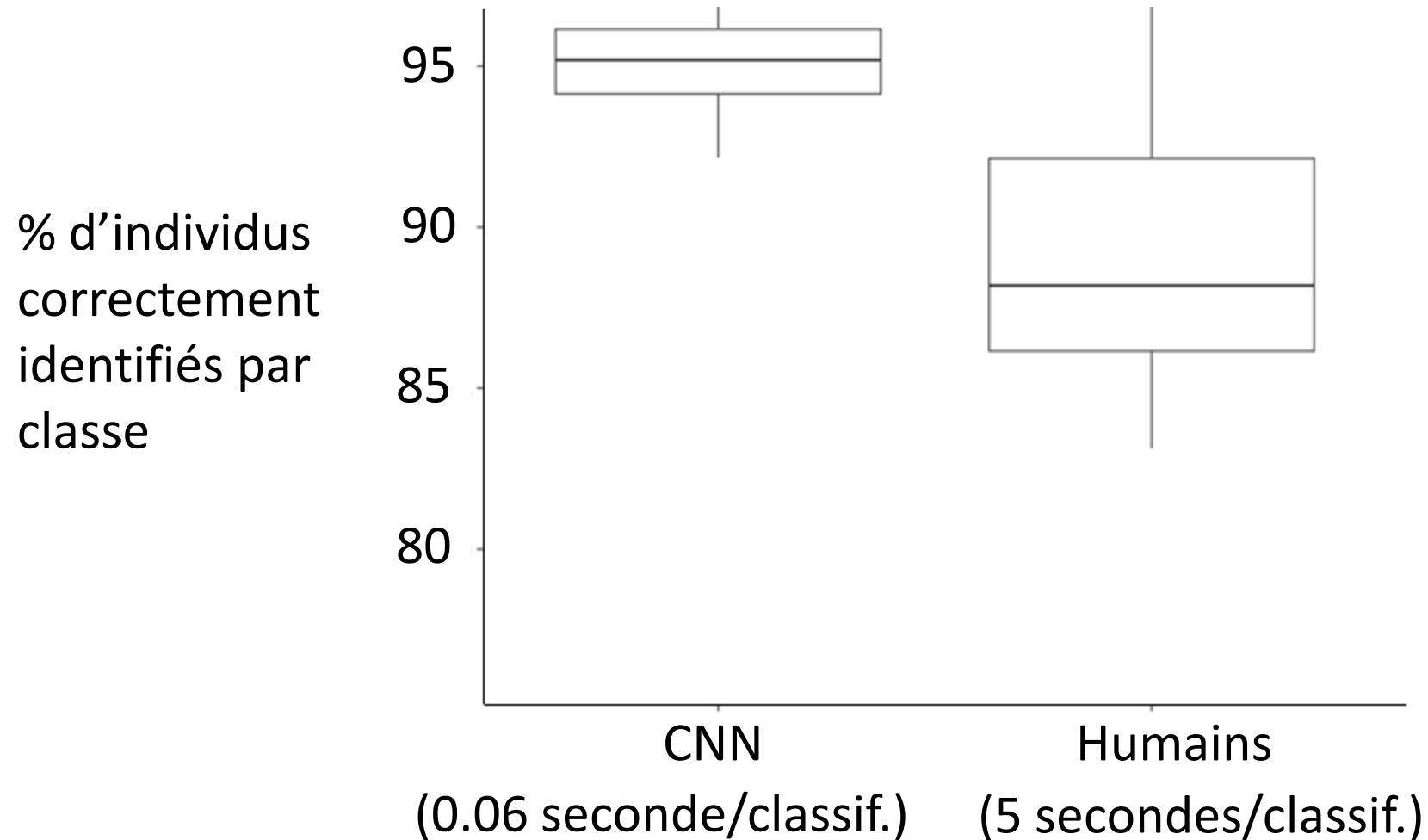


9 espèces



# Expertise humaine et Deep Learning

Comparaison CNN et Humain : **6% plus précis** et **100 fois plus rapide**.



## Résultats

Corrects uniquement pour le CNN



Corrects uniquement pour l'humain

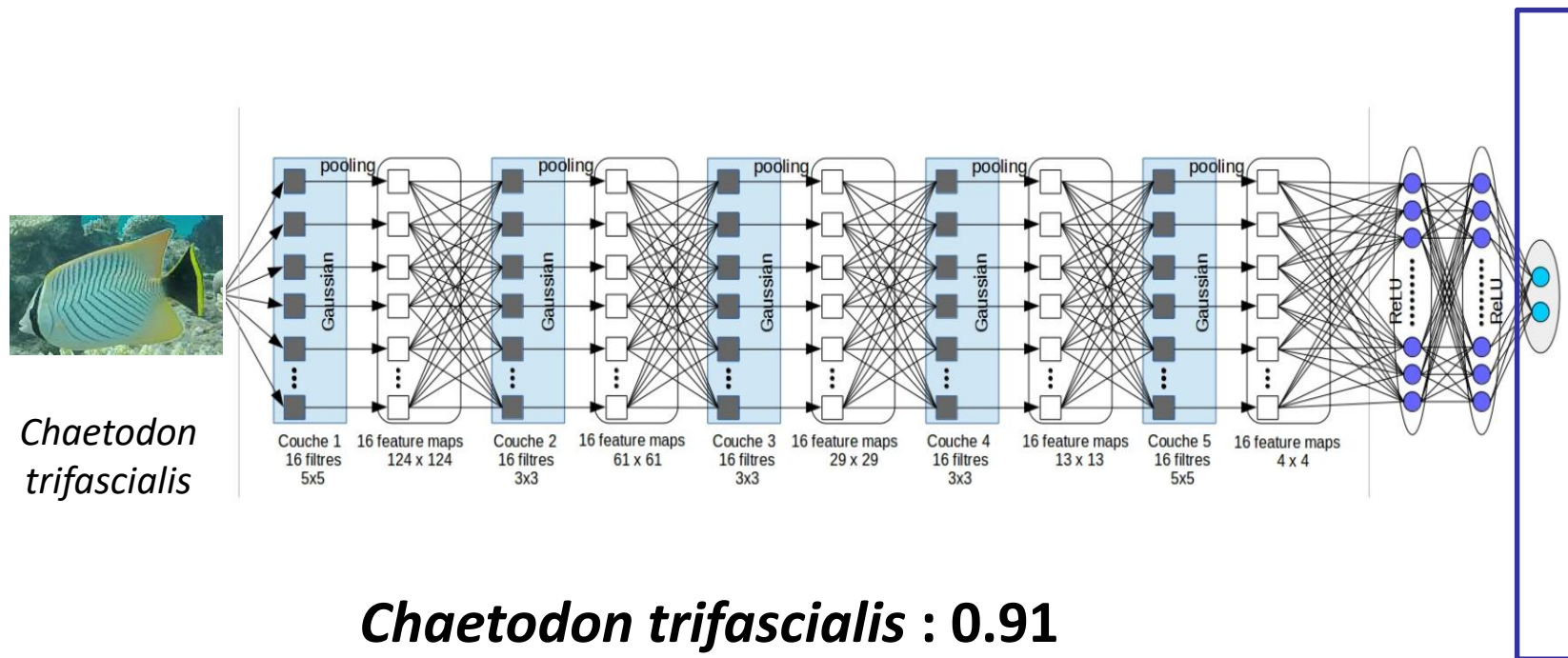


Incorrects pour les deux approches



- Quelle méthode choisir pour le recensement automatique?
- Le traitement automatique est-il aussi efficace que l'humain pour l'identification d'espèces de poissons?
- **Comment gérer les mauvaises classifications du traitement automatique**

# Gestion des erreurs des CNNs



***Chaetodon trifascialis* : 0.91**

*Pomacentrus sulfureus* : 0.08

- 1) Les scores ne sont PAS des probabilités
- 2) La majorité des scores Top 1 retournés par les CNNs sont élevés (Juste ou Faux)

# Gestion des erreurs des CNNs

## Métriques de performances

Vérité terrain

Prédictions



*Zanclus  
cornutus*

Vrai positif (VP)



*Chromis  
viridis*

Faux négatif (FN)

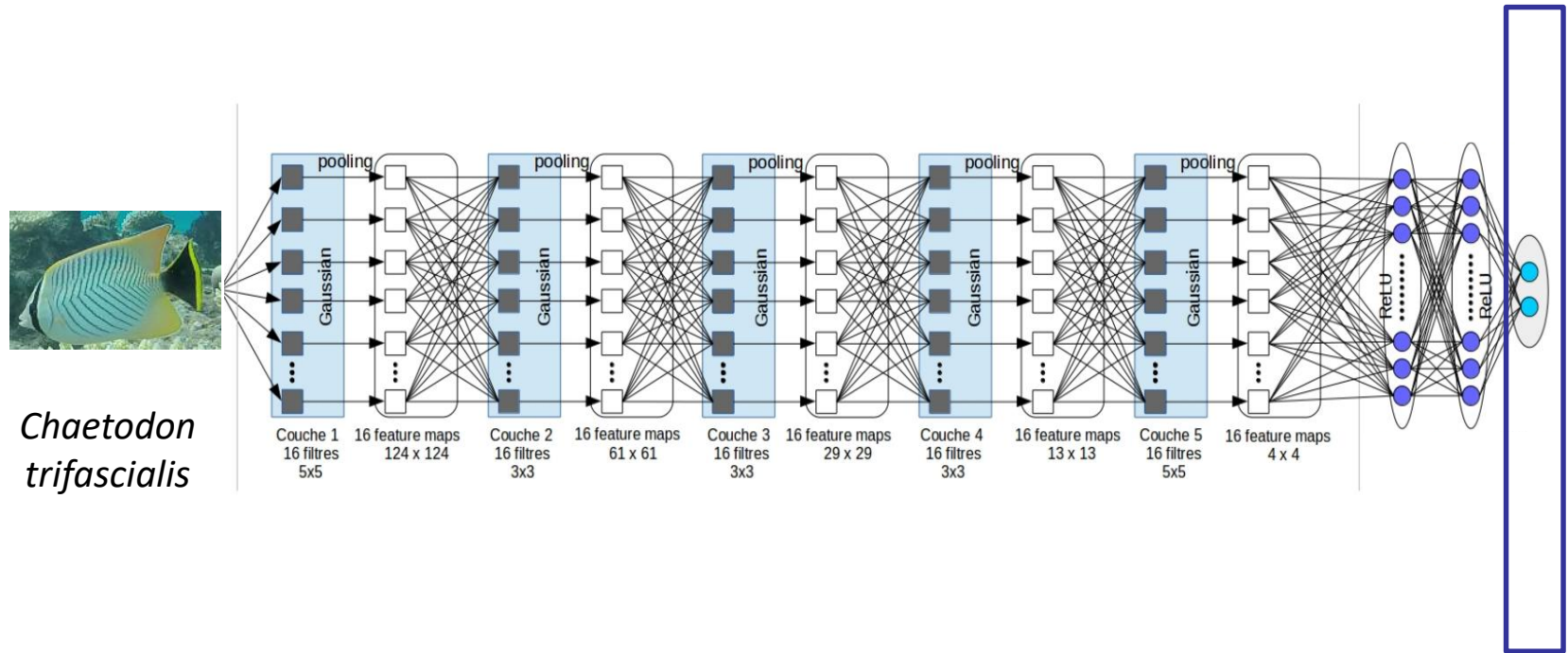


*Zanclus  
cornutus*

Faux positif (FP)



# Gestion des erreurs des CNNs



*Chaetodon trifascialis*

*Chaetodon trifascialis*  
*Pomacentrus sulfureus*

- 
- 
- 

« Je ne sais pas »

Ajout d'une classe  
« Unsure »

Deux méthodes principalement utilisées :

La calibration : transformer le score en probabilité



Problème de précision, problématique différente

*Platt, J. "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods." Advances in large margin classifiers 10.3 (1999).*

*Zadrozny, B., and Charles E. "Transforming classifier scores into accurate multiclass probability estimates." Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining. (2002).*

*Guo, C., et coll. "On calibration of modern neural networks." Proceedings of the 34th International Conference on Machine Learning-Volume 70. JMLR. org, (2017).*

*Nixon, J, et coll. "Measuring Calibration in Deep Learning." arXiv preprint arXiv:1904.01685 (2019).*

# Gestion des erreurs des CNNs

Deux méthodes principalement utilisées :

La calibration : transformer le score en probabilité

L'apprentissage d'une fonction de rejection



Trop rigide pour notre application

*Cortes, C., et coll. "Boosting with abstention." Advances in Neural Information Processing Systems. (2016).*

*Geifman, Y, and Ran E.. "Selective classification for deep neural networks." Advances in neural information processing systems. (2017).*

*Corbière, C., et coll. "Addressing Failure Prediction by Learning Model Confidence." arXiv preprint arXiv:1910.04851 (2019).*

Comment gérer les mauvaises classifications du Deep Learning?

- 1) Comprendre les sorties d'un CNN
- 2) Détecter les erreurs de classification

# Gestion des erreurs des CNNs

Réseau : **ResNet**

130 vidéos, 69 169 images, 20 espèces.

