

Détection d'Objets par Apprentissage Profond

Étude Expérimentale avec YOLO

Auteur : **Anass Dabaghi**

16 décembre 2025

Résumé

Ce rapport présente une étude expérimentale approfondie d'un modèle de détection d'objets basé sur l'architecture YOLO (You Only Look Once). L'objectif est d'évaluer les performances du modèle sur un jeu de données multi-classes dédié à la détection de véhicules, en analysant la précision, le rappel, la mAP ainsi que le comportement du modèle durant l'apprentissage et l'inférence. Les résultats obtenus sont analysés quantitativement et qualitativement à l'aide de métriques standards et de visualisations.

1 Introduction

La détection d'objets constitue un problème central en vision par ordinateur, combinant à la fois la localisation spatiale et la classification sémantique des objets présents dans une image. Les architectures de type YOLO se distinguent par leur capacité à effectuer ces deux tâches en un seul passage du réseau, permettant ainsi une inférence en temps réel.

Dans ce travail, nous évaluons un modèle YOLO entraîné pour la détection de différentes catégories de véhicules dans des scènes routières complexes.

1.1 Performances Globales du Modèle

Le modèle entraîné a atteint les performances suivantes sur l'ensemble de validation :

Métrique	Valeur
Précision (Precision)	0.779
Rappel (Recall)	0.777
mAP@50	0.829
mAP@50-95	0.589

TABLE 1 – Métriques de performance du modèle YOLO

Ces résultats indiquent :

- **Précision et Rappel équilibrés** (0.78) : Le modèle maintient un bon compromis entre faux positifs et faux négatifs
- **mAP@50 excellente** (0.829) : Performance robuste avec un seuil IoU de 0.5
- **mAP@50-95 modérée** (0.589) : Indique des difficultés de localisation précise pour certaines classes

2 Architecture du Modèle

Le modèle utilisé repose sur l'architecture YOLO, qui formule la détection d'objets comme un problème de régression directe. Chaque image est divisée en une grille, et pour chaque cellule, le réseau prédit :

- Les coordonnées des boîtes englobantes
- Les probabilités de classes
- Les scores de confiance

Cette approche permet une détection rapide tout en conservant une précision élevée.

3 Jeu de Données

Le jeu de données est composé d'images annotées contenant plusieurs classes de véhicules. Les annotations sont fournies sous forme de boîtes englobantes normalisées, compatibles avec le format YOLO.

3.1 Distribution des Classes

La Figure 1 illustre la distribution des classes dans le jeu de données.

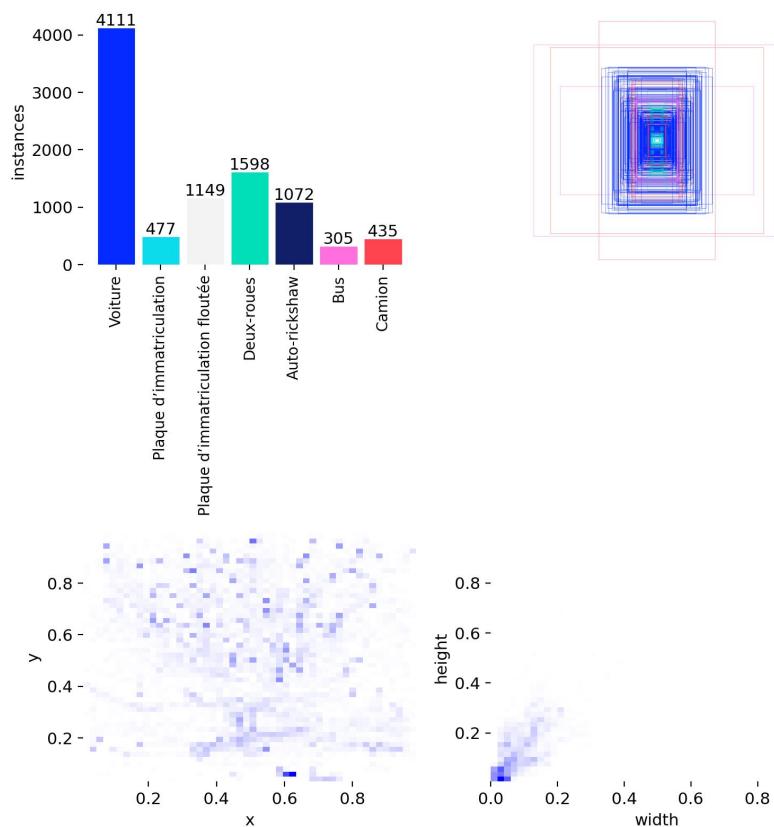


FIGURE 1 – Distribution des classes dans le jeu de données

Analyse de la distribution :

