

*Inria*

CIP: Classification  
d'Images  
Patrimoniales

Équipe Linkmedia – INRIA Rennes

Laurent Amsaleg

Yannis Avrithis

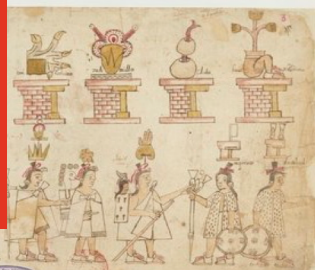
Ewa Kijak

Florent Michel

Miaoqing Shi

# Inria

## CIP: Classification d'Images Patrimoniales



## Base des manuscrits enluminés :

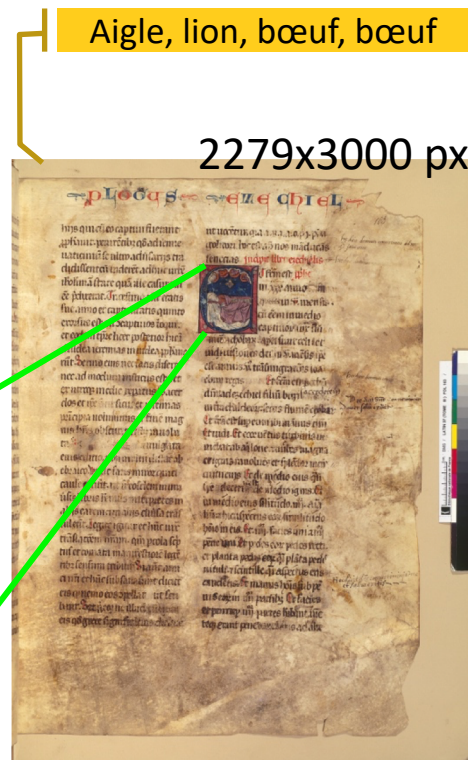
- période -1600 à 1900
- 200 000 enluminures décrites
- cible : la zoologie
  - 23 663 images
  - 52 336 annotations
  - 397 classes
  - 9 super classes (mollusques, poissons, mammifères, reptiles, ...)

## Tâche :

- classification automatique des enluminures vers ces classes

## Difficultés

- annotations :
  - au niveau de la page
  - aucune localisation des objets
  - variété visuelle pour une même classe, ambiguïtés,
  - populations classes très déséquilibrées
- caractéristiques visuelles :
  - taille objet / taille page
  - grande variabilité tailles
  - images ne montrant rien de contemporain



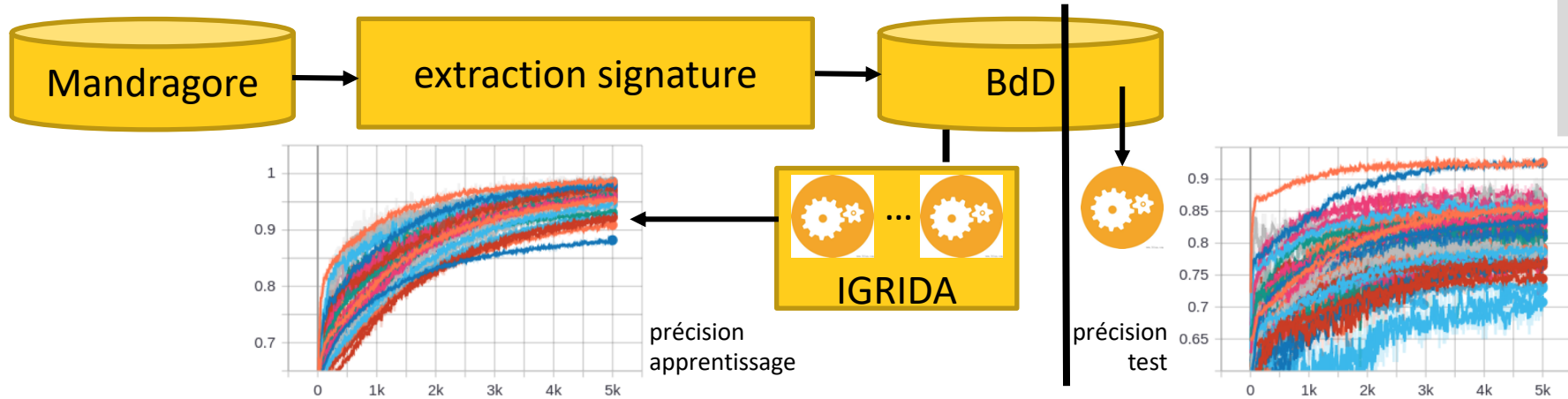
## Solution sur étagère :

- imagenet : 1M d'images contemporaines ; 1000+ classes dont zoologie
- réseau profond Xception (google)

## Supervision faible, classification multi-labels

- mise en correspondance classes imagenet / Mandragore
- augmentation de données (retournements, ...)
- amputation Xception, greffe de multiples classifieurs binaires
- tests de 10 modèles (fonctions d'activation, *dropout*, dimensionalités, ...)