5 augmentations de données → 691 690 images  
(Contraste 60%, 80%, 120%, 140% et flips).

## Entraînement du CNN

Base de données  
d'entraînement (**T0**) :



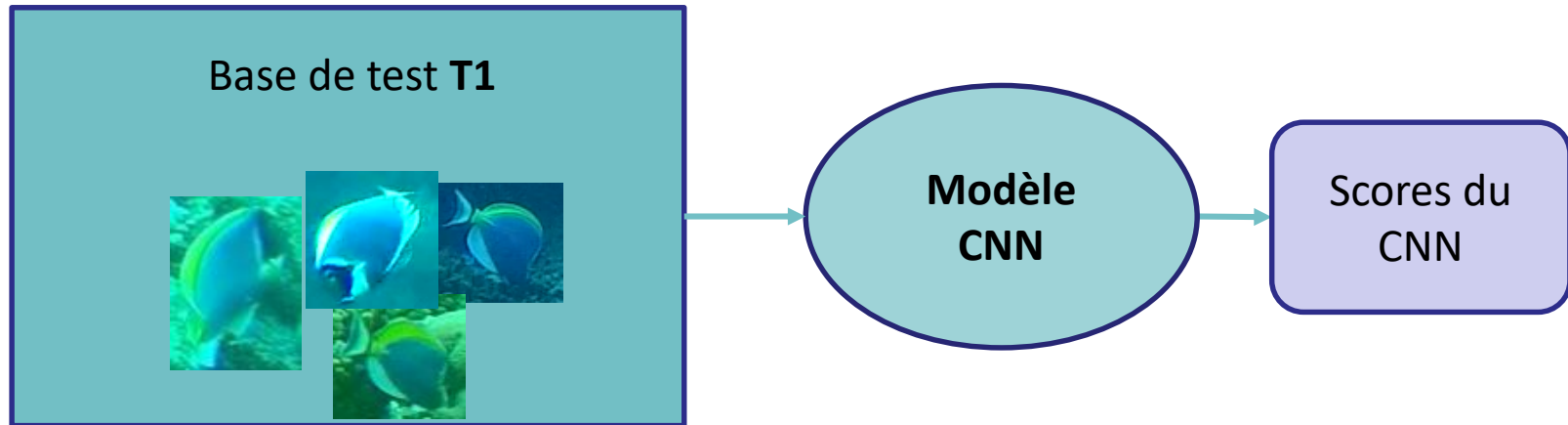
Entraînement du  
CNN

Modèle CNN



# Gestion des erreurs des CNNs

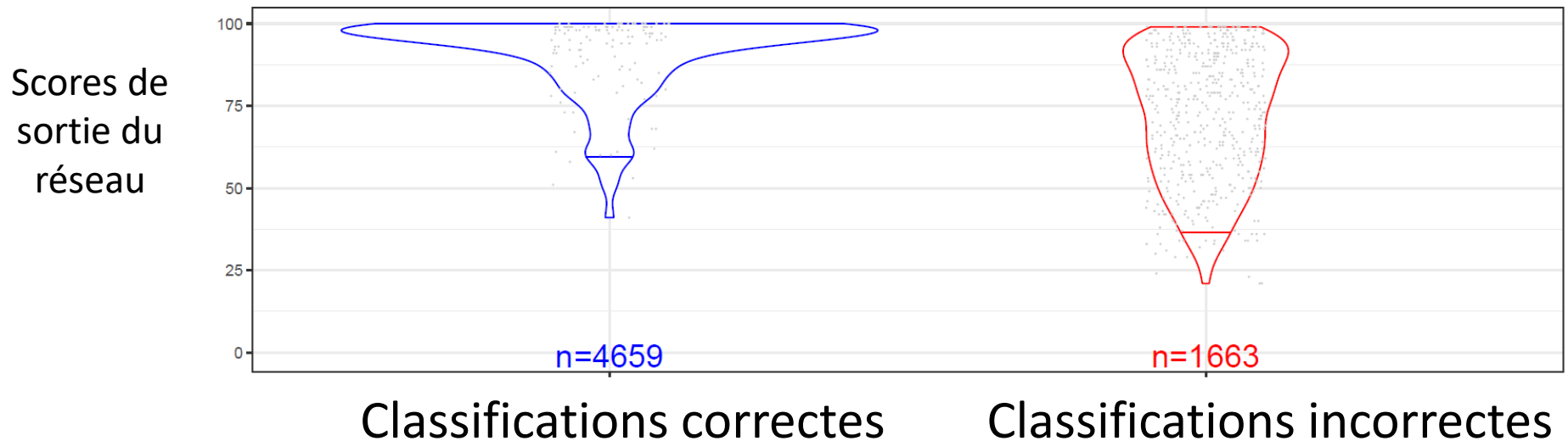
Test indépendant : 6 320 images, 20 espèces



# Gestion des erreurs des CNNs

Observations sur la distribution des classifications

Résultats des classifications, toutes espèces.



# Gestion des erreurs des CNNs

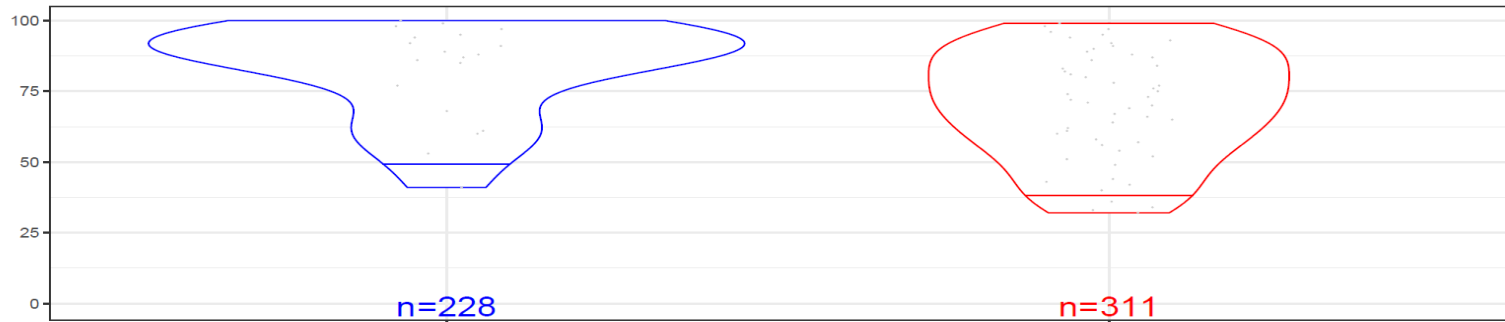
## Observations sur la distribution des classifications



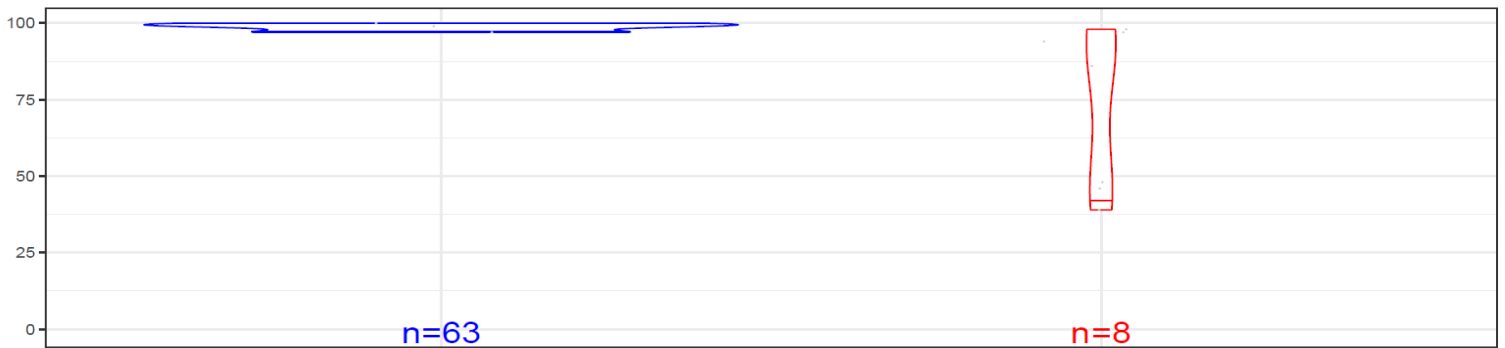
Scores de sortie du réseau



*Naso brevirostris*



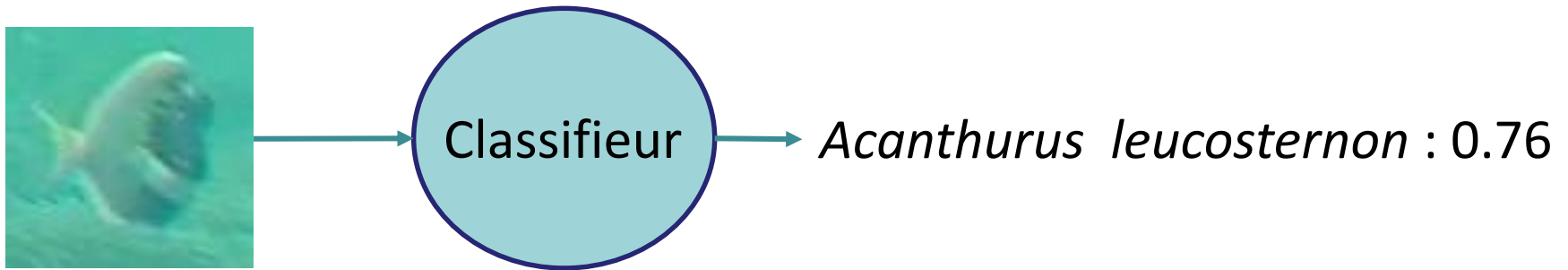
*Chaetodon trifasciatus*



Classifications correctes

Classifications incorrectes

## Définition de seuils



## Application du seuil de l'espèce

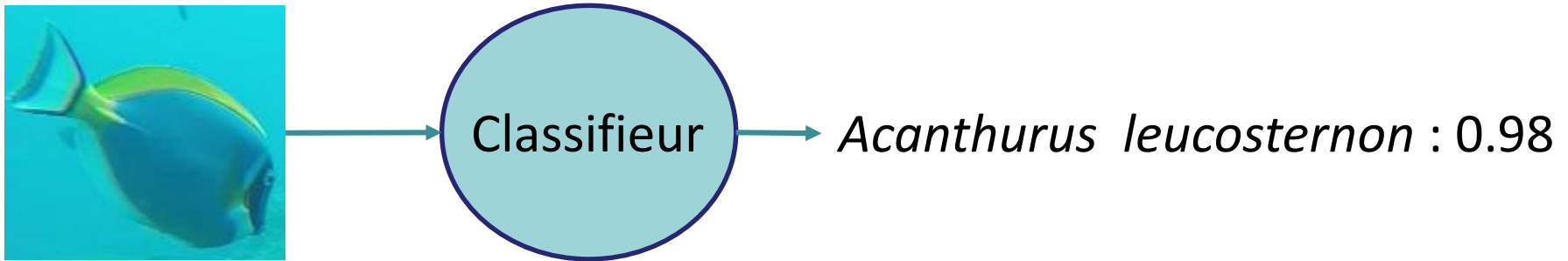
$0.76 < \tau (A. leucosternon)$

## Réponse finale

Unsure

(*Acanthurus leucosternon*)