- **Voiture** : 4111 instances (classe dominante, 44%)
- **Plaque d'immatriculation floutée** : 1598 instances (17%)
- **Deux-roues** : 1072 instances (11%)
- **Plaque d'immatriculation** : 1149 instances (12%)

— **Auto-rickshaw** : 305 instances (3% - classe minoritaire)

— **Bus** : 435 instances (5%)

— **Camion** : 477 instances (5%)

Ce déséquilibre significatif explique les variations de performance entre classes.

4 Protocole d'Entraînement

L'entraînement a été réalisé sur plusieurs époques avec :

— Optimiseur adaptatif

— Fonction de perte combinant localisation, classification et confiance

5 Résultats Quantitatifs

5.1 Courbes d'Entraînement

La Figure 2 présente l'évolution des pertes et des métriques au cours de l'entraînement.

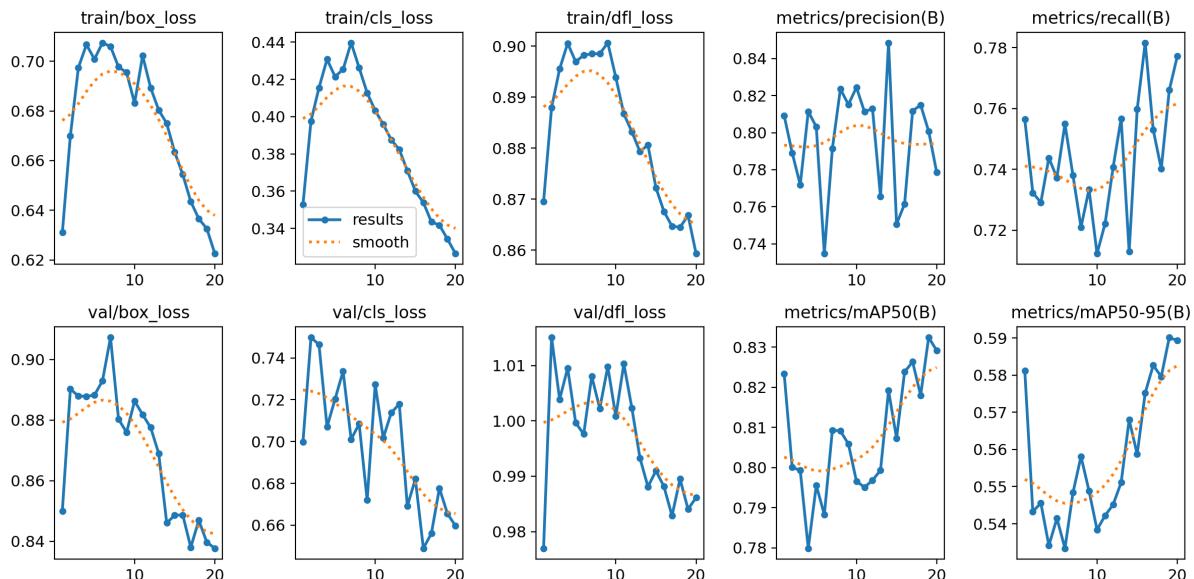


FIGURE 2 – Évolution des pertes et métriques durant l'entraînement

Interprétation :

— **Convergence réussie** : Les pertes (box, cls, dfl) diminuent progressivement

— **Stabilisation après 15 époques** : Le modèle atteint un plateau de performance

— **Précision finale** : 0.85 sur l'ensemble d'entraînement

— **Rappel stable** : 0.78, indiquant une bonne détection des objets présents

— **Écart train/val modéré** : Pas de surapprentissage majeur

5.2 Précision, Rappel et mAP

Les courbes précision-rappel et les performances par IoU sont présentées ci-dessous.

