



Pattern Spotting and deep learning object detection

Khellouf Leila

Encadrants : Laurent HEUTTE

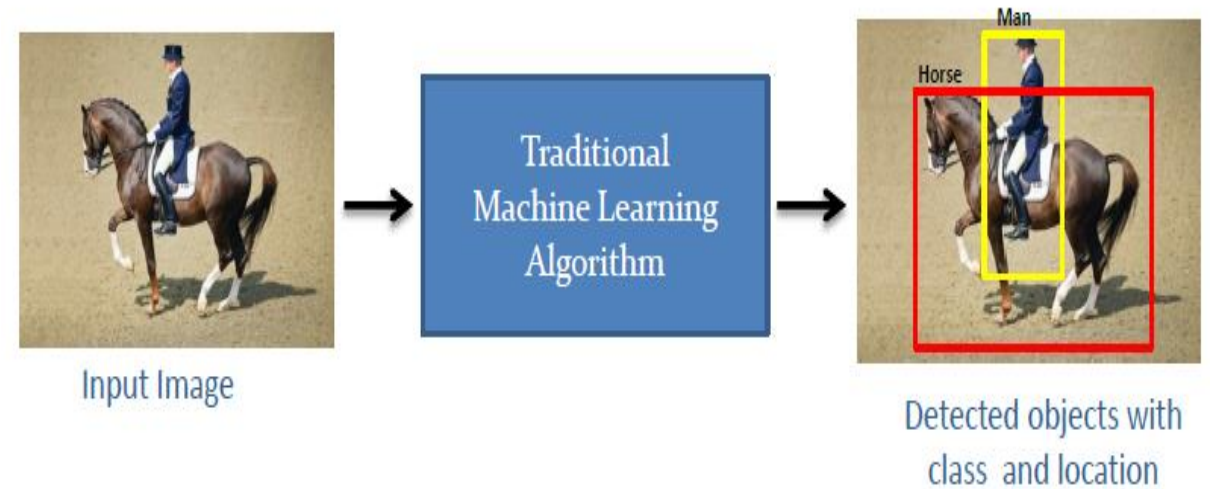
Stéphane Nicolas

L'objectif:

- D'étudier les modèles à l'état de l'art en détection d'objets, en segmentation sémantique, et en segmentation d'instances, dans les images de scènes naturelles
- Appliquer ou bien adapter ces méthodes à la détection d'objets graphiques (ornementations, lettrines, bandeaux,...) dans les images de documents anciens (manuscripts médiévaux).

Détection d'objet:

- On désigne par **détection d'objet** (ou **classification d'objet**) une méthode permettant de détecter la présence d'une instance (**reconnaissance d'objet**) ou d'une *classe d'objets* dans une image numérique



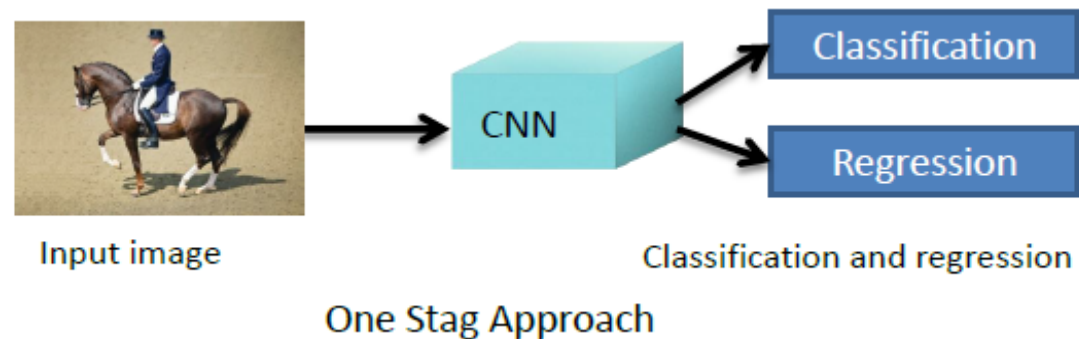
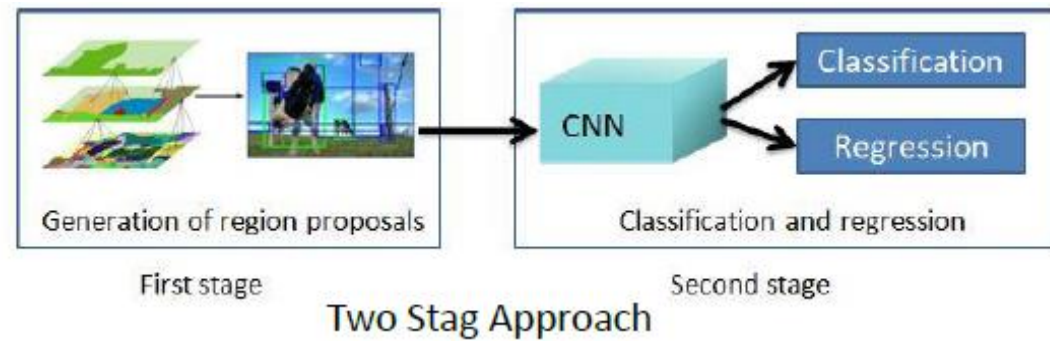
l'état de l'art en détection d'objets:

One stage approach:

- Extraire les features avec la génération de region proposals
 - Effectuer une classification et régression
 - CNN, fast RCNN,

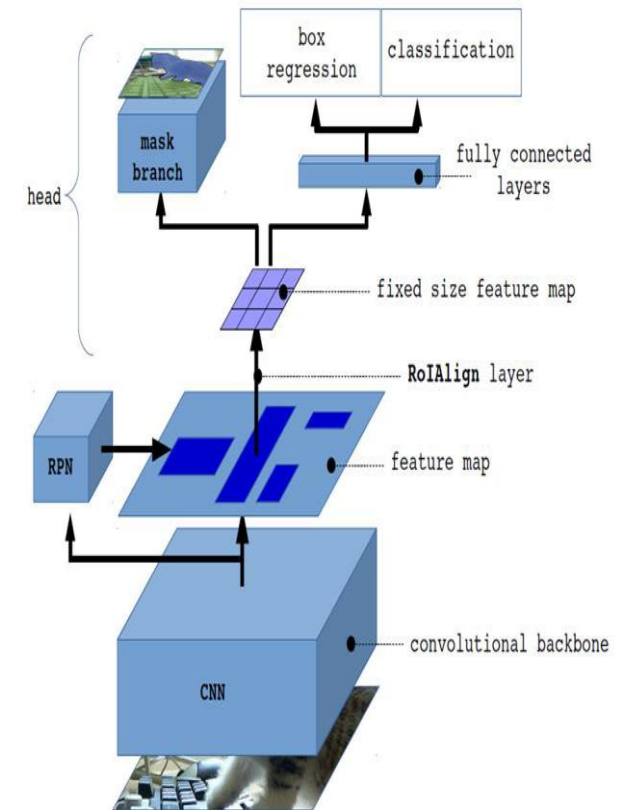
Two stage approach:

- Comme entrée: image en entier
- Effectuer une classification et régression
 - YOLO, RetinaNet , RetinaDet,



Mask RCNN

- **Backbone Model:** On peut utiliser l'architecture ResNet 50 pour l'extraction des features a partir de image
- Ces feature maps sont ensuite transmises au **RPN** qui retourne des boites englobantes candidates.
- Les regions obtenues du RPN pourraient être de différentes formes, On applique **RoIAlign** pour convertir tous les régions à la même forme
- **Fully connect network:** prédire l'etiquette de la classe et les boîtes englobantes .



Mask RCNN on DocExplore:

On a appliqué le Mask RCNN sur DocExplore et on a obtenu les résultats suivants:



Les différentes classes de DocExplore

- **1500 images annotées**
- **1447 requêtes possibles , le nombre d'occurences dans chaque object graphique varie de 2 à plus de 100.**
- **les requêtes sont petites et leurs zones peuvent être aussi basses que 220 (20* 11) pixels .**
- **35 classes**



Losange



Marqueur



Croix

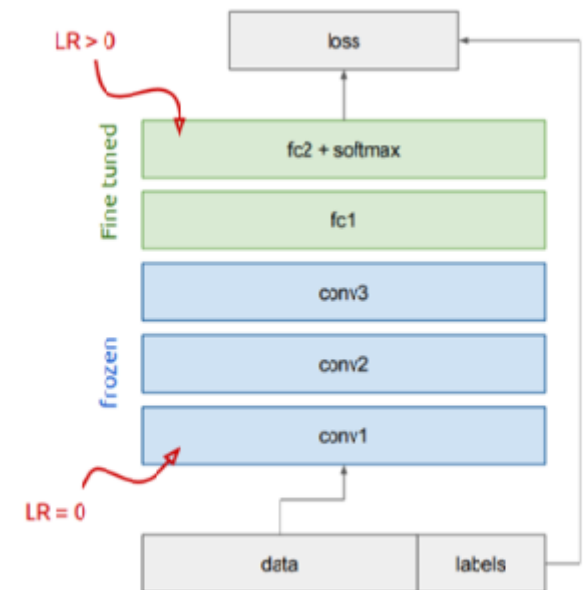


Grand A

Fine Tuning :

Le Fine tuning est un processus qui permet de prendre un modèle de réseau qui a déjà été formé pour une tâche donnée et de lui effectuer une deuxième tâche similaire.