## Définition de seuils



## Application du seuil de l'espèce

## Réponse finale

0.98 >  $\tau (A. leucosternon)$  → *Acanthurus leucosternon*

## Base de données

Base d'entraînement

691 690 imageries  
20 espèces

Base indépendante  
de création des seuils

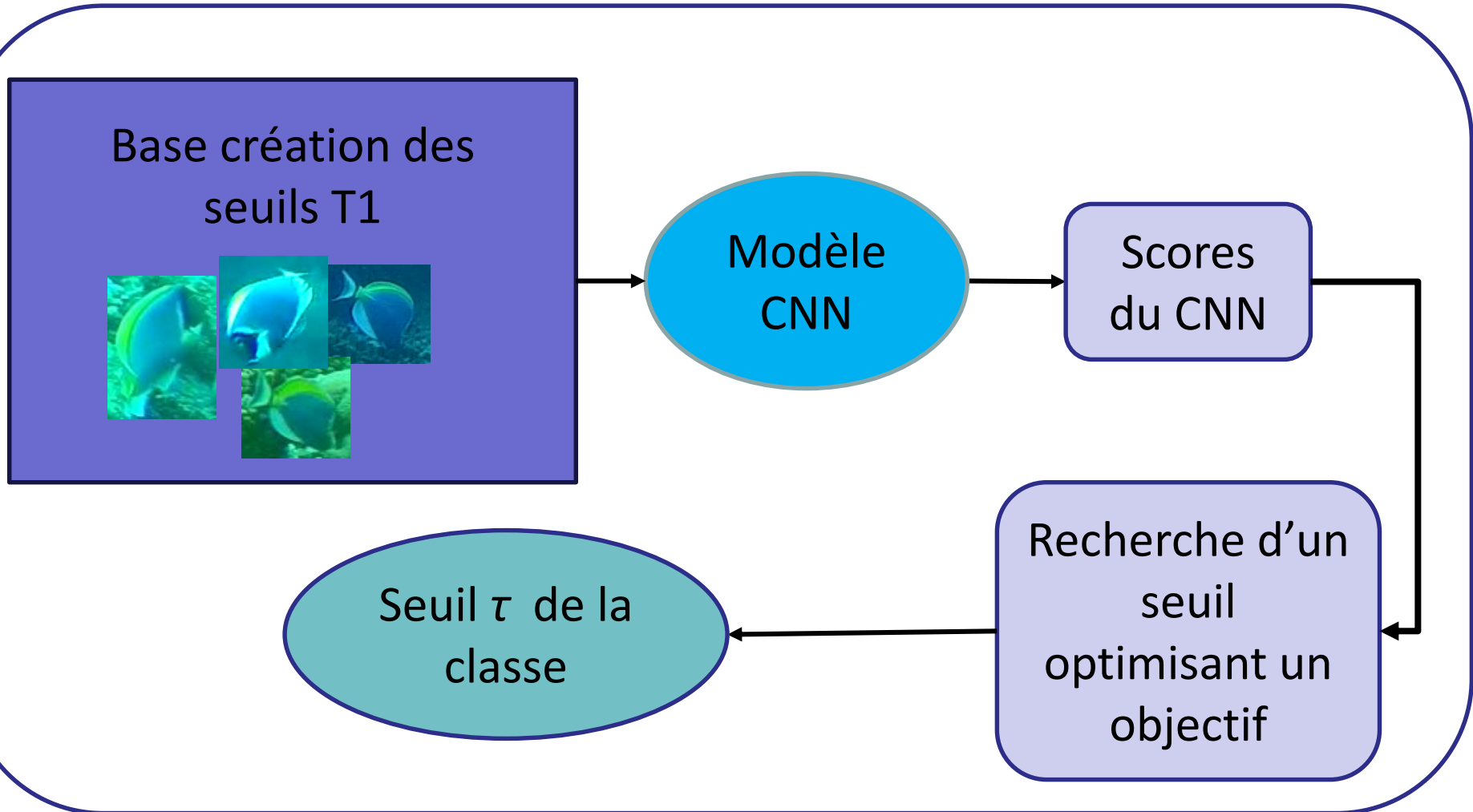
6 320 imageries  
20 espèces

Base indépendante  
de test

13 232 imageries  
20 espèces



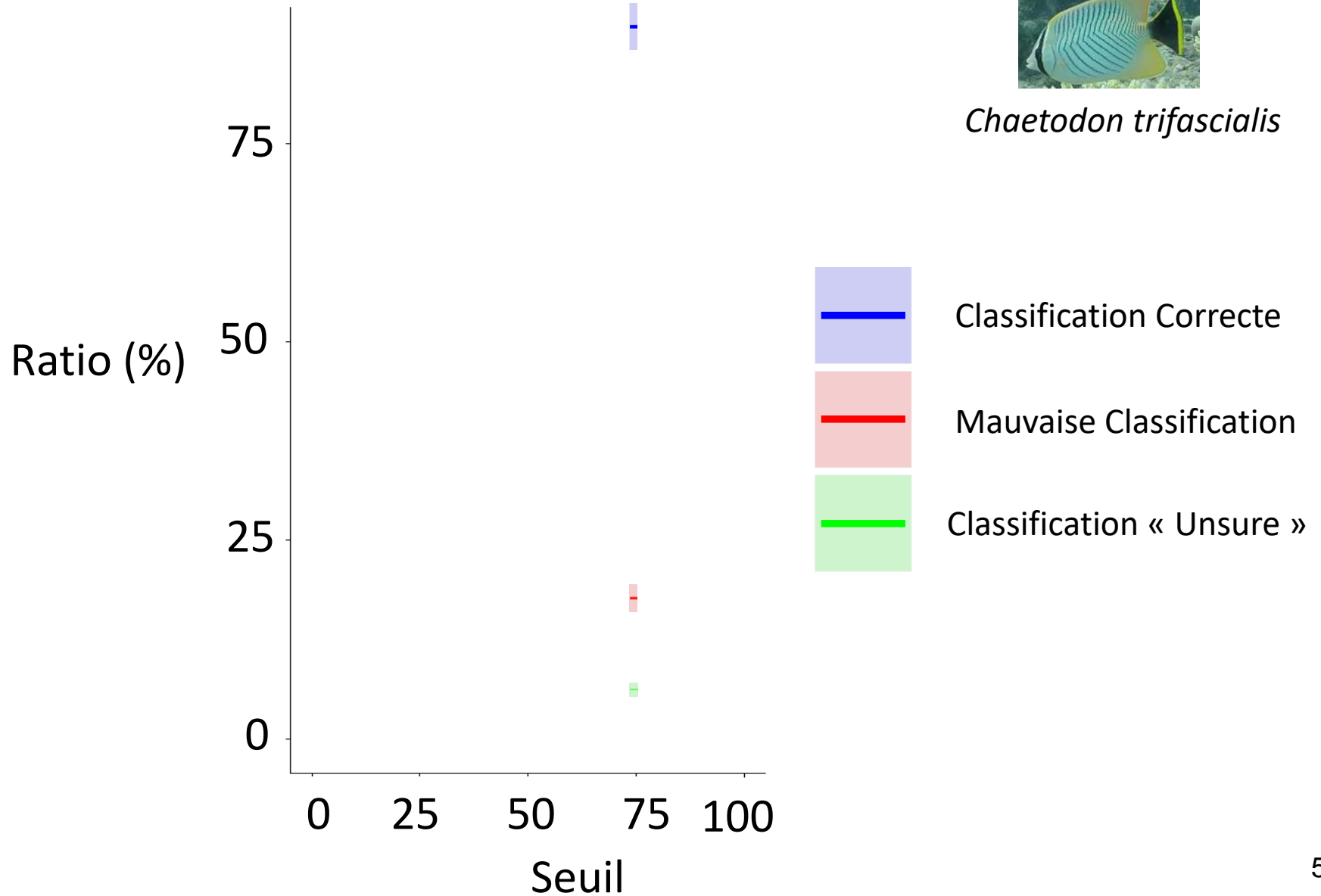
## Recherche des seuils



# Gestion des erreurs des CNNs



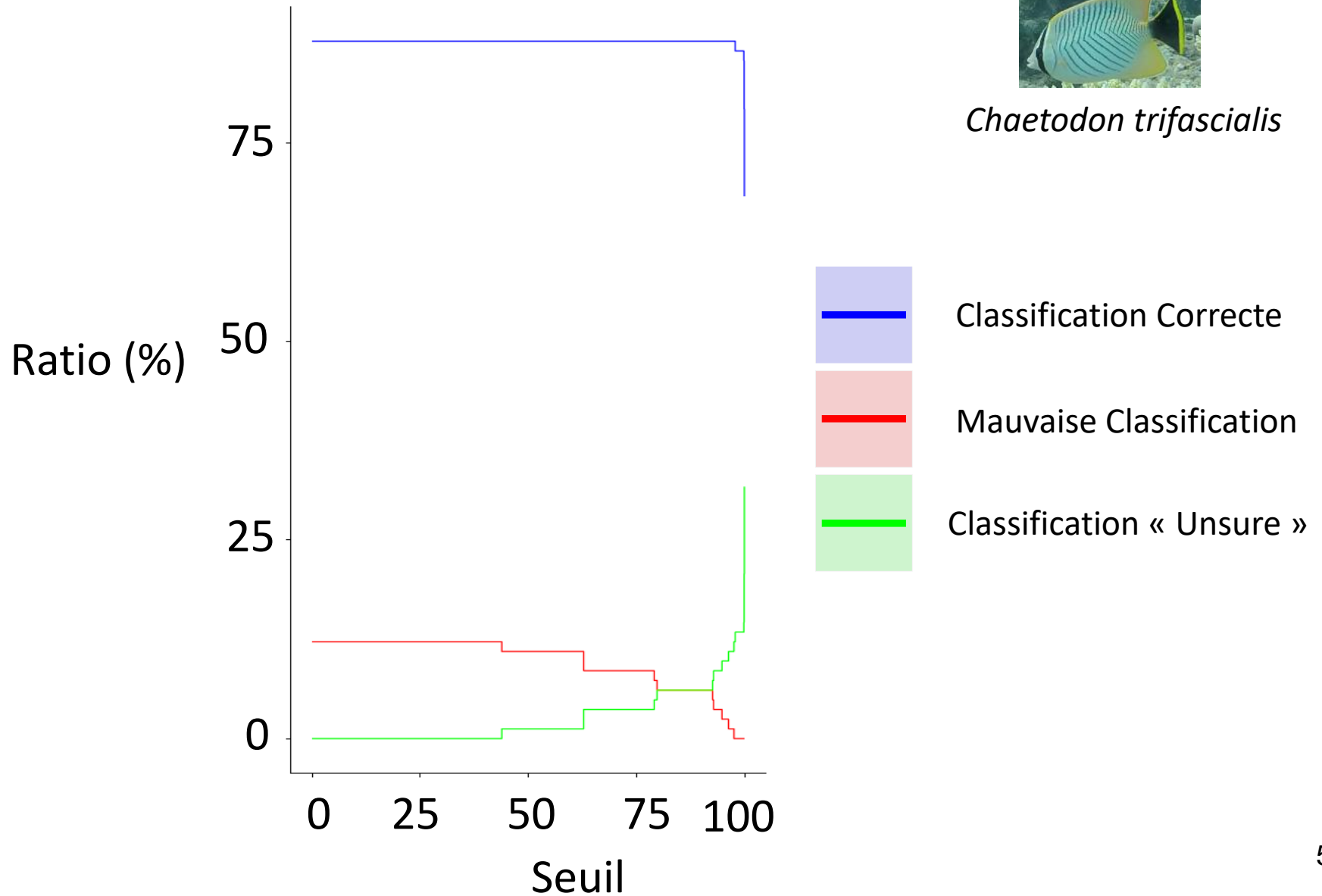
*Chaetodon trifascialis*



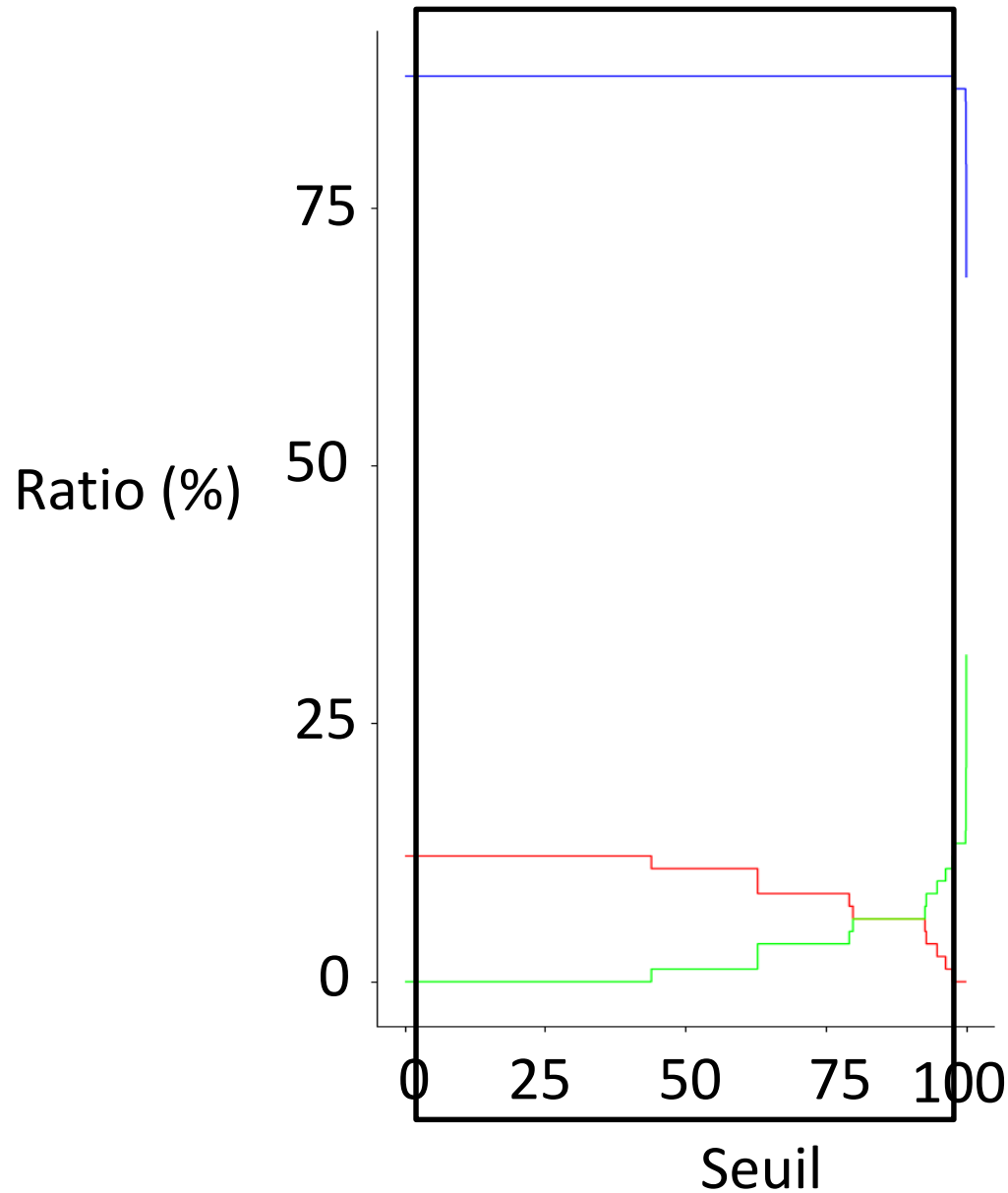
# Gestion des erreurs des CNNs



*Chaetodon trifascialis*



# Gestion des erreurs des CNNs



Objectif n° 1 : Maximiser le taux de bonnes classifications

$$Se_{g1} = \arg \max_{\tau} CC_i(\tau)$$

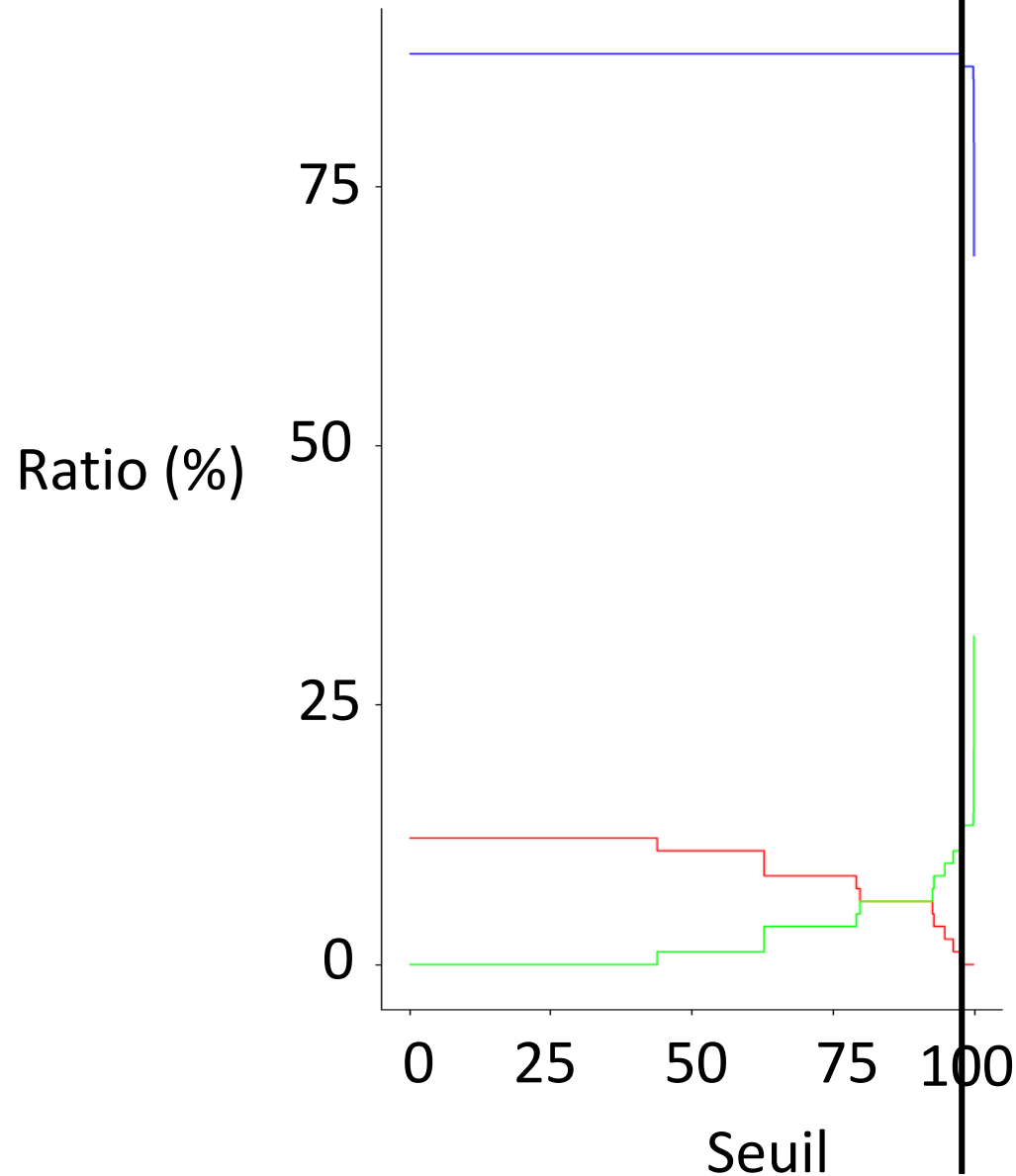
# Gestion des erreurs des CNNs



Objectif n° 1 : Maximiser le taux de bonnes classifications

$$Se_{g1} = \arg \max_{\tau} CC_i(\tau)$$

$$\tau_i = \arg \min_{\tau' \in Se_{g1}} MC_i(\tau')$$

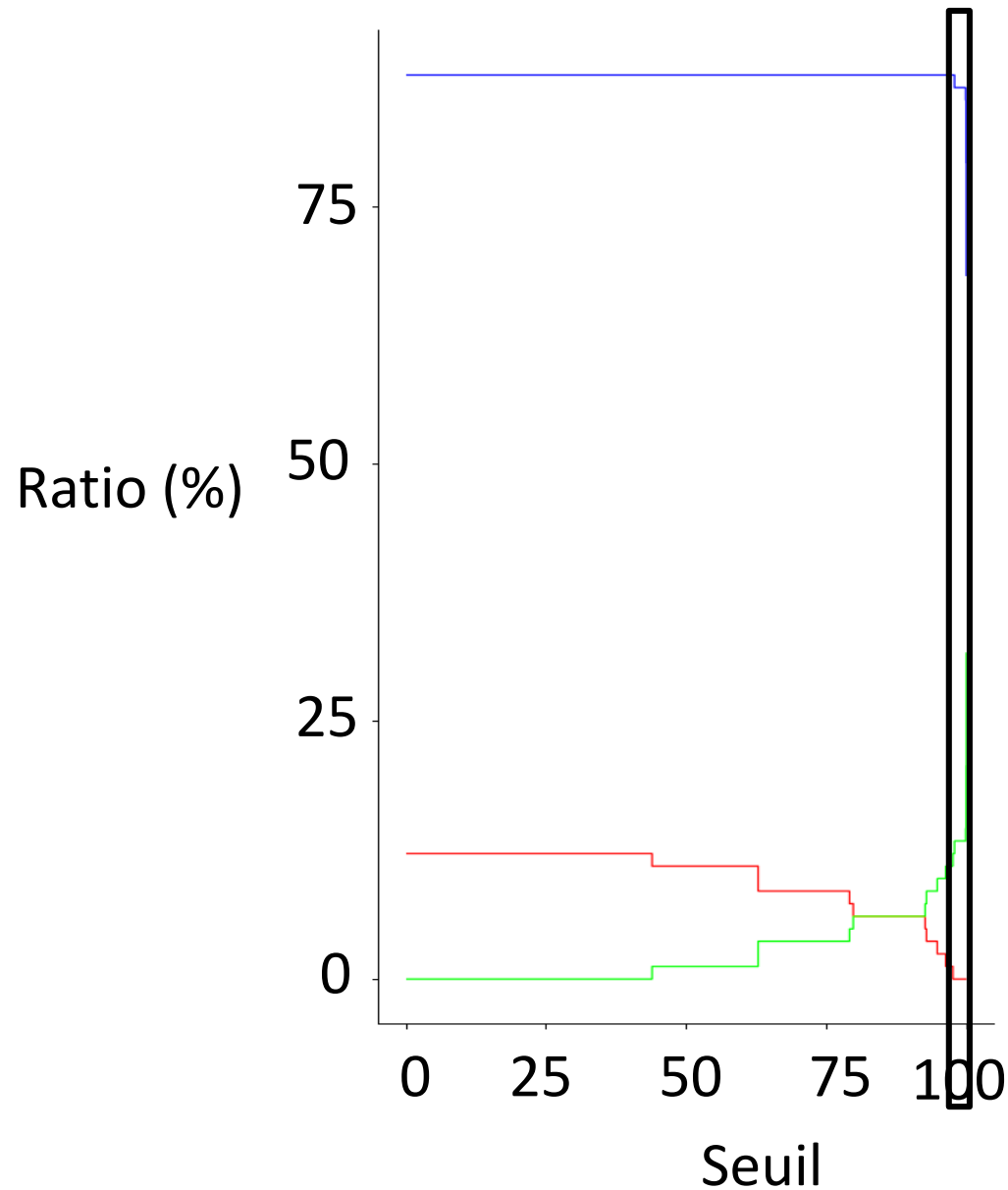


