



Machine-learning assisted phenotyping: From fungal morphology to mode of action hypothesis

Sarah Laroui, Eric Debreuve, Xavier Descombes, François Villalba, Florent Villiers, Aurélia Vernay

► To cite this version:

Sarah Laroui, Eric Debreuve, Xavier Descombes, François Villalba, Florent Villiers, et al.. Machine-learning assisted phenotyping: From fungal morphology to mode of action hypothesis. IUPAC - 14th International Congress of Crop Protection Chemistry, May 2019, Ghent, Belgium. hal-02403936

HAL Id: hal-02403936

<https://hal.archives-ouvertes.fr/hal-02403936>

Submitted on 11 Dec 2019

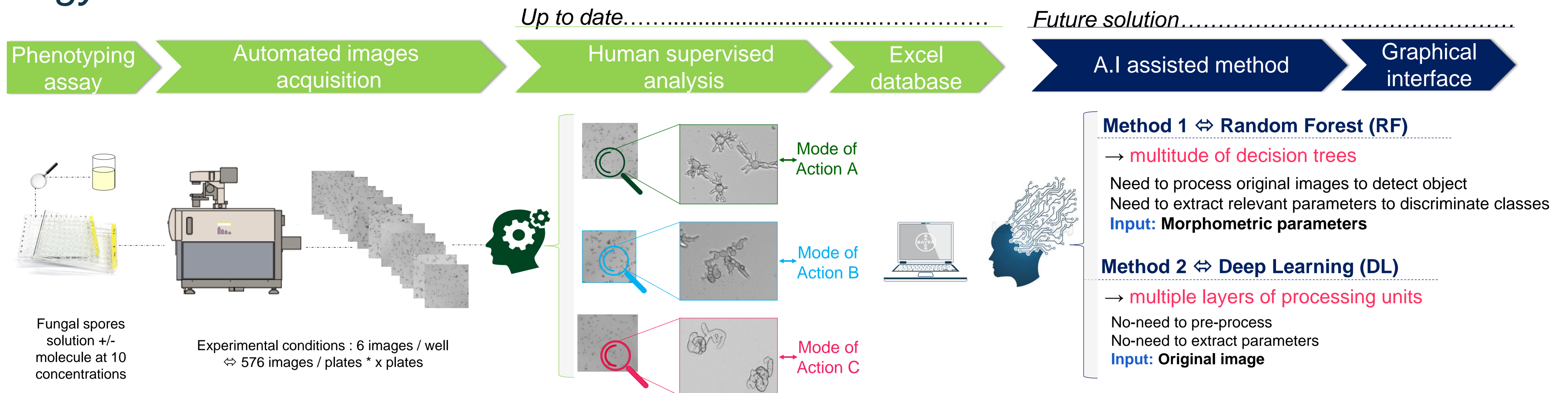
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Introduction

Beyond growth inhibition, fungicides can also trigger specific morphological modifications visualized under transmitted light microscopy. These morphological changes result from the activity of a given compound *via* the inhibition of a molecular target, commonly named as its mode of action (MoA). We are hence able to classify different molecules into their respective MoA by observing their phenotypic signature, and even to detect new MoA with unknown phenotypic effect for further deconvolution. The aim of the presented work is to develop a robust method for automated recognition and classification of these phenotypic signatures in order to lead to a Mode of Action hypothesis. We compare two machine-learning methods (Random forest and Convolutional Neural Network) for direct processing of images generated on the grey mold *Botrytis cinerea* subjected to different antifungal molecules.

Strategy



*From morphology recognition to Mode of Action prediction :
 the power of artificial intelligence to automatize fungal precision phenotyping.*

Method 1

Morphometric parameters extraction and Random Forest-based classification

- Object detection and individualization**
 OBJECT DETECTION SOBEL GRADIENT → IMAGE THRESHOLD & BINARY IMAGE CREATION → BINARY MASK OPTIMIZATION → FIND CONTOUR (from OPT MASK) → FINAL MASK + CELL BY CELL LABEL
- Skelitization and graph generation**
 Extra-branches created → Skeletonization → Graph generation
- Morphometric parameters computation***
 *To be further developed
 From all object :
 • **GLOBAL parameters:** skeleton length variance, skeleton total length, object number, gradient threshold.
 • **SPECIFIC parameters:** weighted length, skeleton length, distance on skeleton mean, distance on skeleton variance, object area, number of nodes, mean branch distance, median branch distance, variance branch distance, length of the longest branch, number of small branch.
- Random Forest for phenotypes clustering**
 (1) Training set : trees generation on **known** phenotypes object
 (2) Test set : trees application on **sample** phenotypes object ⇔ Prediction
 (3) Classifier efficiency evaluation : comparison A.I prediction vs. human expertise

To date, n = 500 trees with 15 features

Sample = 1 object

Morphometric parameter 1 > α
 Morphometric parameter 2 > β
 Morphometric parameter 3 > γ