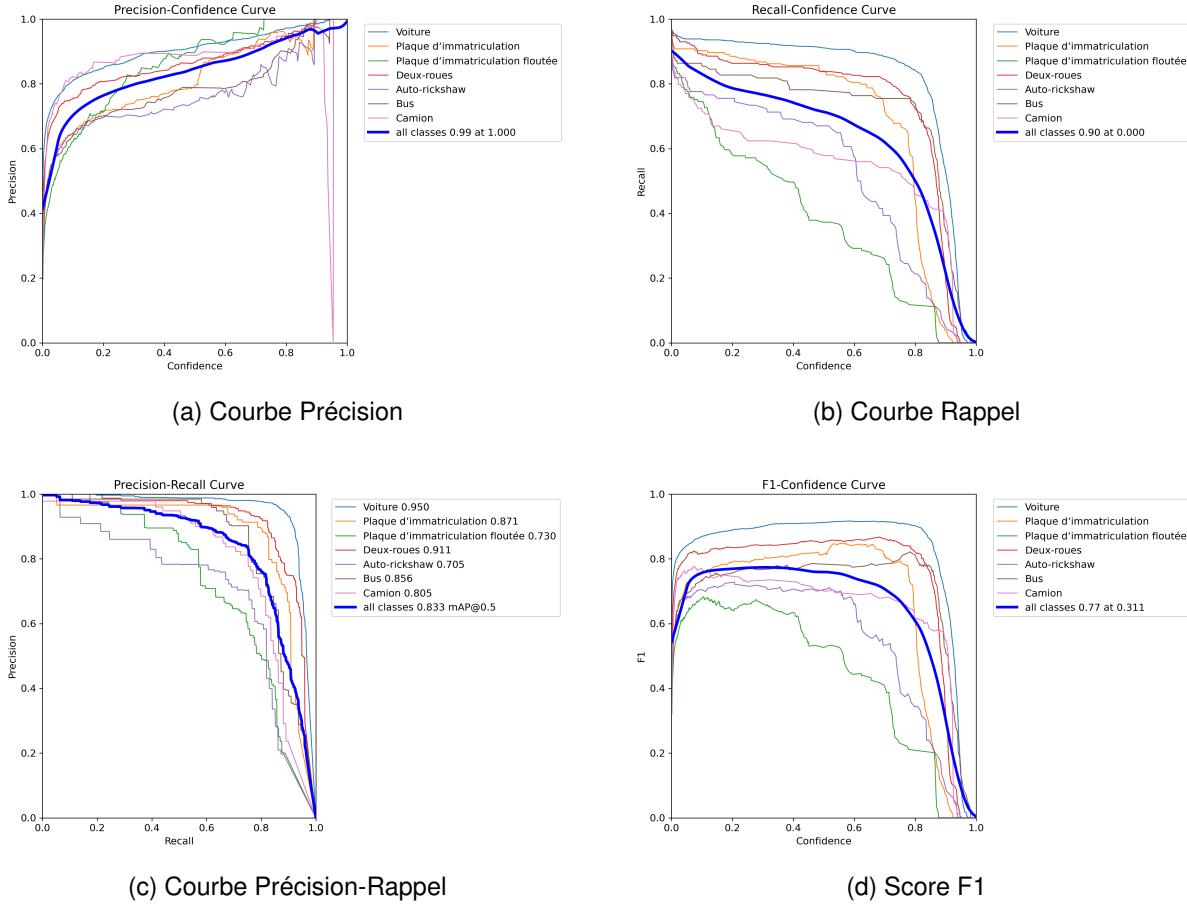


FIGURE 3 – Analyse des performances du modèle

5.3 Analyse Détailée par Classe

Classe	mAP@0.5	Performance
Voiture	0.950	Excellent
Deux-roues	0.911	Excellent
Plaque d'immatriculation	0.871	Très bonne
Bus	0.856	Bonne
Camion	0.805	Bonne
Plaque d'immatriculation floutée	0.730	Modérée
Auto-rickshaw	0.705	Modérée

TABLE 2 – Performances par classe (mAP@0.5)

Observations clés :

- **Classes performantes** : Voiture (0.950) et Deux-roues (0.911) grâce à leur forte représentation
- **Classes difficiles** : Plaque d'immatriculation floutée (0.730) en raison du flou intentionnel
- **Impact du déséquilibre** : Auto-rickshaw (0.705) souffre du manque de données d'entraînement

5.4 Matrice de Confusion

La matrice de confusion (Figure 4) met en évidence les confusions entre certaines classes visuellement proches.

