

Projet 7 : Développez une preuve de concept

Objectif : Améliorer les performances de MobileNetV3 en traitant les données déséquilibrées

Plan de travail prévisionnel

1. Exploration du jeu de données
 - ☒ Sélection du jeu de données (CWD30 - images et étiquettes de classification)
 - ☒ Chargement des données
 - ☒ Analyse exploratoire des données (E.D.A.)
 - ☒ Visualisation des données déséquilibrées
2. Préparation des données
 - ☒ Prétraitement des données
 - ☒ Division des données en ensembles d'apprentissage, de validation et de test
3. Modèle MobileNetV3Small - Baseline
 - ☒ Chargement du modèle pré-entraîné MobileNetV3
 - ☒ Fine-tuning du modèle avec l'ensemble d'apprentissage
 - ☒ Recherche des hyperparamètres les plus performants pour ce modèle
 - ☒ Évaluation des performances sur l'ensemble de validation
4. Pondération des échantillons lors de l'entraînement
 - ☒ Gestion du déséquilibre des classes par pondération
 - ☒ Fine-tuning du modèle avec l'ensemble d'apprentissage
5. Sous-échantillonnage des classes majoritaires
 - ☒ Technique de rééquilibrage par sous-échantillonnage
 - ☒ Fine-tuning du modèle avec l'ensemble d'apprentissage
6. Implémentation de DenseWeight & LMFLOSS
 - ☒ Définition du modèle avec DenseWeight pour la croissance des plantes et LMFLOSS pour les espèces
 - ☒ Fine-tuning du modèle avec l'ensemble d'apprentissage
 - ☒ Recherche des hyperparamètres les plus performants pour ce modèle
 - ☒ Évaluation des performances sur l'ensemble de validation
7. Comparaison des performances
 - ☒ Comparaison des performances entre les différents modèles
 - ☒ Analyse des résultats obtenus
8. Conclusion
 - ☒ Rédaction d'un rapport détaillé sur les résultats, les méthodes utilisées et les conclusions tirées
 - ☒ Documentation du code et des expériences réalisées pour faciliter la réutilisation et la compréhension du projet

Sources bibliographiques :

- CWD30: A Comprehensive and Holistic Dataset for Crop Weed Recognition in Precision Agriculture : <https://arxiv.org/pdf/2305.10084.pdf> (<https://arxiv.org/pdf/2305.10084.pdf>)
- MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications : <https://arxiv.org/pdf/1704.04861.pdf> (<https://arxiv.org/pdf/1704.04861.pdf>)
- Searching for MobileNetV3 : <https://arxiv.org/pdf/1905.02244.pdf> (<https://arxiv.org/pdf/1905.02244.pdf>)
- LMFLOSS: A Hybrid Loss For Imbalanced Medical Image Classification : <https://arxiv.org/pdf/2212.12741.pdf> (<https://arxiv.org/pdf/2212.12741.pdf>)
- Focal Loss for Dense Object Detection : <https://arxiv.org/pdf/1708.02002.pdf> (<https://arxiv.org/pdf/1708.02002.pdf>)
- Learning Imbalanced Datasets with Label-Distribution-Aware Margin Loss : <https://arxiv.org/pdf/1906.07413.pdf> (<https://arxiv.org/pdf/1906.07413.pdf>)
- Density-based weighting for imbalanced regression : <https://link.springer.com/article/10.1007/s10994-021-06023-5> (<https://link.springer.com/article/10.1007/s10994-021-06023-5>)
- Review of Image Classification Algorithms Based on Convolutional Neural Networks : <https://www.mdpi.com/2072-4292/13/22/4712> (<https://www.mdpi.com/2072-4292/13/22/4712>)
- Label-Imbalanced and Group-Sensitive Classification under Overparameterization : <https://arxiv.org/abs/2103.01550> (<https://arxiv.org/abs/2103.01550>)