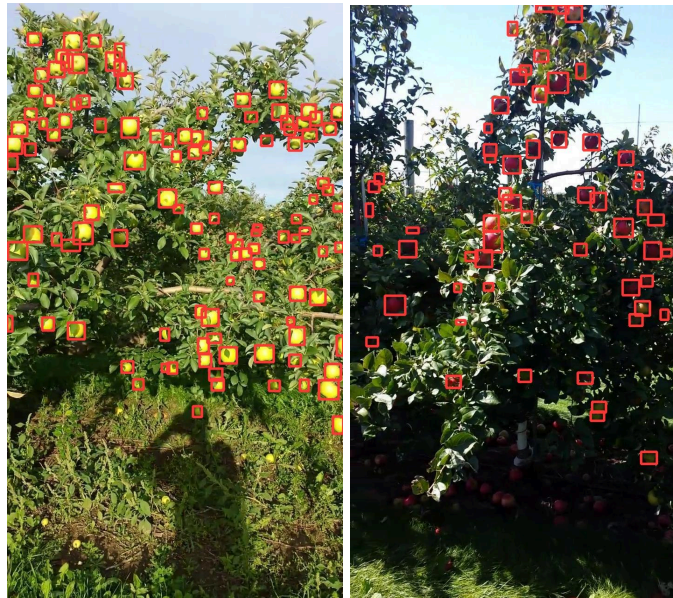


Séance 3 - Transfert learning or Foundation model ?

Herearii Metuarea

L'objectif de ce TP est de savoir si un modèle de fondation est mieux qu'un modèle supervisé. Une grande partie du TP sera fait sur un notebook de Google Colab. Un compte rendu individuel (avec figures et analyse) devra être déposé jusqu'au 17 Février 2025. Chaque exercice de ce TP constitue les parties de votre compte rendu.

Le sujet de ce TP est d'étudier la question suivante : **Est-ce qu'un modèle de fondation est meilleur qu'un modèle supervisé réentraîné ?** Nous répondrons à cette question en comparant les performances du modèle de fondation DINO et du modèle de détection Yolo (version 11) avec transfert learning. Notre étude repose sur la détection de pommes. Nous utiliserons la base de données MinneApple. Le code notebook travaillé dans cette séance peut être réutilisé.



Echantillon de la base de données MinneApple.

1- Introduction

Rappeler le sujet de votre TP dans le compte rendu.

2- Description de la base de données

Combien d'images contient la base de données ?

Présenter deux échantillons issus de cette base.

Décrivez la qualité de vos images : est-ce que la luminosité est équilibrée ? quelles sont les tailles de vos images ? est-ce que la taille des cellules sont visuellement équilibrées ?

3- Train, Validation, Test

Combien d'images contient votre train, validation et test ?

4- Modèles de détection

Expliquer en une phrase l'architecture YOLO (version 11) et le modèle DINO. Quelle est la différence entre les deux modèles ?

5- Métrique et fonction de perte

Définir la métrique et la fonction de perte utilisées. Justifier votre choix.

Mentionner votre méthode de calcul du gradient. Si adam choisi, mentionner la valeur fixée pour le taux d'apprentissage ou learning rate.

6- Méthode d'apprentissage

Comme indiqué dans le titre de ce TP, nous allons seulement réentraîner le modèle Yolo. En effet, on rappelle que l'objet de ce TP est de comparer les performances d'un modèle réentraîner et d'un modèle de fondation. À partir des poids du modèle entraîné sur les images COCO, vous allez réentraîner le modèle à partir des images MineApple. Or, nous avons vu une technique complémentaire pour garantir une meilleure performance de détection: la data augmentation. On rappelle que la data augmentation génère synthétiquement de nouvelle image par opération mathématique. Elle permet de combler des manques d'hétérogénéités dans votre base de données, mais aussi d'augmenter la taille de votre base de données. Lorsque vous regardez la taille de votre jeu de données, pensez-vous avoir suffisamment d'images pour entraîner votre modèle? Pensez-vous utiliser la data augmentation? Si vous utilisez la data augmentation, appliquer seulement un flip horizontal et un brightness contrast.

7- Résultat

- a) Faire l'inférence des images test sur le modèle Yolov11. Quel est le score de détection ?
- b) Réentraîner le modèle Yolov11 avec les images MinneApple.
 - i) Quelle est la qualité de votre apprentissage ? Y-a-t-il eu du surapprentissage, du sous-apprentissage ?
- c) Faire l'inférence des images test sur votre nouveau modèle entraîné.
- d) Faire l'inférence en générant les prédictions des images test sur le modèle DINO en utilisant DeepDataSpace.
- e) Point bonus : résultat quantitatif des images test du modèle DINO

8- Analyse

Analyser visuellement les résultats à partir des images prédits avec Yolo réentraînée : que remarquez-vous ?

Comparer visuellement les performances du modèle Yolo réentraînée et du modèle DINO.

9- Conclusion

Conclure en répondant à la question du TP : **Est-ce qu'un modèle de fondation est meilleur qu'un modèle supervisé réentraîné ?**