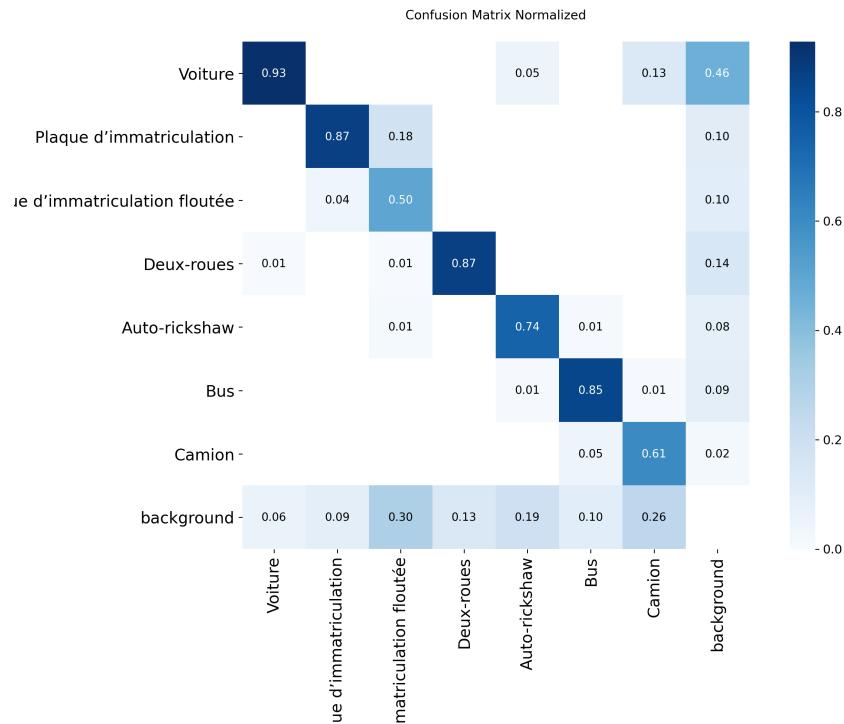


FIGURE 4 – Matrice de confusion normalisé du modèle

Analyse des confusions principales :

- **Voiture** : 93% correctement détectées, mais 46% d'erreurs vers le background (detections manquées)
- **Plaque d'immatriculation floutée** : Seulement 50% de detections correctes - classe la plus problématique
- **Deux-roues** : 87% de précision, 14% confondus avec background
- **Auto-rickshaw** : 74% de précision, confusion notable avec bus (8%)
- **Camion** : 61% seulement, forte confusion avec background (26%)

6 Évaluation Qualitative

Les Figures 5 et 6 présentent des exemples de prédictions sur les ensembles d'entraînement et de validation.