# Gestion des erreurs des CNNs



Objectif n° 2 : Minimiser le taux de mauvaises classifications

$$Se_{g2} = \arg \min_{\tau} MC_i(\tau)$$

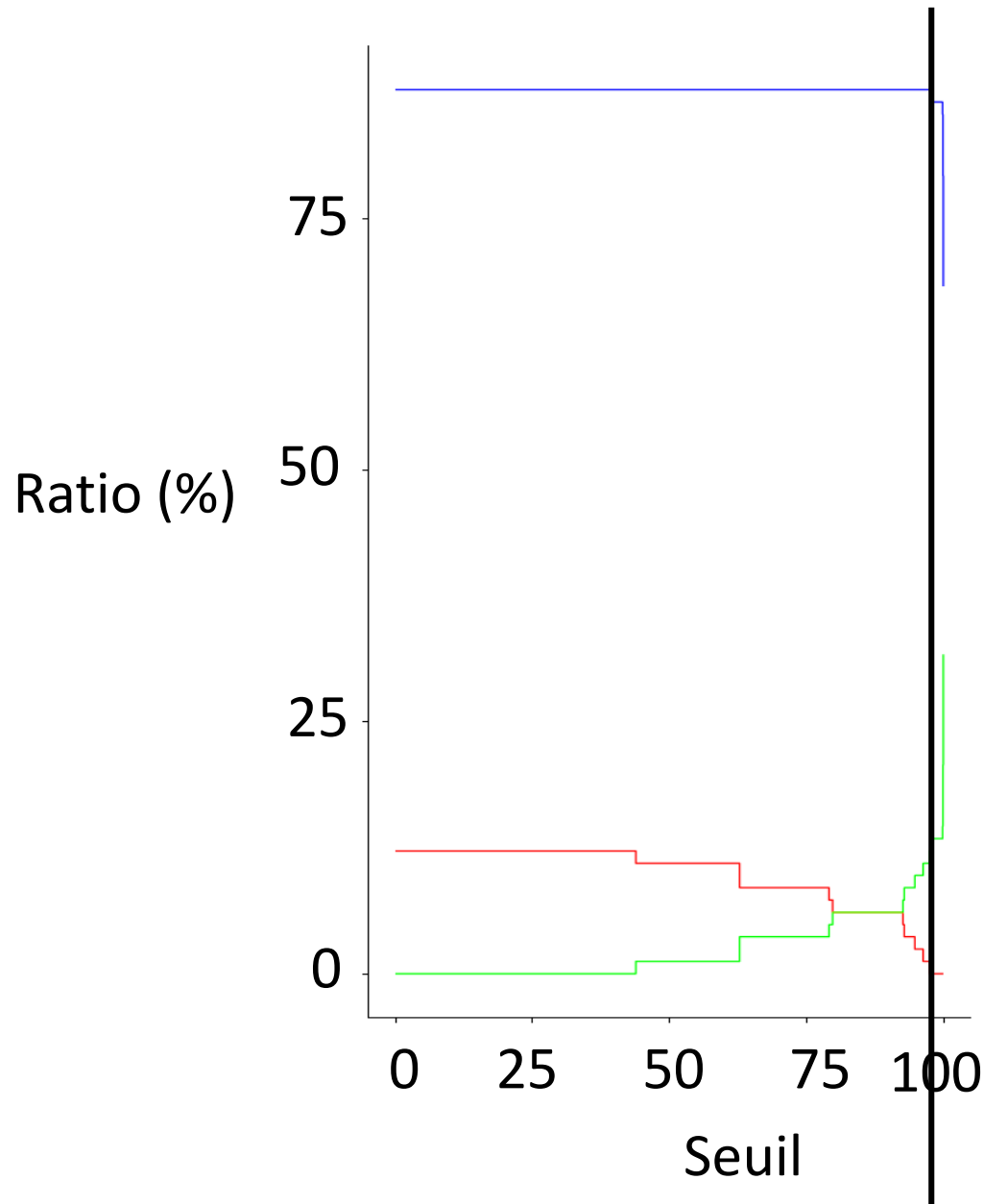




# Gestion des erreurs des CNNs



Objectif n° 2 : Minimiser le taux de mauvaises classifications

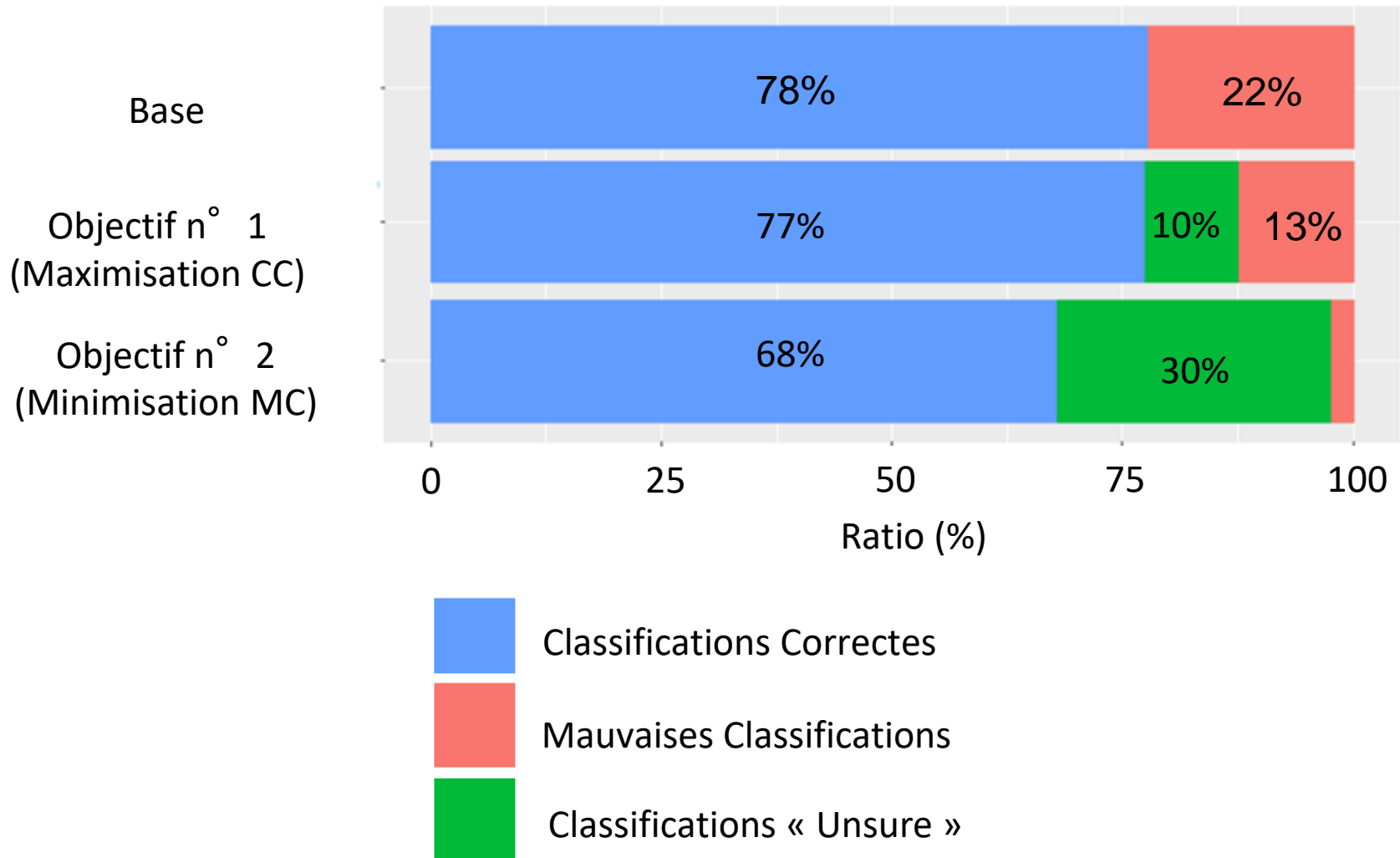


$$Se_{g2} = \arg \min_{\tau} MC_i(\tau)$$

$$\tau_i = \arg \max_{\tau' \in Se_{g2}} CC_i(\tau')$$

# Gestion des erreurs des CNNs

## Résultats

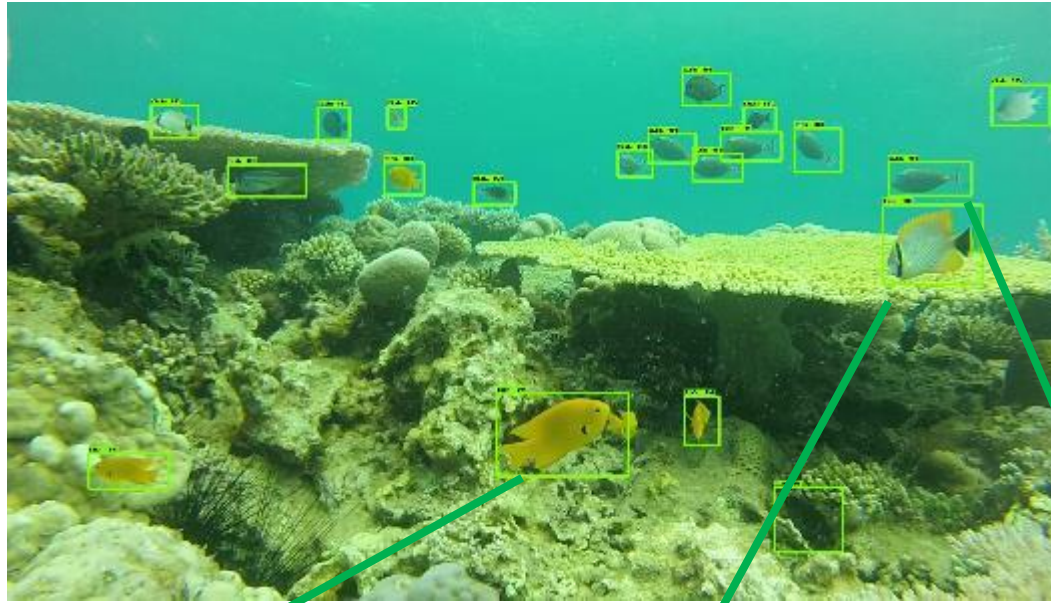


- Quelle méthode choisir pour le recensement automatique?

Comment améliorer la base d'entraînement?

- Le traitement automatique est-il aussi efficace que l'humain pour l'identification d'espèces de poissons?
- Comment gérer les mauvaises classifications du traitement automatique
- **Vers un outil opérationnel de recensement automatique?**

# Vers un outil complet : Localisation et Identification



*Pomacentrus sulfureus*



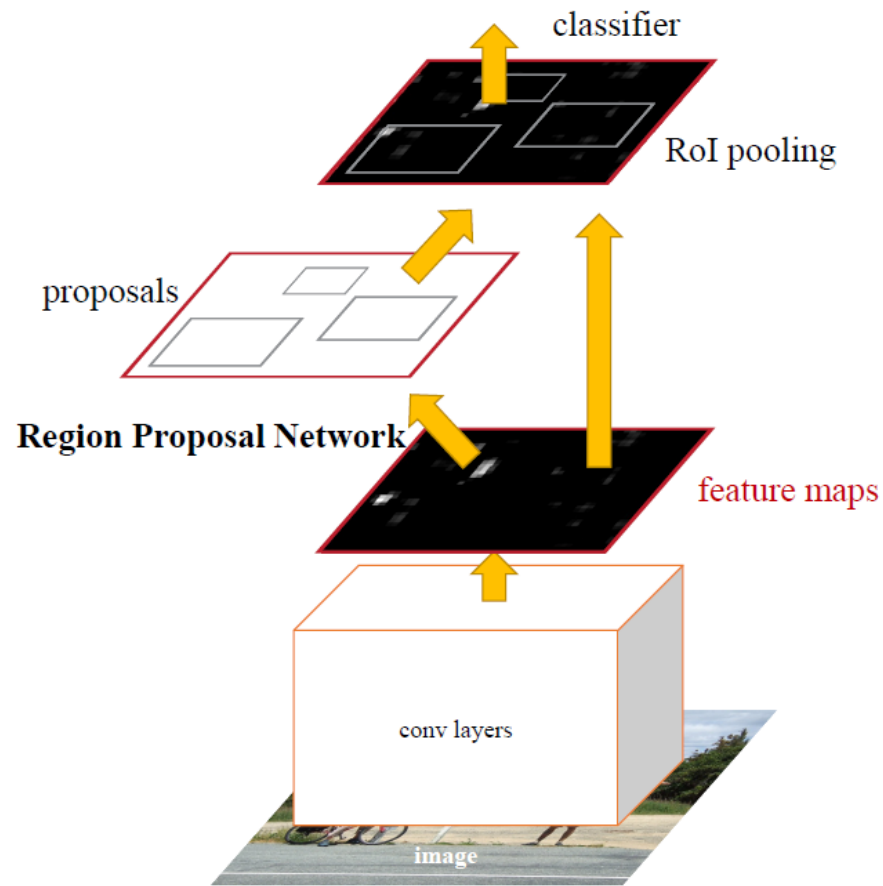
*Chaetodon trifascialis*



*Naso elegans*

# Vers un outil complet : Localisation et Identification

## Le réseau de localisation



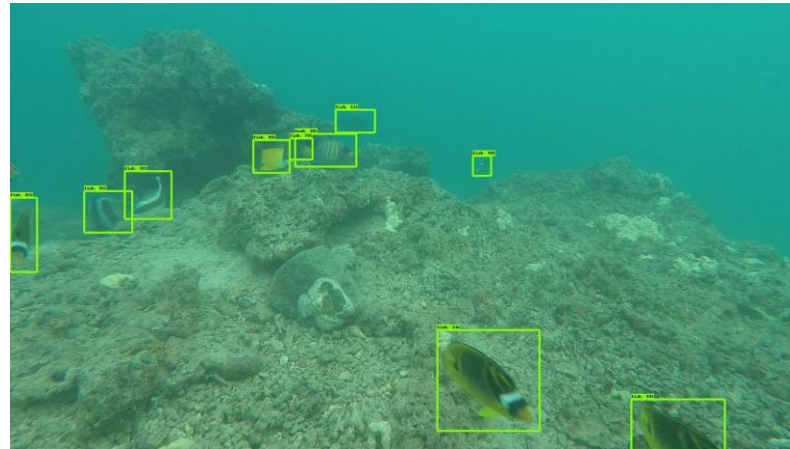
