

# You Only Look Once (YOLO)

BEX Roméo, RIVALDI Tristan, LAMURE Maxence

Université de Montpellier  
Master 2 Statistiques et Sciences des Données

2024 - 2025



**STATISTIQUE**  
**SCIENCE DES DONNÉES MIND**  
UNIVERSITÉ DE MONTELLIER

# Table des matières

Introduction

Méthodologie

Comparaison des Méthodes

Limites et Faiblesses

Conclusion

Conclusion

# Introduction

La **classification d'images** consiste à associer des étiquettes à une image en fonction de son contenu visuel. C'est un problème clé de la *vision par ordinateur*, souvent résolu via l'apprentissage supervisé. Applications courantes :

- ▶ **Reconnaissance faciale**
- ▶ **Voitures autonomes**

# Image sans Boîte



Figure: Image d'entrée.

# Image avec Boîtes Englobantes

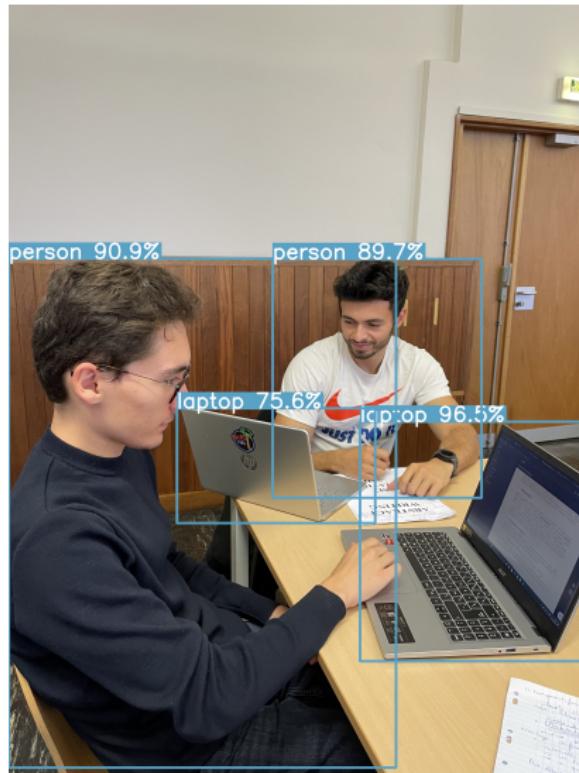


Figure: Image avec boîtes englobantes prédites par YOLOv8.

## Formulation du problème

- ▶ YOLO divise l'image en une grille  $S \times S$ .
- ▶ Chaque cellule prédit des **boîtes englobantes** et des **probabilités de classes**.
- ▶ YOLO utilise une **régression directe**, prédisant en une seule étape les coordonnées des objets et leurs classes, ce qui permet un traitement rapide.

# Structure du réseau YOLO

- ▶ Basé sur un réseau de convolution profond (24 couches + 2 couches entièrement connectées).
- ▶ Traitement de l'image entière en un seul passage.
- ▶ Sortie : tenseur de taille  $S \times S \times (B \times 5 + C)$ .

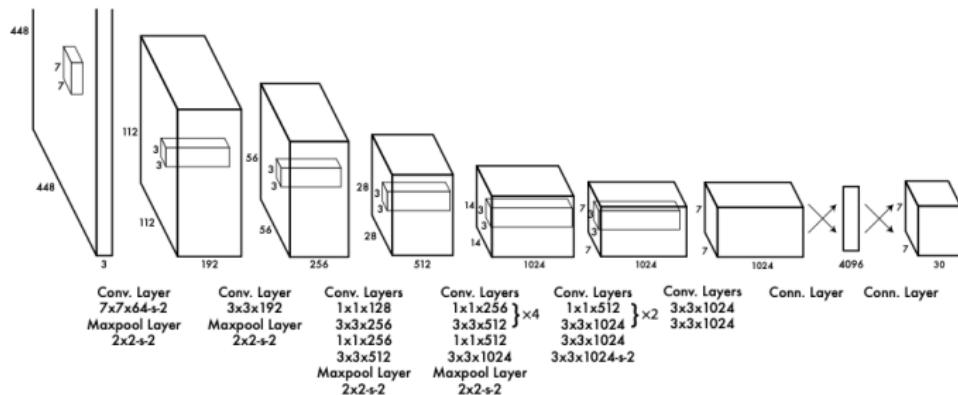


Figure: Architecture du réseau YOLO.

## Fonction de perte

La fonction de perte de YOLO prend en compte les erreurs de localisation et de classification. Elle utilise deux hyperparamètres  $\lambda_{\text{coord}}$  et  $\lambda_{\text{noobj}}$  pour pondérer les différentes parties de la perte :

$$\lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \quad (1)$$

$$+ \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} \left[ (\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \right] \quad (2)$$

$$+ \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \quad (3)$$

Où :

- ▶  $1_{ij}^{\text{obj}}$  est un indicateur qui vaut 1 si l'objet est présent dans la cellule de la grille  $i$  et 0 sinon.
- ▶  $x_i, y_i$  sont les coordonnées du centre de la boîte prédite.

## Fonction de perte (suite)

Suite des explications :

- ▶  $C_i$  est le score de confiance de la boîte.
- ▶  $\lambda_{\text{coord}} = 5$  et  $\lambda_{\text{noobj}} = 0.5$  sont les pondérations des différents termes de la perte.
- ▶ Les termes en  $\lambda_{\text{coord}}$  pondèrent l'erreur de localisation ( coordonnées et dimensions des boîtes).
- ▶ Les termes en  $\lambda_{\text{noobj}}$  réduisent l'impact des fausses prédictions dans les cellules sans objet.

## Comparaison avec d'autres méthodes

- ▶ YOLO est comparé à Faster R-CNN et Fast R-CNN.
- ▶ YOLO est plus rapide avec un bon équilibre entre vitesse et précision.

Méthode	mAP (%)	FPS
Fast R-CNN	70.0	0.5
Faster R-CNN	73.2	7
YOLO	63.4	45
Fast YOLO	52.7	155

Table: Comparaison des méthodes sur PASCAL VOC 2007

# Limites et Faiblesses

- ▶ Problèmes de localisation pour les petits objets.
- ▶ Sensibilité aux petites boîtes, impactant l'IoU (précision des boîtes).
- ▶ Difficulté à généraliser à des objets aux proportions inhabituelles.



Figure: Limites de YOLOv8 sur une scène complexe.

## Limites et Faiblesses - Exemple Supplémentaire



# Conclusion

- ▶ YOLO simplifie les processus classiques de détection d'objets en intégrant toutes les étapes en une seule.
- ▶ Avantages: Rapidité et facilité d'utilisation.
- ▶ Limitations: Localisation imprécise des petits objets.
- ▶ Applications futures: robotique, surveillance, réalité augmentée.

# Conclusion

- ▶ YOLO simplifie la détection d'objets en intégrant toutes les étapes en une seule.
- ▶ Avantages : rapidité, simplicité d'utilisation.
- ▶ Limites : localisation des petits objets.
- ▶ Applications : robotique, surveillance,...

# Questions ?

Merci pour votre attention !  
Des questions ?