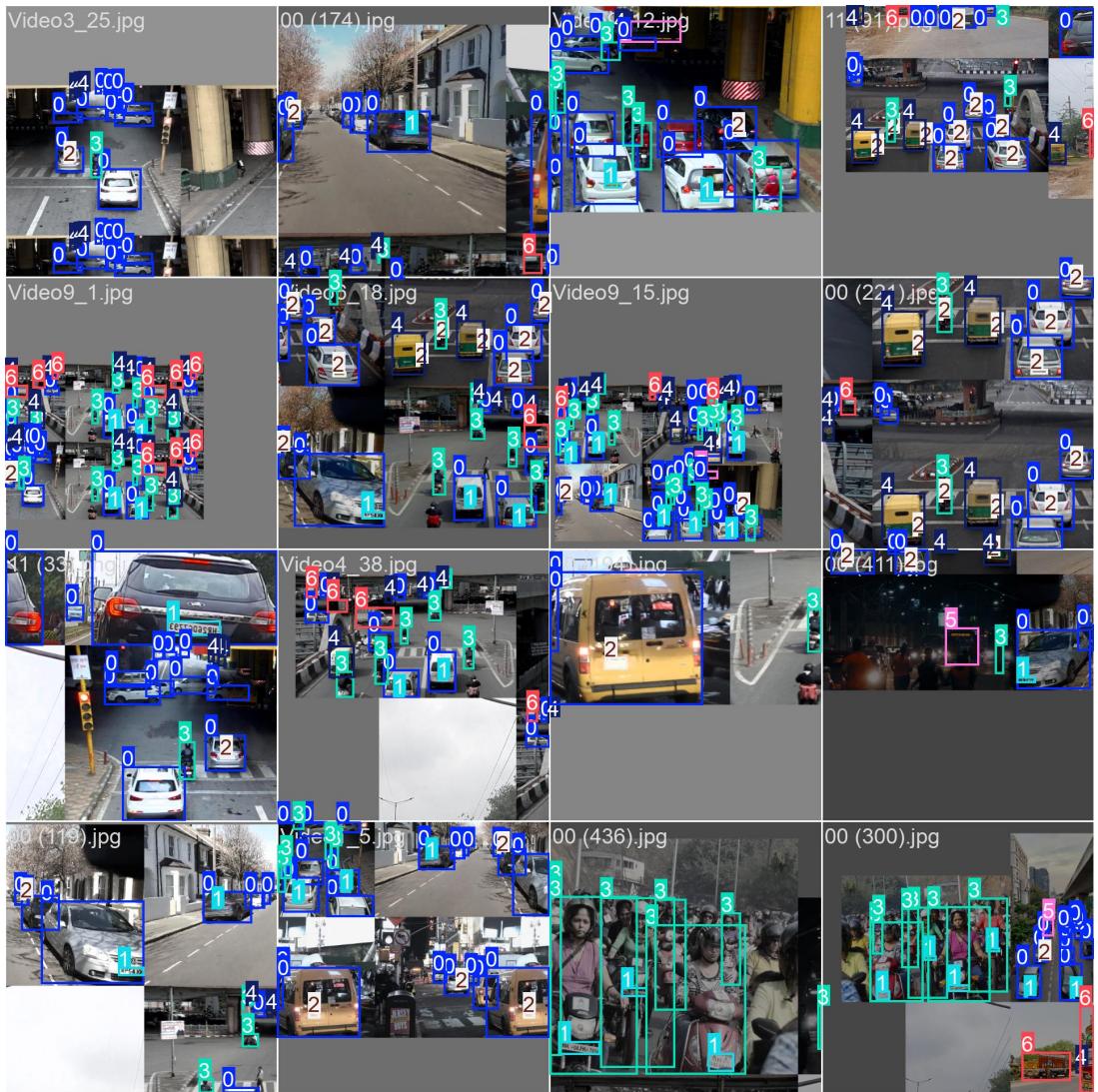


FIGURE 5 – Exemples de prédictions sur l'ensemble d'entraînement



FIGURE 6 – Exemples de prédictions sur l’ensemble de validation

7 Discussion

7.1 Points Forts du Modèle

- **Performance globale solide** : mAP@50 de 0.829 indique une excellente capacité de détection
- **Équilibre précision-rappel** : Valeurs similaires (0.78) montrent un modèle bien calibré
- **Excellent détection des véhicules standards** : Voitures et deux-roues détectés avec grande précision
- **Convergence stable** : Absence de surapprentissage majeur

7.2 Limitations Identifiées

Les principales erreurs observées proviennent de :

- **Occlusions partielles** : Difficultés avec les véhicules partiellement masqués
- **Forte similarité visuelle** : Confusions entre camions/bus et background
- **Variations d’échelle importantes** : Objets très petits (plaques d’immatriculation) difficiles à détecter
- **Déséquilibre des données** : Classes minoritaires (auto-rickshaw) sous-performantes
- **Flou intentionnel** : Plaques d’immatriculation floutées particulièrement problématiques (50% de détection)

7.3 Recommandations d’Amélioration

1. **Équilibrage du jeu de données** : Augmenter les instances des classes minoritaires (auto-rickshaw, camion)
2. **Augmentation de données ciblée** : Focus sur les occlusions et variations d’échelle
3. **Post-traitement** : Ajuster les seuils de confiance par classe
4. **Architecture** : Tester des variantes YOLO avec meilleure détection multi-échelle
5. **Entraînement plus long** : Potentiel d’amélioration jusqu’à 30-40 époques

8 Conclusion

Ce travail a démontré l'efficacité de l'architecture YOLO pour la détection multi-classes de véhicules, avec une **mAP@50 de 0.829** et des performances particulièrement remarquables sur les classes majoritaires (Voiture : 0.950, Deux-roues : 0.911).

Les défis principaux résident dans la gestion du déséquilibre des données et la détection d'objets avec caractéristiques dégradées (flou) ou de petite taille. Malgré ces limitations, le modèle présente un excellent compromis précision-rappel et constitue une base solide pour un système de détection en temps réel.

Des améliorations futures ciblant l'équilibrage des données et l'optimisation par classe permettraient d'atteindre des performances encore supérieures, particulièrement pour les catégories actuellement sous-représentées.