

# Identification des maladies foliaires du colza : analyse d'image vs les experts

Nicolas Parisey

UMR 1349 IGEPP - Département SPE - INRAe

Décembre 2022

# Équipe Démécologie, Axe "Imagerie Quantitative"

## Qui

Lydia Bousset (CR), Melen Leclerc (CR), Marcellino Parlerme (AI),  
Magali Ermel (TR), Stéphane Jumel (TR)

## Pourquoi

- ▶ Détection, classification, caractérisation dont :
  - ▶ Quantification de sévérité, de risques ;
  - ▶ Analyses morphométriques ;
  - ▶ Mesures fonctionnelles ;
- ▶ Pour abonder à de l'épidémiologie
  - ▶ Construire des estimateurs de mesures d'intérêt ;
  - ▶ Assimiler dans des modèles dynamiques, spatio-temporelles.

# Identification de maladies foliaires

Questions : Malade? De quoi? A quel point?



# Identification de maladies foliaires

## Réponses

- ▶ Images de loin, de près, en séquence ... ;
- ▶ Position, température, humidité, pression, ...
- ▶ ITK, bassin de production, ...

$$P(\text{Disease}_i | F(\text{Img}_i), \\ G(X_i, Y_i), \\ H(\text{Cultiv}_i), \\ \dots))$$



# Identification de maladies foliaires

## Réponses

- ▶ Images de loin, de près, en séquence ... ;
- ▶ Position, température, humidité, pression, ...
- ▶ ITK, bassin de production, ...

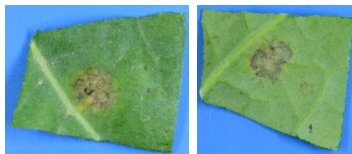
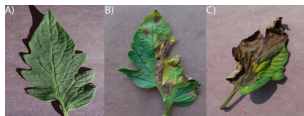
$$P(\text{Disease}_i | F(\text{Img}_i), \\ G(X_i, Y_i), \\ H(\text{Cultiv}_i), \\ \dots))$$



# Identification de maladies foliaires

## En Conditions Contrôlées [CASDAR ATIPICAL]

- ▶ Le plus cité, PlantVillage<sup>1</sup> : 54303 images de feuilles saines et malades réparties en 38 catégories par espèce et par maladie ;
- ▶ Le nôtre, "NoName"<sup>2</sup> : 12836 images de fragments de feuilles saines et malades réparties en 7 catégories attaquant une espèce (le colza).



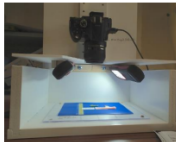
1. (Hugues & Salathe, 2015), cité min. 613 fois directement

2. (Bousset et al, in prep)

# Identification de maladies foliaires

Situation simplifiée : tester l'identification sans la recherche de zones d'intérêt, et avec des images standardisées

1. Collecte de feuilles de colza, découpe de fragments porteurs d'une seule lésion. Fragments imagés avec éclairage, fonds et code barres



Jeu de données équilibré sur 7 classes :

- 6 maladies foliaires d'intérêt :

*Alternaria*

*L. biglobosa*

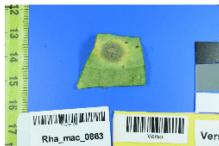
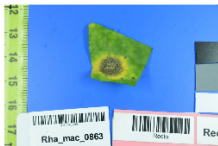
*L. maculans*

Mildiou

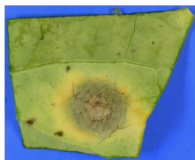
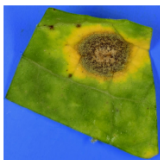
*Mycosphaerella*

- fragments asymptotiques

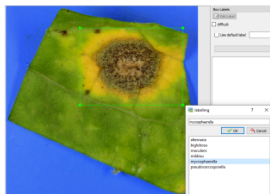
2. Images d'origine Recto et verso



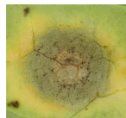
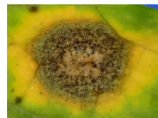
3. Recentré sur le fragment de feuille



4. Annotation des boîtes dans LabelImg



5. ROI : recadrage sur les boîtes, recto et verso



6. Deep learning pour l'identification

# Identification de maladies foliaires

## Déjà des questions intéressantes à cette étape

1. Peut il y avoir des biais? <sup>3</sup>
2. Qu'est ce qui se passe si on retourne la feuille?
3. Qu'est ce qui se passe si même les experts ne savent pas?



