

Analyse des résultats de nos 3 modèles

Hugo Dehurtevent

Juin 2024

1 Mise en contexte

Dans cette expérience, nous avons utilisés 3 modèles différents avec 2 bases d'images différentes pour voir les différentes évolutions de nos modèles. Dans notre cas, nous allons utiliser les modèles suivants :

1. Faster R-CNN
2. RetinaNet
3. FCOS

Pour les bases d'images nous allons utiliser les suivantes :

1. BDD100K
2. COCO

2 Format de l'expérience

Pour cette expérience, nous avons décider d'utiliser les modèles pré-entraîné. Les mesures que nous allons analyser sur ces modèles sont les suivantes :

- Precision
- Recall
- mAP

La précision représente la proportion de détection correcte pour tous les positifs. Nous pouvons le calculer comme ceci : $\frac{TP}{TP+FP}$.

Le Recall représente la proportion de vrai détection positive parmi tous les vrai positif. Nous pouvons le calculer comme ceci : $\frac{TP}{TP+FN}$

La précision moyenne (mAP) est une métrique utilisée pour évaluer les modèles de détection d'objets. La moyenne des valeurs de précision moyenne (AP) est calculée sur les valeurs de rappel de 0 à 1 par rapport à toutes les classes.

Notre objectif est d'avoir un modèle avec le plus de précision et de recall possible.

Pour cela, nous allons d'abord évaluer les modèles simplement sur les différents datasets. Puis nous allons faire des fusions de modèles, c'est à dire que pour l'instant nous allons faire une fusion simple sans traitement derrière, une autre en faisant un traitement Non Max Suppression (NMS).

3 Tableau des résultats

Table 2 Tableau de résultats pour 100 images pour chaque dataset.

Model	BDD100K			COCO		
	precision	recall	mAP	precision	recall	mAP
Faster R-CNN	0.6900	0.5331	14.2811	0.5881	0.7857	53.3767
RetinaNet	0.9028	0.3176	10.9700	0.8415	0.5504	54.7104
FCOS	0.8225	0.3962	12.5330	0.6908	0.6541	45.4001
Faster R-CNN, RetinaNet Cat	0.6615	0.5348	10.8924	0.5479	0.7927	48.0539
Faster R-CNN, RetinaNet NMS	0.6797	0.5336	10.5462	0.5840	0.7885	46.2454
Faster R-CNN, FCOS Cat	0.6270	0.5462	11.5580	0.4898	0.8095	47.4916
Faster R-CNN, FCOS NMS	0.6588	0.5405	10.9920	0.5564	0.8081	46.7206
RetinaNet, FCOS Cat	0.7742	0.4105	8.7381	0.6310	0.6779	41.1140
RetinaNet, FCOS NMS	0.8100	0.4082	9.5651	0.6814	0.6709	42.1277
Faster R-CNN, RetinaNet, FCOS Cat	0.6187	0.5468	11.2171	0.4861	0.8081	47.2838
Faster R-CNN, RetinaNet, FCOS NMS	0.6522	0.5410	10.9920	0.5516	0.8081	46.7206

TABLE 1 – Résultats pour 100 images.

3.1 Analyse des résultats

Observons d’abord les combinaisons de modèles, on peut voir qu’ils permettent généralement d’améliorer légèrement le recall par rapport aux modèles combinée mais cela diminue la précision du modèle.

— Comparaison des modèles individuels

Pour BDD100K, RetinaNet a la plus haute précision (0.9028) mais un recall très faible (0.3176), indiquant qu’il détecte moins de vrais positifs. FCOS a une meilleure balance entre précision (0.8225) et recall (0.3962) par rapport à RetinaNet. Faster R-CNN présente un recall le plus élevé (0.5331) parmi les trois, mais une précision plus modeste (0.6900).

Pour COCO, RetinaNet a également la plus haute précision (0.8415), mais avec un recall modéré (0.5504). Faster R-CNN a un excellent recall (0.7857) mais une précision plus basse (0.5881). FCOS est équilibré avec une précision (0.6908) et un recall (0.6541).