# Vers un outil complet : Localisation et Identification



YOLO



SSD



Faster R-CNN

Ren, S., et coll. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems*. (2015).

Liu, W., et coll. "Ssd: Single shot multibox detector." *European conference on computer vision*. (2016).

Redmon, J., et coll. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. (2016).



Possibilité 1 :

1 algorithme, identification+localisation

+ Synergie entre la localisation et l'identification

+ Optimisation du temps de calcul

- Annotation exhaustive pour localisation et identification



# Vers un outil complet : Localisation et Identification

Possibilité 2 :  
2 algorithmes, 1 localisation 1 identification  
- Temps de calcul plus long  
+ Annotation exhaustive pour localisation  
+ Équilibrage possible des classes

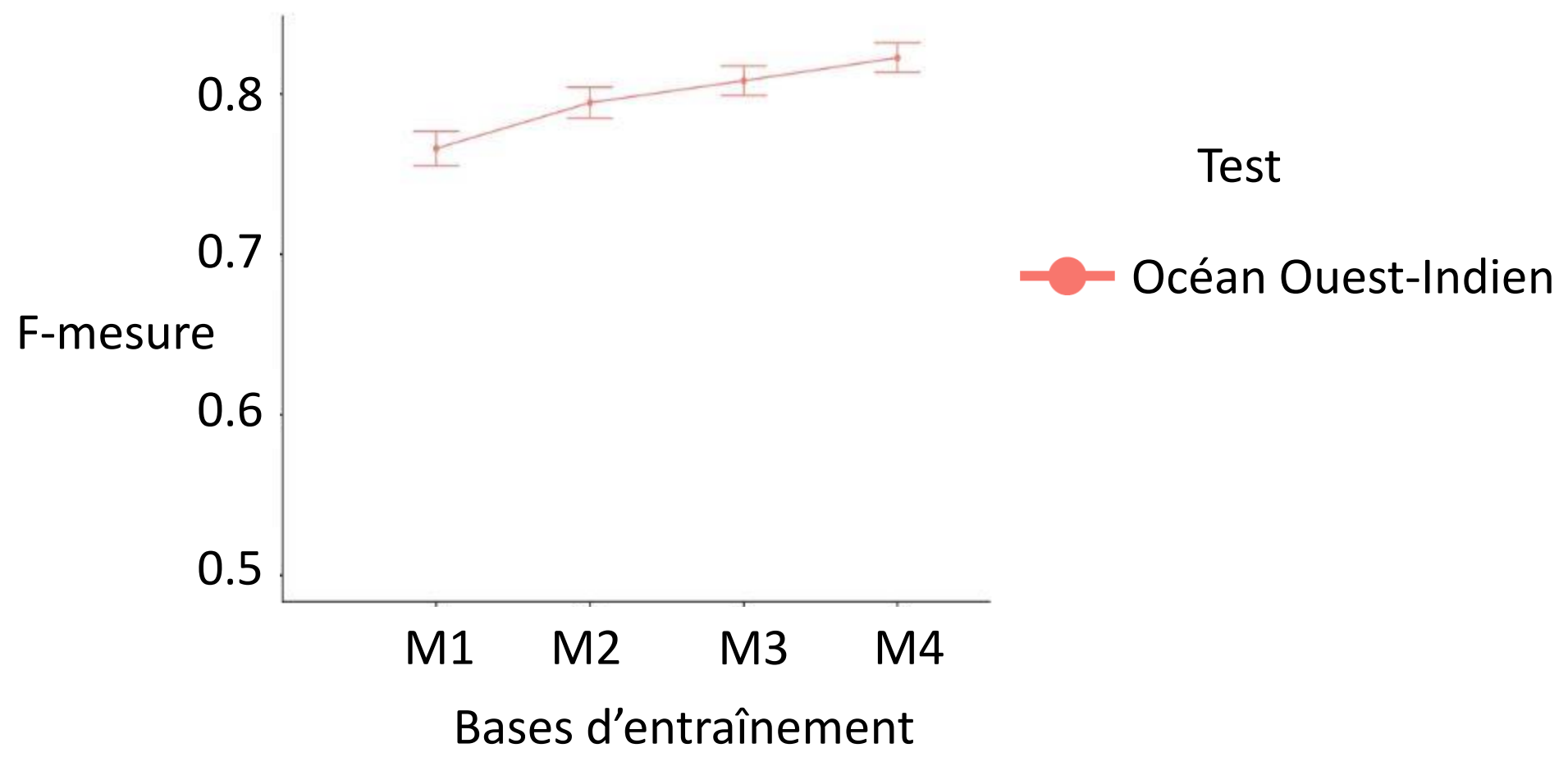


# Vers un outil complet : Localisation et Identification

## Océan Ouest-Indien

<b>Base de données</b>	<b>Annotations</b>	<b>Classes</b>	<b>Temps d'annotation (heures)</b>
<b>M1</b>	3 603	40	9
<b>M2</b>	6 146	57	16
<b>M3</b>	11 230	97	29
<b>M4</b>	17 708	125	46