# Identification de maladies foliaires

## Préambule

- ▶ Une bonne acquisition c'est déjà 80% du boulot

# Identification de maladies foliaires

## Préambule

- ▶ Une bonne acquisition c'est déjà 80% du boulot
- ▶ On doit prévoir la généralisation
  - ▶ Transfer Learning<sup>4</sup> → ImageNet, PlantCLEF, PlantVillage,...
  - ▶ Multi-Task Learning<sup>4</sup> → Fusion conditionnel
  - ▶ Etc<sup>4</sup> →  $\geq 16$  types de méthodes

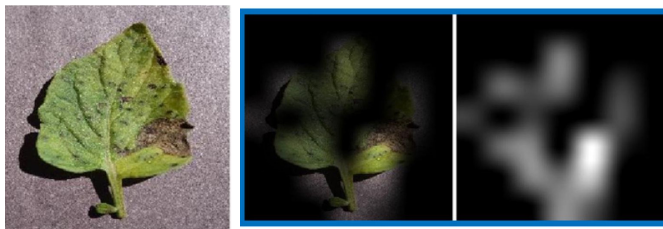
Learning paradigm	Training data	Test data	Condition	Test access
Multi-task learning	$S^1, \dots, S^n$	$S^1, \dots, S^n$	$Y^i \neq Y^j, 1 \leq i \neq j \leq n$	✓
Transfer learning	$S^{src}, S^{tar}$	$S^{tar}$	$Y^{src} \neq Y^{tar}$	✓
Domain adaptation	$S^{src}, S^{tar}$	$S^{tar}$	$P(\mathcal{X}^{src}) \neq P(\mathcal{X}^{tar})$	✓
Meta-learning	$S^1, \dots, S^n$	$S^{n+1}$	$Y^i \neq Y^j, 1 \leq i \neq j \leq n+1$	✓
Lifelong learning	$S^1, \dots, S^n$	$S^1, \dots, S^n$	$S^i$ arrives sequentially	✓
Zero-shot learning	$S^1, \dots, S^n$	$S^{n+1}$	$Y^{n+1} \neq Y^i, 1 \leq i \leq n$	×
Domain generalization	$S^1, \dots, S^n$	$S^{n+1}$	$P(S^i) \neq P(S^j), 1 \leq i \neq j \leq n+1$	×



# Identification de maladies foliaires

Explainable AI<sup>5</sup>

(1) Peut il y avoir des biais ?<sup>6</sup>



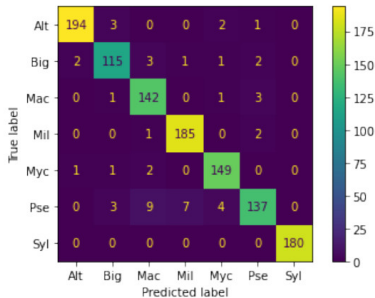
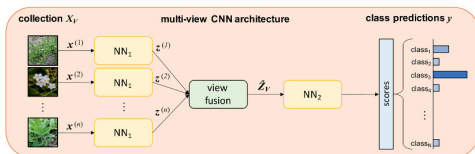
Prédiction  $\sim$  symptôme | espèce (nervures)

5. (Yosinski et al, 2015 ; Selvaraju et al, 2016))

6. (Lee et al, 2020)

# Identification de maladies foliaires

(2) Qu'est ce qui se passe si on retourne la feuille? <sup>7</sup>



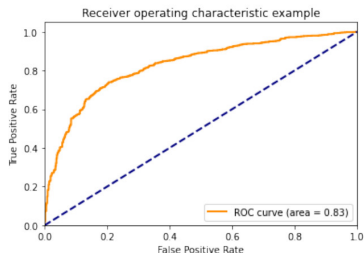
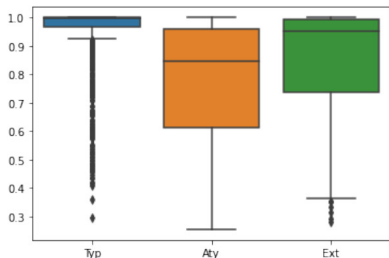
Validation Accuracy : 0.93 (Mono)  $\rightarrow$  0.97 (Multi)

# Identification de maladies foliaires

## (3) Qu'est ce qui se passe si même les experts ne savent pas ?

- ▶ Subset : Vote majoritaire multi-experts Versus QPCR
  - ▶ accuracy 'typique' 0.88 (+0.04 vs best expert)
  - ▶ accuracy 'atypique' 0.36 (-0.07 vs best expert)