# Approche #1 : supervision faible

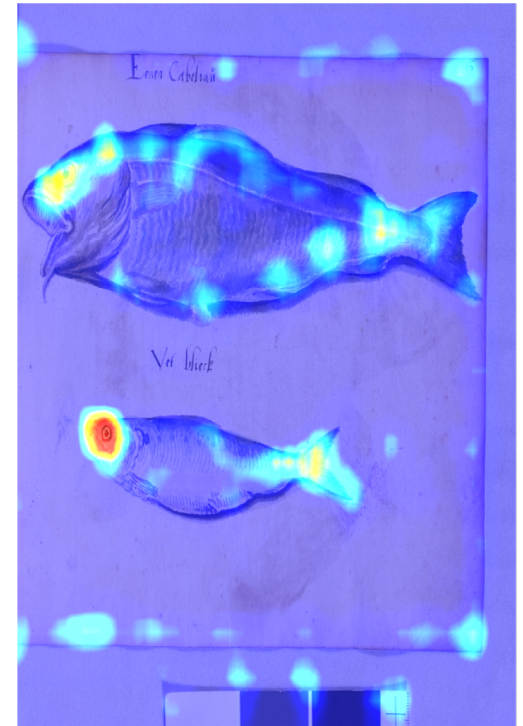


## Excellents résultats

- très bonnes reconnaissances
- parallélisme massif donc très bonnes performances
- trop beau...

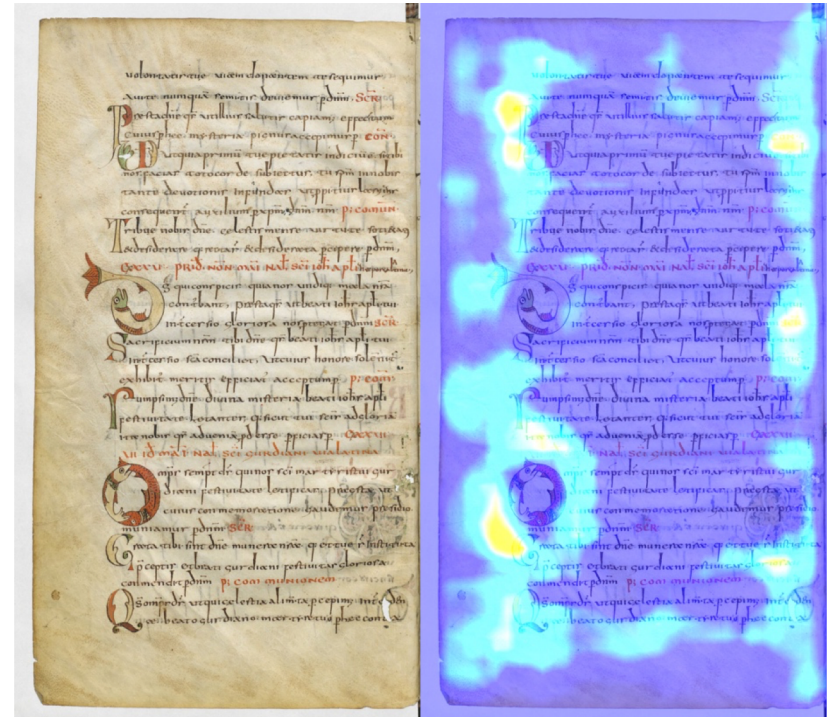
## Examen des cartes d'activation :

- origine classification



# Approche #1 : supervision faible

## Examen des cartes d'activation :





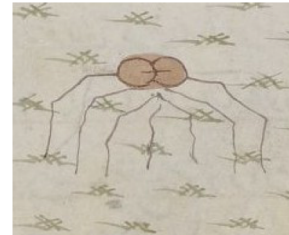
## Conclusions temporaires :

- impossible d'utiliser une solution sur étagère
- supervision trop faible
- images très différentes
- alignement classes très complexe, classes adverses