# Vers un outil complet : Localisation et Identification





# Vers un outil complet : Localisation et Identification



# Vers un outil complet : Localisation et Identification

○ Base de test Océan Ouest Indien

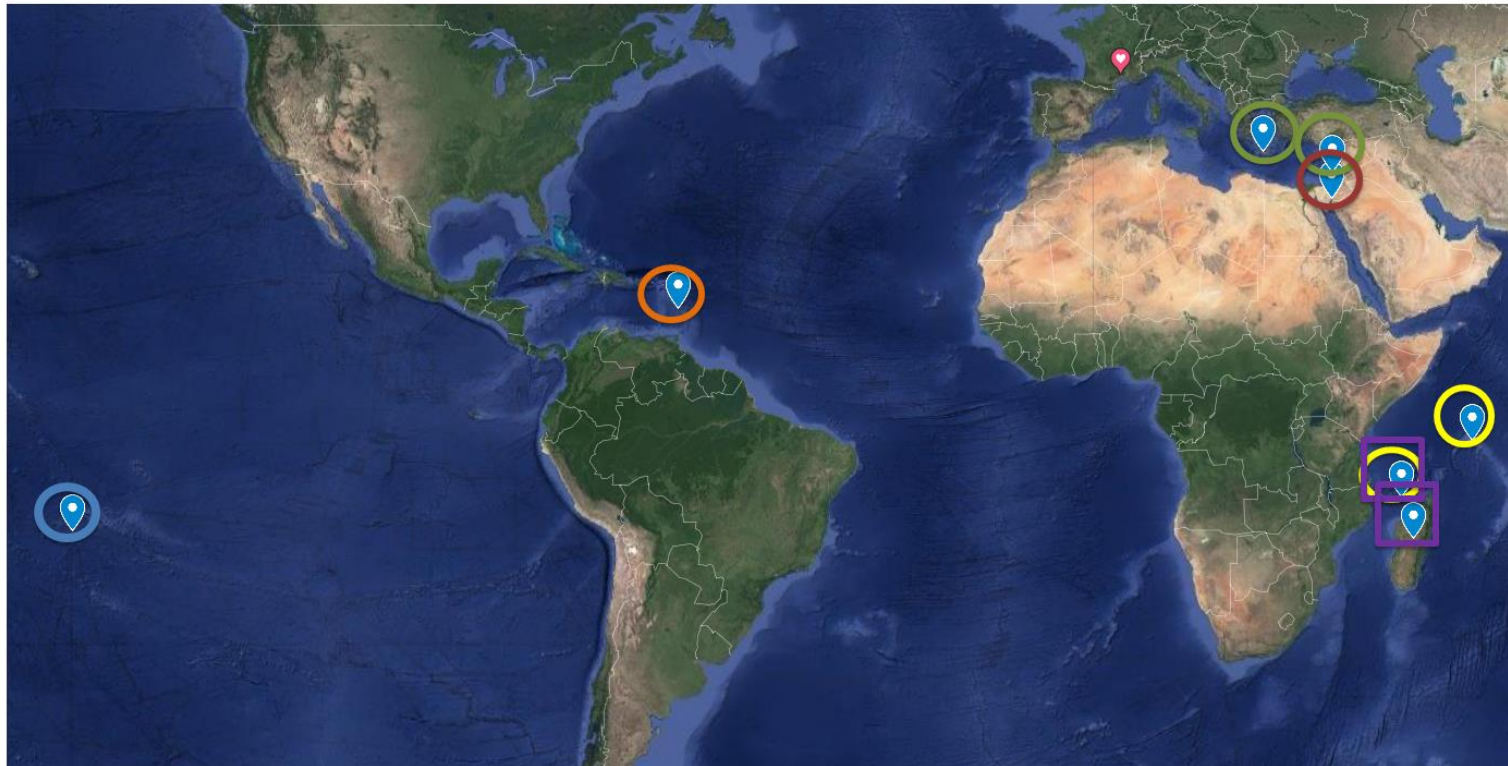
○ Base de test Caraïbes

○ Base de test Mer Rouge

○ Base de test Pacifique Sud

○ Base de test Méditerranée

□ Base d'entraînement Océan Ouest-Indien

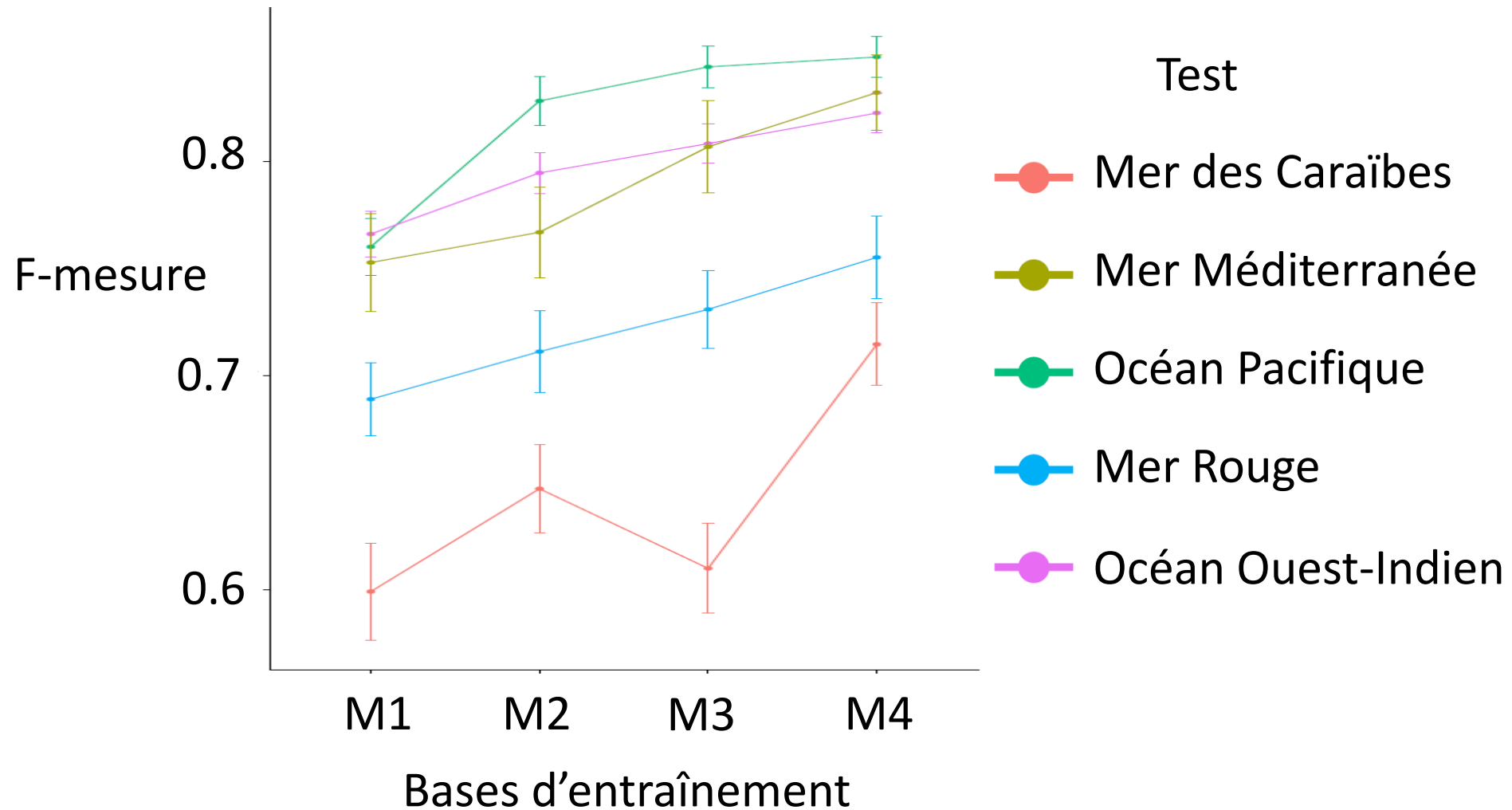




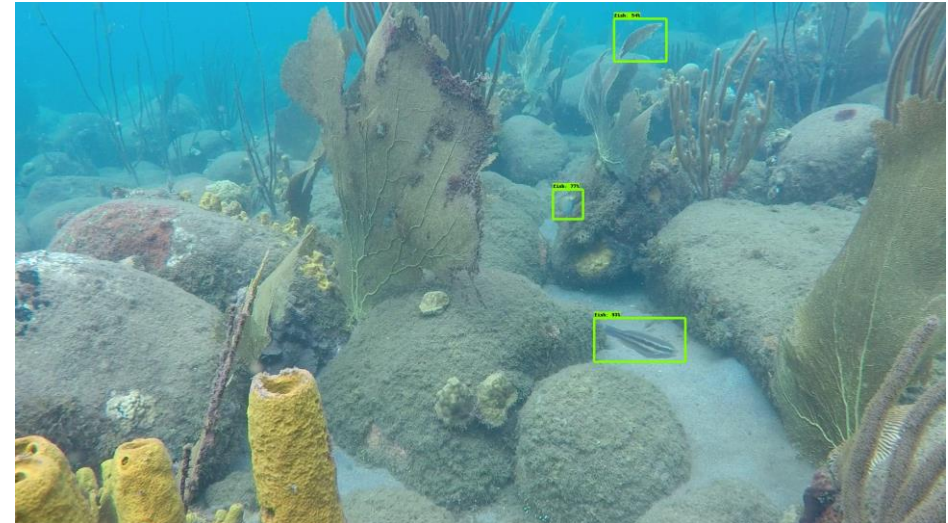
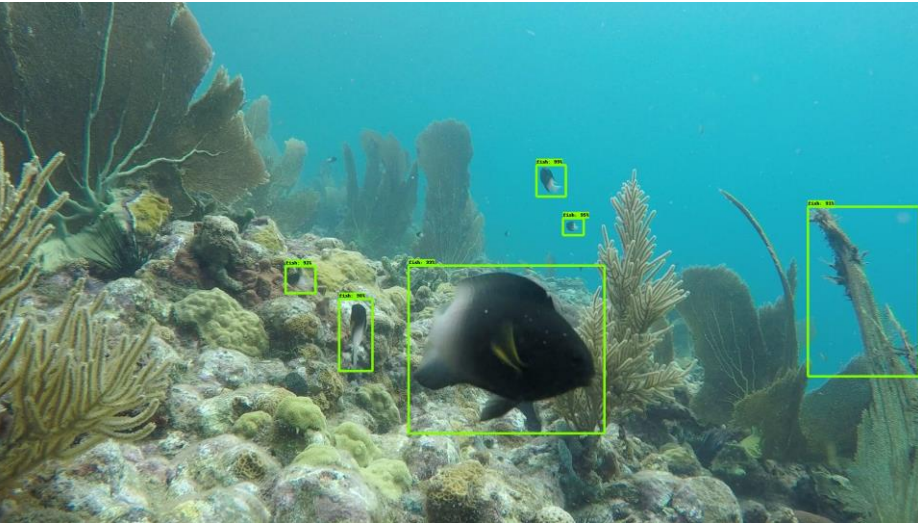
# Vers un outil complet : Localisation et Identification