Et si on reprend le réseau précédent en mode 'Novelty Detection' ?<sup>8</sup>



# Identification de maladies foliaires

## Conclusions & Perspectives

- ▶ En conditions contrôlés : premiers résultats très encourageants

---

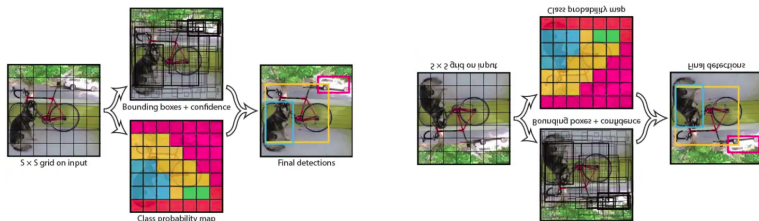
9. "leaf disease" en 2022, COMPAG 44 articles, EcolInf 12 articles, IEEE Explore 120 articles, ...

10. (Redmon et al 2015 ; Ren et al, 2015)

# Identification de maladies foliaires

## Conclusions & Perspectives

- ▶ En conditions contrôlés : premiers résultats très encourageants
- ▶ Du labo au champ : beaucoup de choix<sup>9</sup> e.g. 'ligne de base' par fusion tardive de modèle classique<sup>10</sup>



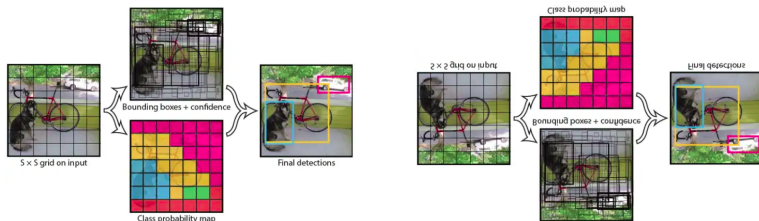
9. "leaf disease" en 2022, COMPAG 44 articles, EcolInf 12 articles, IEEE Explore 120 articles, ...

10. (Redmon et al 2015 ; Ren et al, 2015)

# Identification de maladies foliaires

## Conclusions & Perspectives

- ▶ En conditions contrôlés : premiers résultats très encourageants
- ▶ Du labo au champ : beaucoup de choix<sup>9</sup> e.g. 'ligne de base' par fusion tardive de modèle classique<sup>10</sup>



- ▶ Continuer la récolte, préparer les suites. . .

9. "leaf disease" en 2022, COMPAG 44 articles, EcolInf 12 articles, IEEE Explore 120 articles, . . .

10. (Redmon et al 2015 ; Ren et al, 2015)



## ...dont le PEPR PI@ntAgroEco

Responsables : **Alexis Joly** (INRIA, établ. porteur) & **Pierre Bonnet** (CIRAD)

2023 – 2027 "Nouvelles perspectives sur la caractérisation des maladies des plantes"

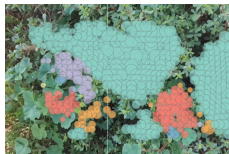
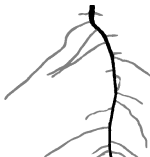
WP1 New image-based AI models for agro-ecology

WP2 Knowledge aggregation and data production

WP3 Integration in PI@ntNet infrastructure and maintenance

WP4 Interactions with other platforms and initiatives

WP5 Citizen science programs, animation and communication



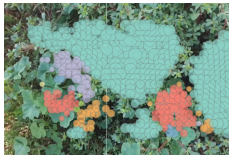
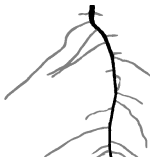
## ...dont le PEPR PI@ntAgroEco

Responsables : **Alexis Joly** (INRIA, établ. porteur) & **Pierre Bonnet** (CIRAD)

2023 – 2027 "Nouvelles perspectives sur la caractérisation des maladies des plantes"

### WP1 New image-based AI models for agro-ecology

- ▶ Partners : INRAE (IGEPP [Lead.], SAVE (JM Armand), LISAH (F Vinatier)), Inria (ZENITH), CIRAD (AMAP), UM (IMAG), IRD (DIADE), RHS, Sw
  1. Plant disease recognition (AMAP)
  2. Estimation of symptom severity (IGEPP)
  3. Infra-specific identification (ZENITH)
  4. Species ID in multi-specimen images (LISAH)



Merci beaucoup pour votre  
attention !