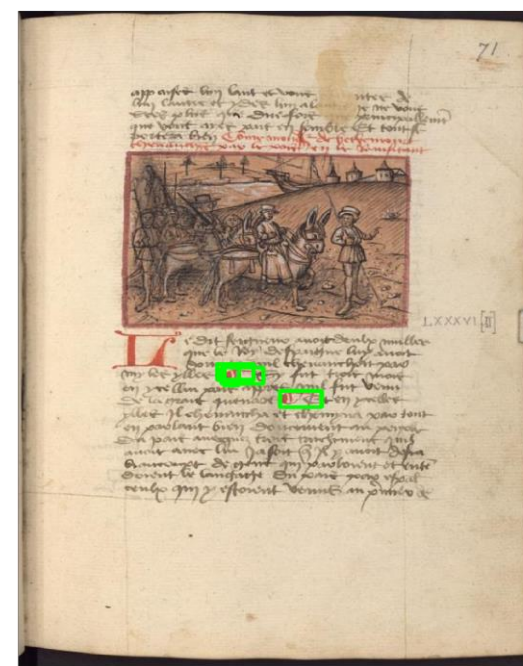
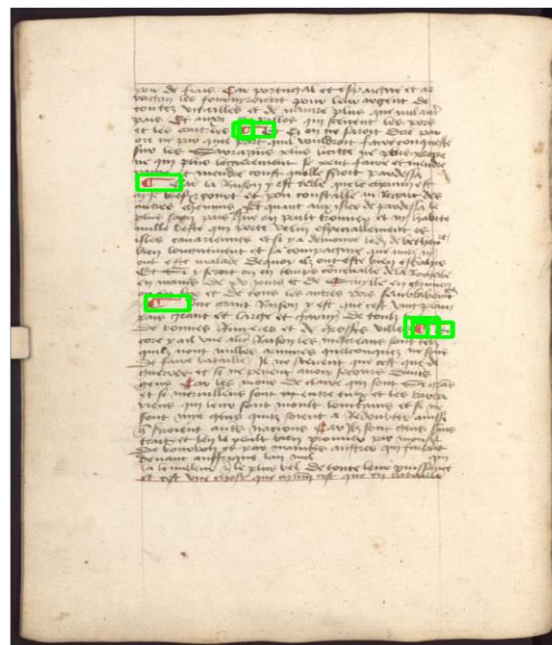
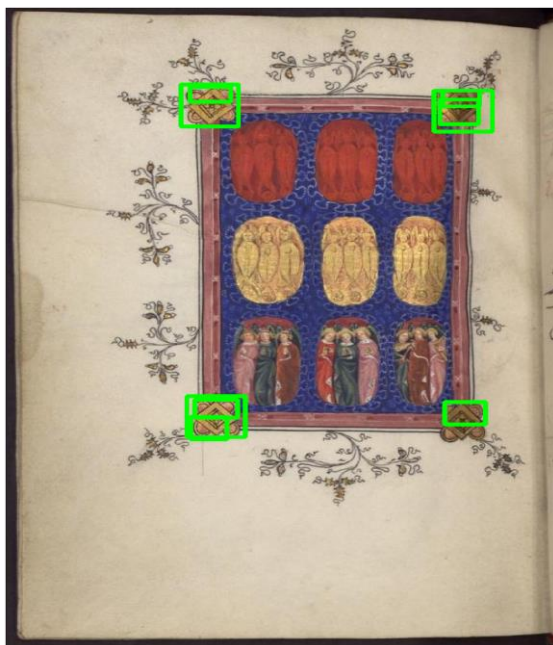
Toute la complexité de création de CNN peut être évitée en adaptant des réseaux pré-entraînés



Fine tuning sur DocExplore:

L'expérience présentée dans l'article a été réalisée en partant du classique maskrcnn-resnet50-fpn entraîné sur la dataset MS COCO . Ensuite ré-entraîné plus finement (Fine Tuning) le modèle pour traiter notre problème de classification des classes « marqueur » et « losange » avec le nombre d'image =161

Résultat sur Test Dataset



Question