<b>Bases de test</b>	<b>Paysage marin</b>	<b>Nombre de vidéos</b>	<b>Classes partagées avec l'entraînement</b>
<b>Ouest-Indien</b>	Coraux "durs"	15	52/66
<b>Mer Rouge</b>	Coraux "durs"	4	7/28
<b>Pacifique</b>	Coraux "durs"	4	17/36
<b>Caraïbes</b>	Coraux "mous"	4	1/21
<b>Méditerranée</b>	Algues	5	1/13

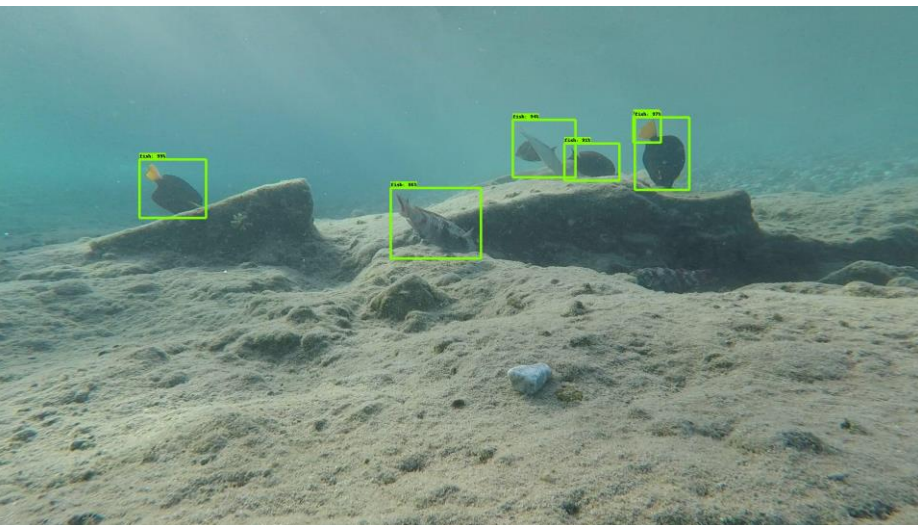
# Vers un outil complet : Localisation et Identification



# Vers un outil complet : Localisation et Identification



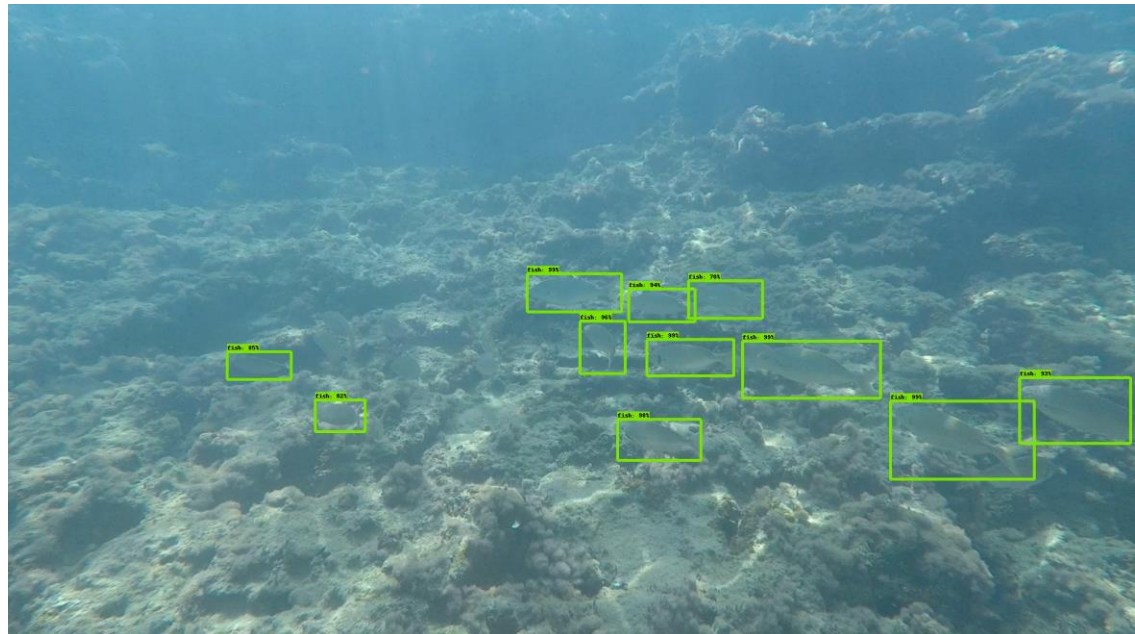
Caraïbes



Mer rouge

# Vers un outil complet : Localisation et Identification

Méditerranée (Crête)



# Conclusion

Le deep learning est-il efficace par rapport au machine learning?

**Oui!**

Le Deep Learning est-il aussi efficace que l'humain pour l'identification d'espèces de poissons?

**Oui, dans le cas de la classification d'images sans contexte**

Comment gérer les mauvaises classifications du Deep Learning?

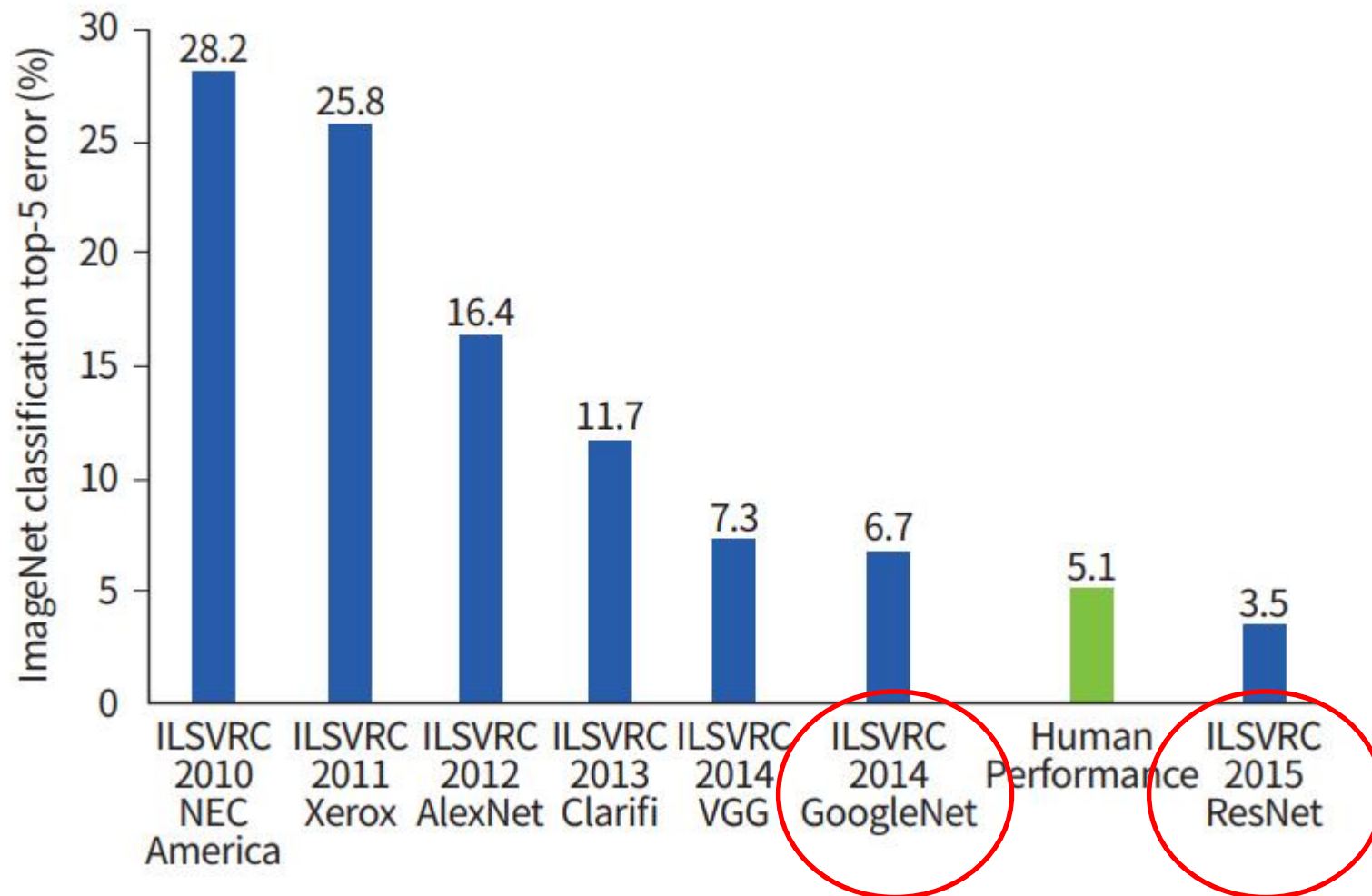
**Oui, avec des méthodes de post-traitement**

Peut-on envisager un outil opérationnel de recensement automatique?

**Un grand effort d'annotation requis**



# Perspectives



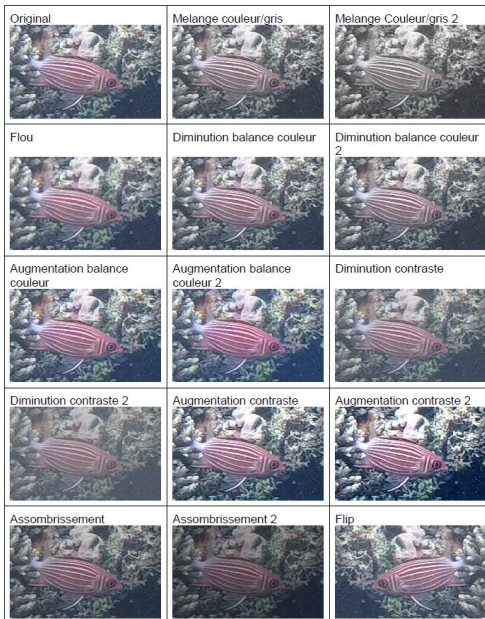
# Perspectives

8 Classes  
5 856 imageries

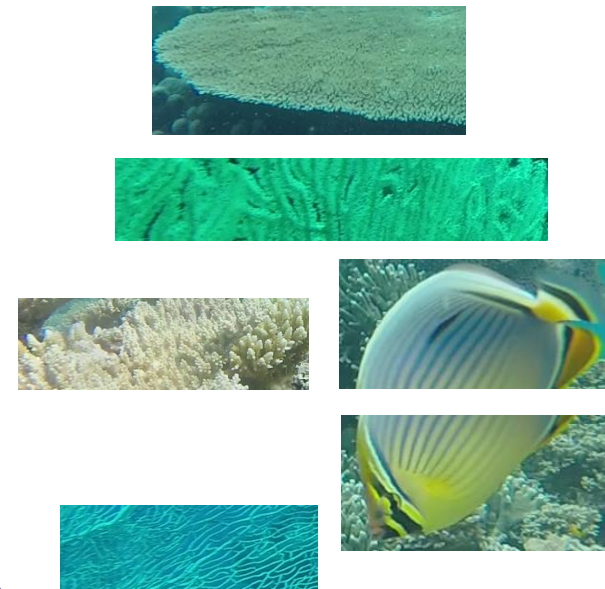
18 Classes  
46 325 imageries

20 Classes  
61 690 imageries

## Enrichissement de données



## Ajout de classes spécifiques





Une première utilisation: Détection d'espèces exotiques



Être capable de localiser des espèces non apprises



Être capable d'identifier des espèces dans un contexte différent

## Fusion d'informations

Informations  
images

Informations  
stéréos

Informations  
vidéos

Informations écologiques :  
Comportement, taille max,  
distribution géographique...

*McLaughlin, N. et coll. Recurrent convolutional network for video-based person re-identification. In Proceedings of the IEEE conference on computer vision and pattern recognition(2016).*

*Chuang, M.C., et coll. "Tracking live fish from low-contrast and low-frame-rate stereo videos." IEEE Transactions on Circuits and Systems for Video Technology 25.1 (2014).*

*Li, Xiaojing, et coll. "Real-Time Underwater Fish Tracking Based on Adaptive Multi-Appearance Model." 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, (2018).*

# Remerciements



Villon, S., Chaumont, M., Subsol, G., Villéger, S., Claverie, T., & Mouillot, D. (2016). **Coral reef fish detection and recognition in underwater videos by supervised machine learning: Comparison between Deep Learning and HOG+ SVM methods.** In *International Conference on Advanced Concepts for Intelligent Vision Systems*

Villon, S., Mouillot, D., Chaumont, M., Darling, E. S., Subsol, G., Claverie, T., & Villéger, S. (2018). **A Deep learning method for accurate and fast identification of coral reef fishes in underwater images.** *Ecological informatics*, 48, 238-244.

Villon, S., Mouillot, D., Chaumont, M., Subsol, G., Claverie, T., & Villéger, S. (2019). **A new method to control error rates in automated species identification with deep learning algorithms.** *Scientific Reports*, soumis.

Villon, S., Mouillot, D., Chaumont, M., Subsol, G., Claverie, T., & Villéger, S. (2019). **Transferability of Deep Learning algorithms for automatic fish detection in underwater videos between different coastal marine ecosystems.** En cours de soumission.