Predicted class: 1 2 3 ... n

In summary :
 Precise analysis and classification of **cellular objects** within images.
 Access to **known and defined** morphometric features values.
 Limited number of morphometric features.
 Run time: **few minutes** / image

Method 2

Application of a pre-trained image recognition Convolutional Neural Network (CNN) for phenotype classification

- Image pre-processing**
 Data augmentation using image cropping in order to ensure analysis in a given pre-defined neural network.
 TIFF → PNG → Grid
- CNN for phenotypes clustering**
 (1) Pre-trained deep neural network choice. Mobilenet provides best results in comparison to Inception or ggNet.
 (2) Training of a new output layer while keeping all previous layers frozen: this last layer ensures custom classification (4 phenotypes) from parameters extracted in the hidden layers.
 (3) Test & classifier efficiency evaluation : comparison A.I prediction vs. human expertise

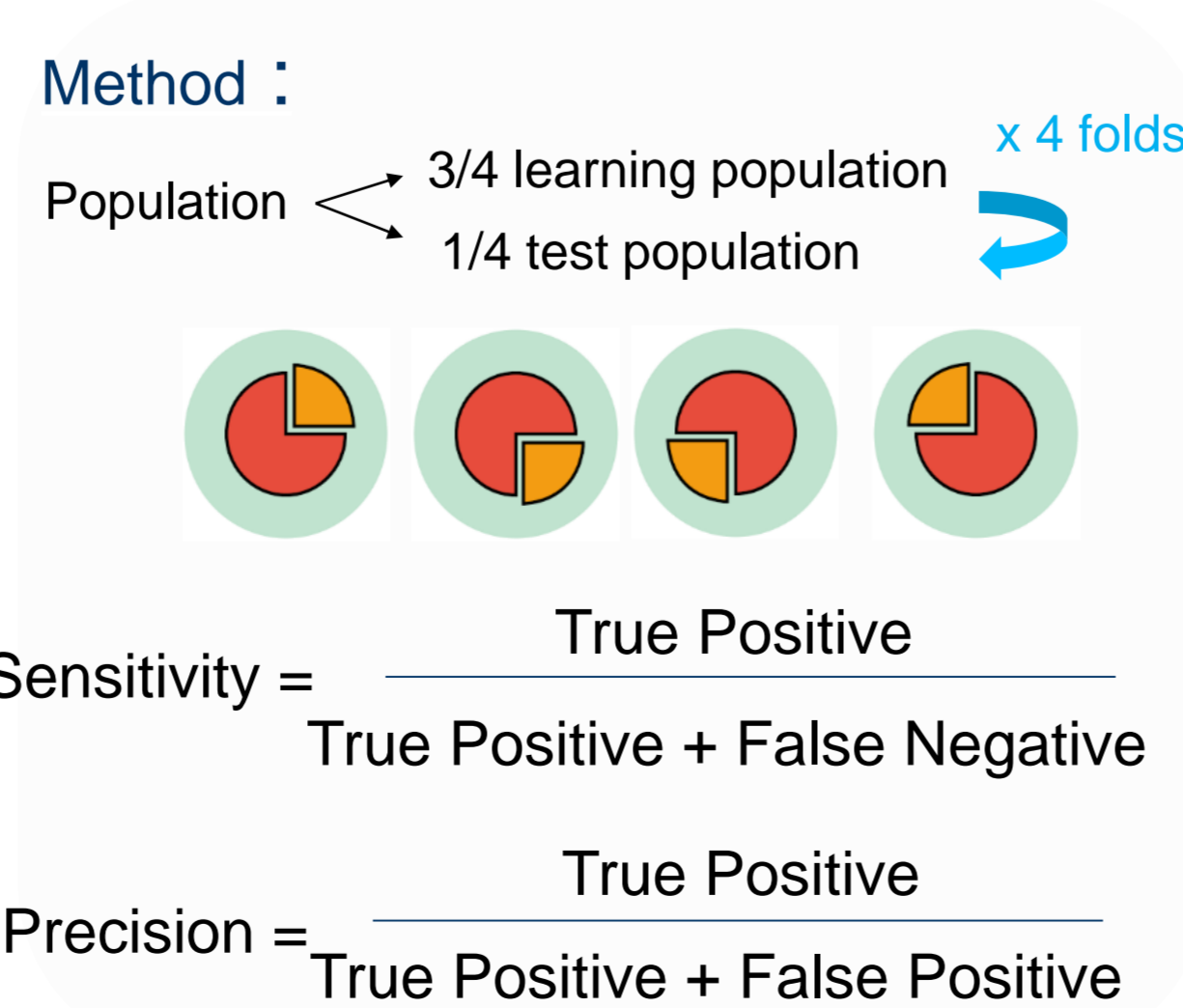
In summary :
 Global analysis and object classification of the **entire image**.
 Access to **unknown features** (morphometric or others ?)
 High number of features (~million)
 Run time: **few seconds** / image