— Comparaison des modèles combinés

Les modèles combinés montrent que l’usage de NMS (Non-Maximum Suppression) ou de la concaténation des résultats (Cat) peut légèrement améliorer le rappel par rapport à l’usage d’un seul modèle. Cependant, la précision a tendance à diminuer.

— mAP

Le mAP sur COCO est nettement plus élevé que sur BDD100K pour tous les modèles, ce qui est normal car ils sont entraînés par cette base d’image. La différence ce fait aussi sur le fait que COCO connaît beaucoup plus de classe (80) et que BDD100K en connaît 10 sur les images que nous avons utilisées.

Table 2 Tableau de résultats pour les datasets entiers de BDD100K et COCO.

Model	BDD100K			COCO		
	precision	recall	mAP	precision	recall	mAP
Faster R-CNN	0.6916	0.5301	30.8229	0.5862	0.7593	53.4689
RetinaNet	0.8983	0.3008	19.7121	0.8550	0.5186	50.8616
FCOS	0.7884	0.3775	27.4476	0.7011	0.6212	49.9437
Faster R-CNN, RetinaNet Cat	0.6640	0.5345	26.4853	0.5512	0.7665	51.4885
Faster R-CNN, RetinaNet NMS	0.6847	0.5317	26.8720	0.5840	0.7621	53.1047
Faster R-CNN, FCOS Cat	0.6169	0.5438	26.7680	0.4938	0.7813	54.2192
Faster R-CNN, FCOS NMS	0.6601	0.5383	27.0701	0.5576	0.7750	53.1047
RetinaNet, FCOS Cat	0.7453	0.3962	19.5895	0.6485	0.6507	46.9224
RetinaNet, FCOS NMS	0.7783	0.3896	21.1537	0.7009	0.6420	47.1866
Faster R-CNN, RetinaNet, FCOS Cat	0.6115	0.5447	26.7680	0.4920	0.7824	54.2192
Faster R-CNN, RetinaNet, FCOS NMS	0.6544	0.5393	27.0701	0.5541	0.7770	53.1047

TABLE 2 – Résultats des différents modèles pour toutes les données.

3.2 Analyse des résultats de la Table 2

Observons d’abord les combinaisons de modèles, on peut voir qu’ils permettent généralement d’améliorer légèrement le recall par rapport aux modèles combinée mais cela diminue la précision du modèle.

— Comparaison des modèles individuels

Pour BDD100K, RetinaNet a la plus haute précision (0.8983) mais un rappel très faible (0.3008), ce qui indique qu’il détecte moins de vrais positifs. FCOS a une meilleure balance entre précision (0.7884) et rappel (0.3775) par rapport à RetinaNet. Faster R-CNN présente un rappel le plus élevé (0.5301) parmi les trois, mais une précision plus modeste (0.6916).

Pour COCO, RetinaNet a également la plus haute précision (0.8550), mais avec un rappel modéré (0.5186). Faster R-CNN a un excellent rappel (0.7593) mais une précision plus basse (0.5862). FCOS est équilibré avec une précision (0.7011) et un rappel (0.6212).

— Comparaison des modèles combinés

Les modèles combinés montrent que l’usage de NMS (Non-Maximum Suppression) ou de la concaténation des résultats (Cat) peut légèrement améliorer le rappel par rapport à l’usage d’un seul modèle. Cependant, la précision a tendance à diminuer.

— mAP

Le mAP sur COCO est nettement plus élevé que sur BDD100K pour tous les modèles, ce qui est normal car ils sont entraînés par cette base d’image. La différence ce fait aussi sur le fait que COCO connaît beaucoup plus de classe (80) et que BDD100K en connaît 10 sur les images que nous avons utilisées.

— Conclusion

Précision vs Rappel : Un modèle avec une haute précision mais un faible rappel (comme RetinaNet sur BDD100K) est très bon pour éviter les faux positifs mais manque de détection de nombreux vrais positifs. À l'inverse, un modèle avec un haut rappel mais une précision modérée (comme Faster R-CNN sur COCO) détecte la majorité des vrais positifs mais génère plus de faux positifs.

Combinaison de modèles : Combiner des modèles peut améliorer les métriques globales, mais il y a un compromis à faire entre précision et rappel. L'usage de techniques comme NMS permet d'atténuer les duplications de détections.