## Décisions :

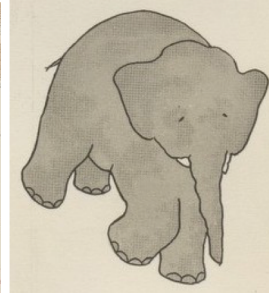
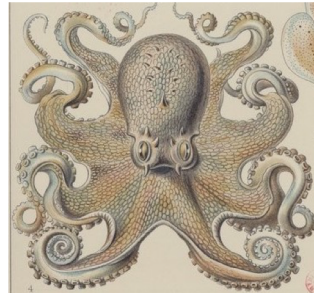
- s'appuyer sur des images plus pertinentes
- augmentation du jeu de données
- augmenter le niveau de supervision
- viser de la détection d'objet, rebondir vers la classification

## iNaturalist

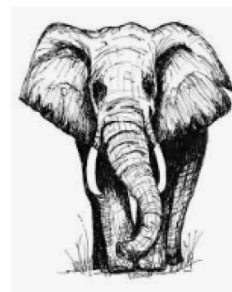
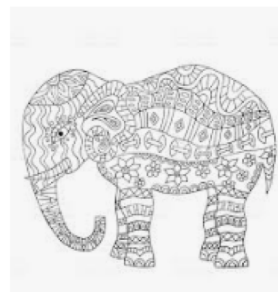
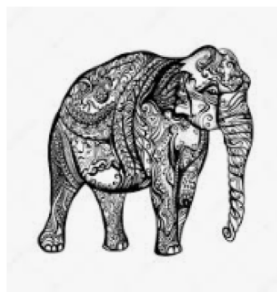
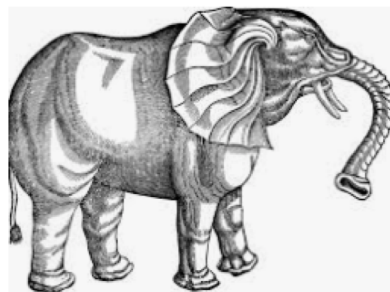


## Dessins, illustrations :

- de la BNF
  - 250+ exemples
  - classes sous représentées



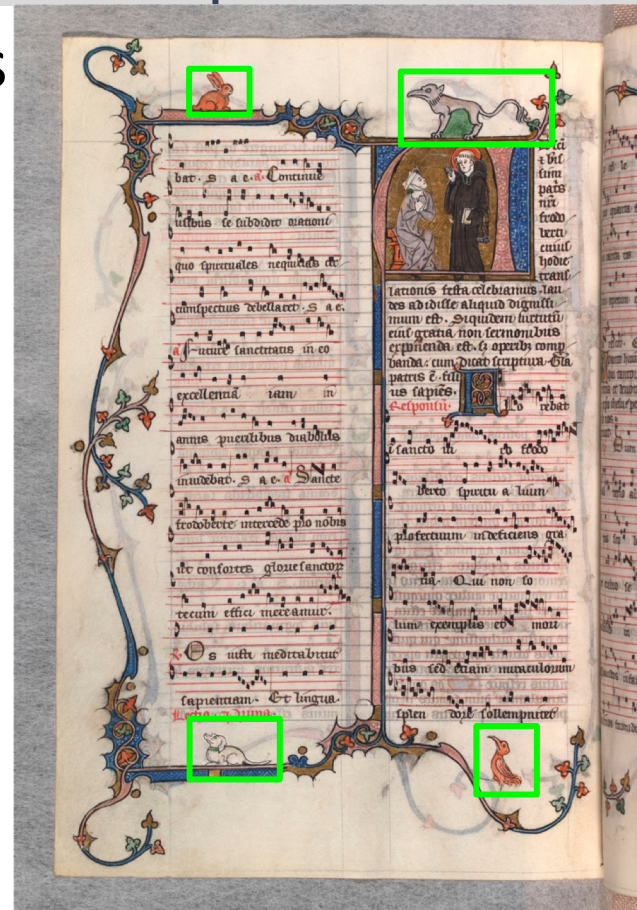
- d'internet



# Augmenter le niveau de supervision

## Ajout manuel de boîtes englobantes

- détourer + attacher annotation
- laborieux
- ambigu (ici : lapin, échassier, chien)