## Utilisation de l'information vidéo

Suivi de poissons

CNNs orientés  
« vidéo »

*Mobahi, H. et coll. "Deep learning from temporal coherence in video." Proceedings of the 26th Annual International Conference on Machine Learning. ACM, (2009).*

*McLaughlin, N. et coll. Recurrent convolutional network for video-based person re-identification. In Proceedings of the IEEE conference on computer vision and pattern recognition(2016).*

*Chuang, M.C., et coll. "Tracking live fish from low-contrast and low-frame-rate stereo videos." IEEE Transactions on Circuits and Systems for Video Technology 25.1 (2014).*

*Li, Xiaojing, et coll. "Real-Time Underwater Fish Tracking Based on Adaptive Multi-Appearance Model." 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, (2018).*

# Gestion du taux d'erreur des CNNs

## Résultats

	Base	Objectif n°1	Objectif n°2
Classifications "unsure" (%)	0	10.2	29.7
Classifications incorrects (%)	22.1	12.3	2.2
Classifications corrects (%)	77.9	77.4	67.9



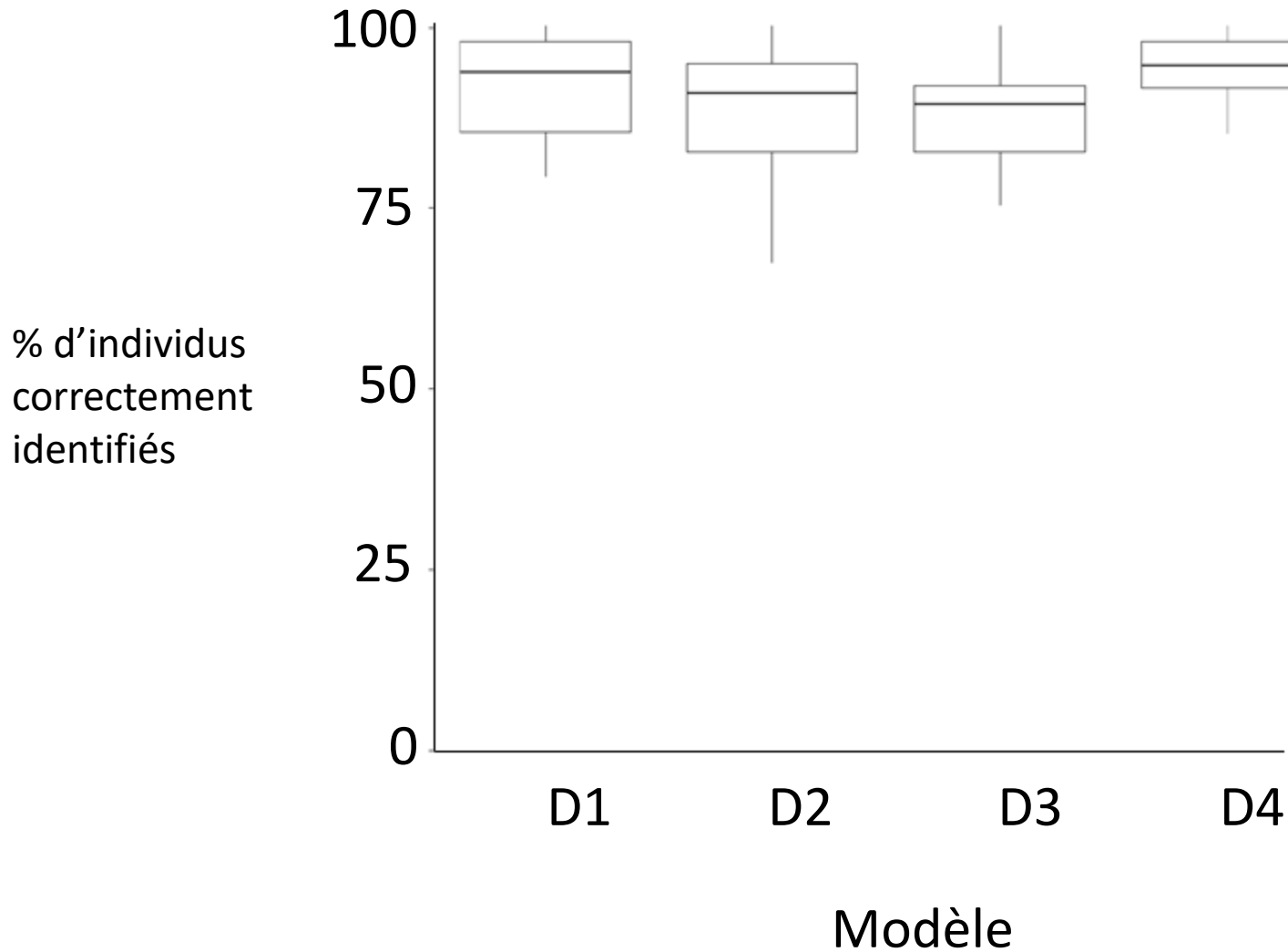
# Gestion du taux d'erreur des CNNs

## Résultats

	Base	Objectif n°1	Objectif n°2
Classifications "unsure" (%)	0	10.2	29.7
Classifications incorrects (%)	22.1	12.3	2.2
Classifications corrects (%)	77.9	77.4	67.9

# Expertise humaine et Deep Learning

**94.1%** de vrais positifs sur les 20 espèces

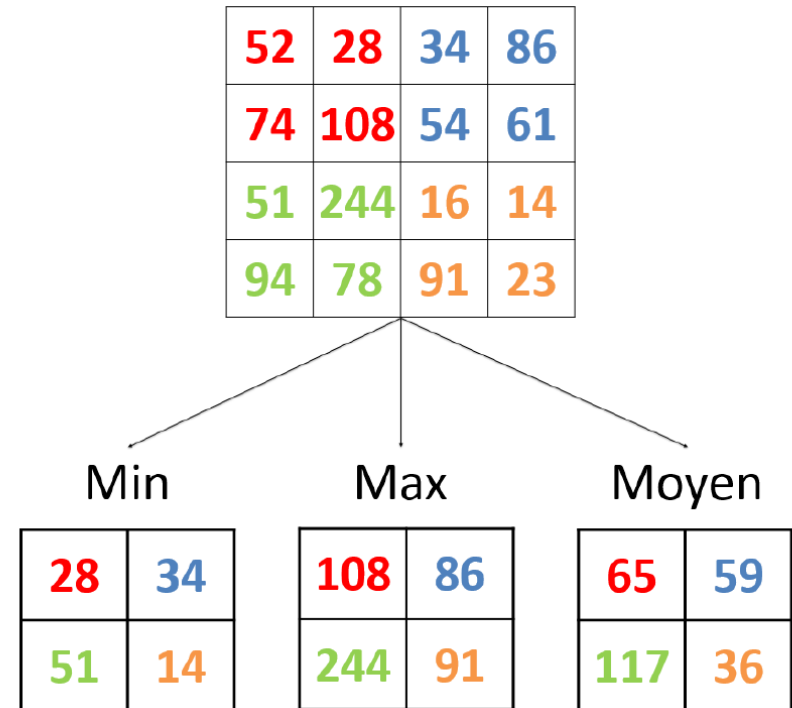


# Identification automatique d'images

## Deep Learning en profondeur : Le pooling

Exemples : *Pooling* Min, Max, et Moyen.

Un *pooling* est défini par 3 variables : la taille du filtre, le type de regroupement et le déplacement (*stride*).



$$CC_i(\tau) = (X = i|Y = i) = \frac{\#(X = i|Y = i) - \#(X = i, S_i < \tau|Y = i)}{\#(Y = i)}$$

Avec :

- $X$  : le label prédit
- $Y$  : le label Vérité Terrain
- $i$  : l'espèce de poisson,
- $S_i$  : le score obtenu
- $\tau$  : le seuil de l'espèce

# Gestion du taux d'erreur des CNNs

$$CC_i(\tau) = (X = i|Y = i) = \frac{\#(X = i|Y = i) - \#(X = i, S_i < \tau|Y = i)}{\#(Y = i)}$$

$$MC_i(\tau) = (X \neq i|Y = i) = \frac{\#(X \neq i|Y = i) - \#(X \neq i, S_i < \tau|Y = i)}{\#(Y = i)}$$

Avec:

- $X$ : le label prédit
- $Y$ : le label Vérité Terrain
- $i$ : l'espèce de poisson,
- $S_i$ : le score obtenu
- $\tau$ : le seuil de l'espèce

# Gestion du taux d'erreur des CNNs

$$CC_i(\tau) = (X = i|Y = i) = \frac{\#(X = i|Y = i) - \#(X = i, S_i < \tau|Y = i)}{\#(Y = i)}$$

$$MC_i(\tau) = (X \neq i|Y = i) = \frac{\#(X \neq i|Y = i) - \#(X \neq i, S_i < \tau|Y = i)}{\#(Y = i)}$$

$$UC_i(\tau) = \frac{\#(X = i, S_i < \tau_i|Y = i) + \#(X \neq i, S_i < \tau|Y = i)}{\#(Y = i)}$$

Avec:  $X$ : le label prédit  
 $Y$ : le label Vérité Terrain  
 $i$ : l'espèce de poisson,  
 $S_i$ : le score obtenu  
 $\tau$ : le seuil de l'espèce



Vers un outil complet d'estimation de la biodiversité



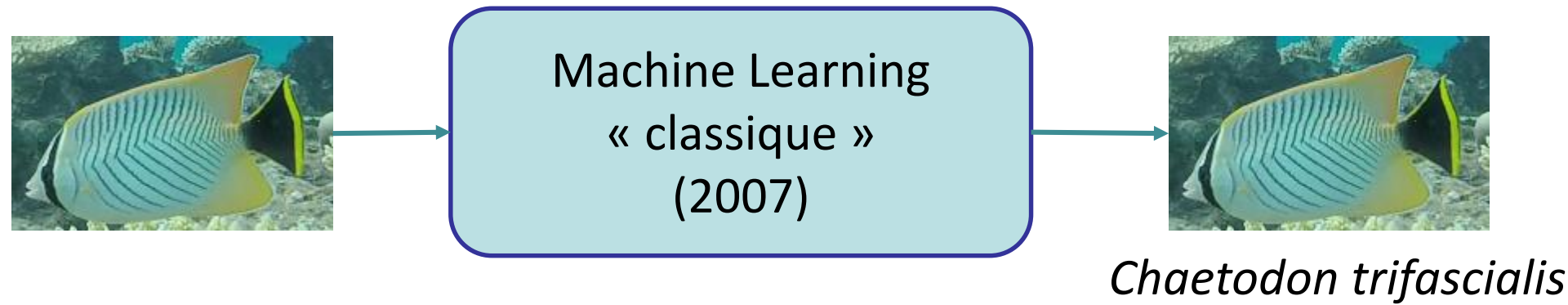
Localisation automatique de poissons



Estimation automatique de biodiversité

# Identification d'images de poisson

## Méthodes de classification par apprentissage

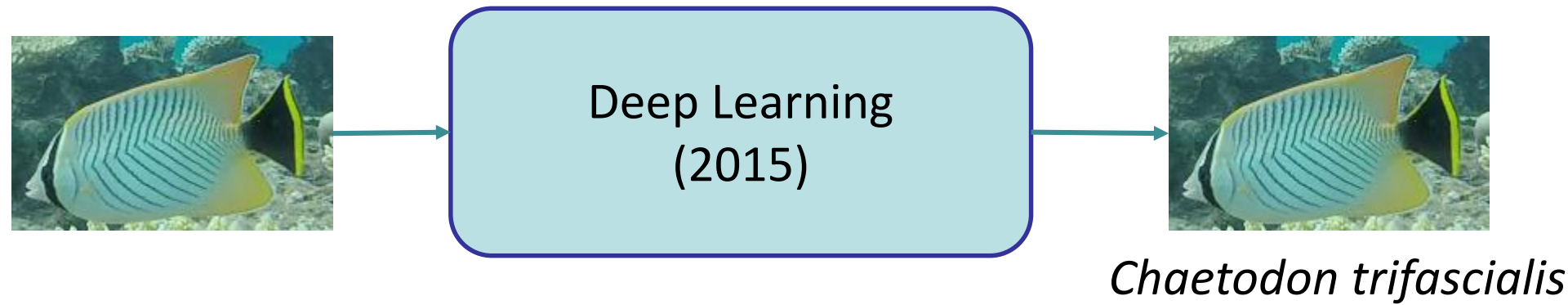


Rova, A., Greg M., and Lawrence M. D. "One fish, two fish, butterfly, trumpeter: Recognizing fish in underwater video." MVA. (2007).

Fisher, R. B., et coll., eds. *Fish4Knowledge: collecting and analyzing massive coral reef fish video data*. Vol. 104. Springer, 2016.

# Identification d'images de poisson

## Classification automatique d'images



Qin, Hongwei, et coll. "DeepFish: Accurate underwater live fish recognition with a deep architecture." *Neurocomputing* (2016).

Rathi, D., et coll. "Underwater fish species classification using convolutional neural network and deep learning." 2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR) (2017).

Deep, B. V., and Ratnakar, D. "Underwater Fish Species Recognition Using Deep Learning Techniques." 2019 6th International Conference on Signal Processing and Integrated Networks (2019).