Results & Conclusions

Cross-validation

Evaluation on 4 different phenotypes
 Nbr molecule / phenotype : Ph 1(X), Ph 2 (x), Ph 3(X) and Ph 4 (X)
 Nbr images / phenotype : Ph 1(162), Ph 2 (180), Ph 3(297) and Ph 4 (414)

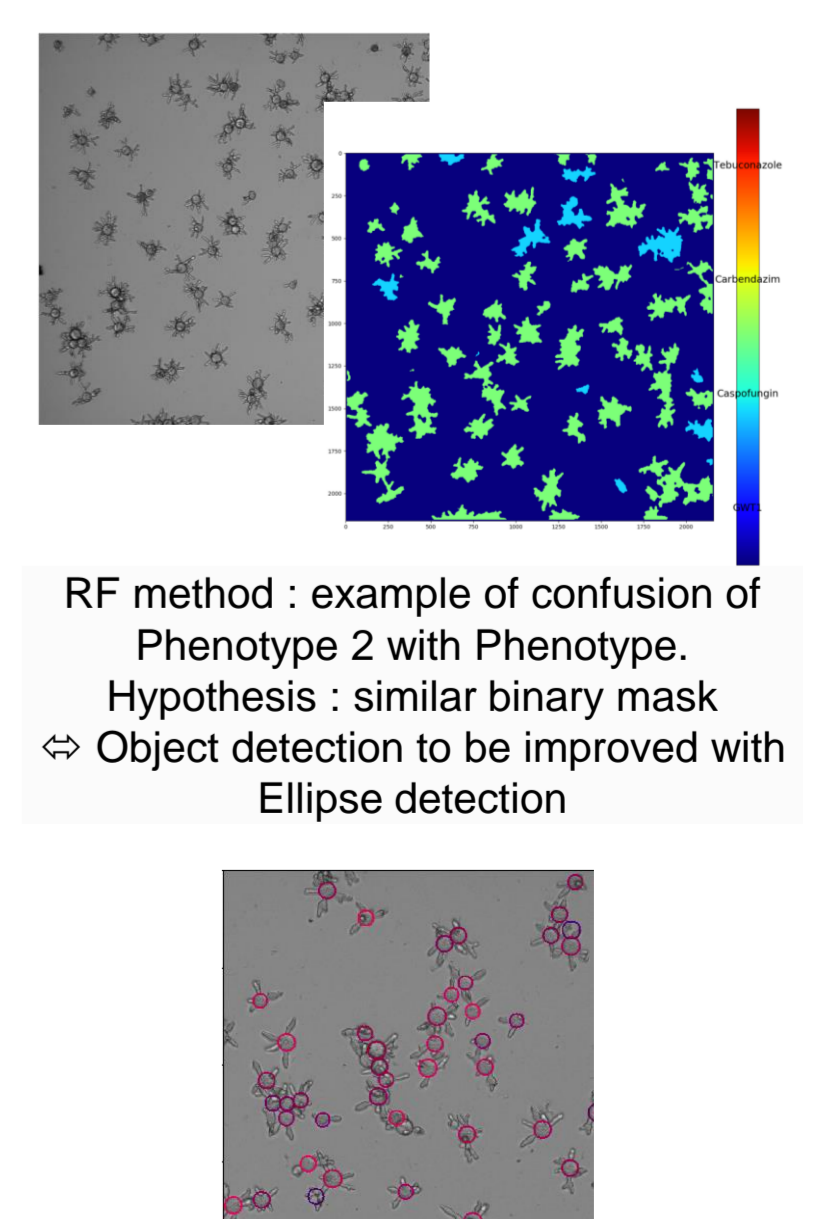
		Fold 1	Fold 2	Fold 3	Fold 4
RF n = 500 trees	Nbr of molecule with correct prediction	19 / 23	20 / 23	17 / 23	14 / 22
	Sensitivity (%)	77			
	Precision (%)	79			
CNN n = 500 training steps	Nbr of molecule with correct prediction	23 / 23	21 / 23	22 / 23	21 / 22
	Sensitivity (%)	92			
	Precision (%)	95			



Robustness intra-experiments

Evaluation on 4 different phenotypes. Nbr of images / experiments
 E1 : Ph 1(99), Ph 2 (99), Ph 3 (189) and Ph 4 (243)
 E2 : Ph 1(63), Ph 2 (81), Ph 3 (99) and Ph 4 (171)

		Training set E1, Test E2	Training set E2, Test E1
RF n = 500 trees	Nbr of molecule with right prediction	15 / 35	17 / 57
	Sensitivity (%)	37	
	Precision (%)	38	
CNN n = 500 training steps	Nbr of molecule with right prediction	31 / 35	54 / 57
	Sensitivity (%)	92	
	Precision (%)	91	



To date, CNN provides better phenotype classification than Random-Forest based method. We are going to further investigate respective approaches in order to evaluate which one, or a combination, will permit to ensure a robust automation of phenotyping recognition in a larger number of phenotype.