## État de l'art : Faster-RCNN

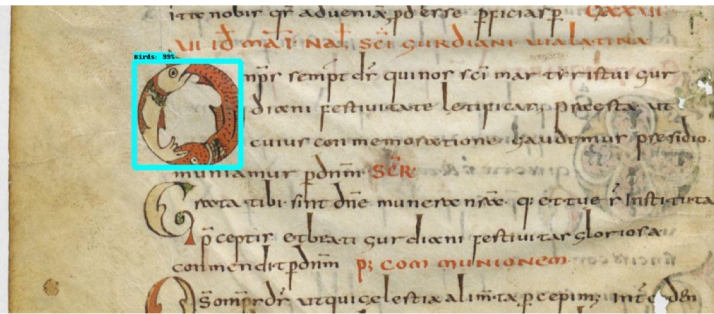
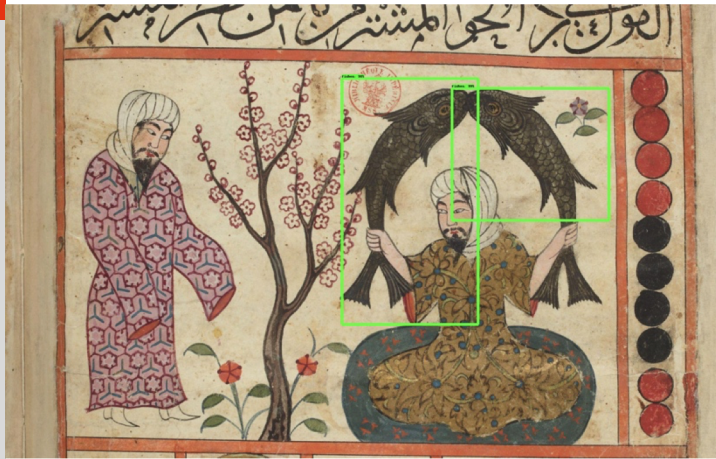
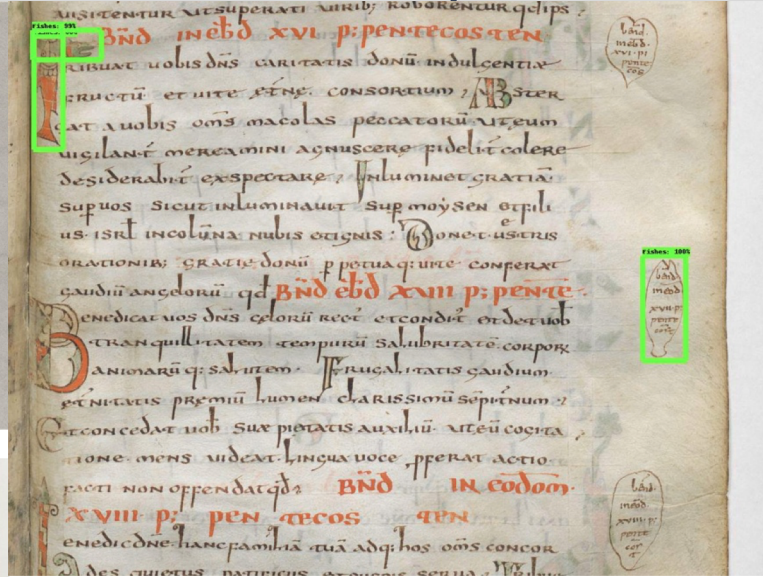
- détection régions candidates
- classification candidats
- ajustement boite englob.



## Application

- supervision manuelle
- 3000 boites créées, 30 classes
- amputation d'iNaturalist : classes d'origine remplacées par classes BNF
- apprentissage et ajustement fin
- tests

# Évaluation CIP avec supervision



## Évaluation qualitative :

- pour le moment très subjective
- difficile
- pas de vérité terrain, annotations niveau page

## Supervision :

- quelle quantité ? Quelle finesse ?
- annoter tout ce qui est présent dans une page
- déséquilibre, classes sur- ou sous-représentées
- boîtes ou masques ?



## Sentiments :

- travail très original
- vraies difficultés, vrai questionnement scientifique

## Difficultés :

- projet très court. Un an passe très vite
- on laisse de côté :
  - autres corpus (Excelsior 1910-1920)
  - définition dynamique de classes
- peu de visibilité sur l'après étude
- définir un cadre pour des projets plus ambitieux ?





## Complémentarité :

- Linkmedia : classification
  - annotation préparatoires : définir classes, donner exemples
  - traitement corpus
- Zénith: apprentissage actif
  - classes dynamiquement définies
  - bouclage de pertinence
  - exemples les plus rentables