Data augmentation for pathogen segmentation in vinewood fluorescence microscopy images

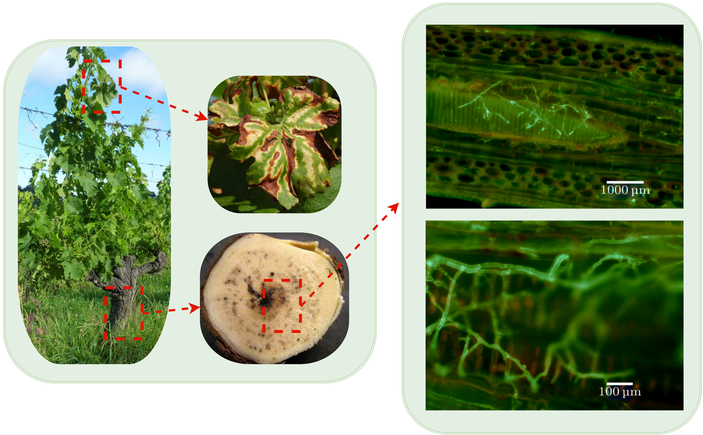
Julie Munsch [1,2], Sonia Ouali [2], Jean-Baptiste Courbot [2], Romain Pierron[3] and Olivier Haeberlé [1]

[1] Eiffage Energie Systèmes - centre expertise IA, Mulhouse, France
[2] IRIMAS, UR 7499, Université de Haute-Alsace, Mulhouse, France
[3] LVBE, UR 3991, Université de Haute-Alsace, Colmar, France
julie.munsch@uha.fr

MOTIVATION-PROBLEMS

Grapevine trunk diseases pose a major problem for vinegrowers worldwide.

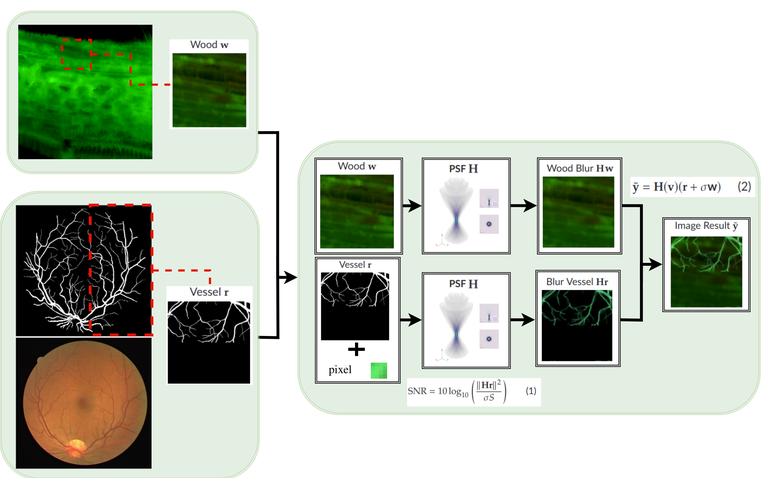
- **13%** of vineyards are unproductive each year in France, resulting in important commercial losses [1].
- We are studying one of the oldest diseases: **Esca**.
- **No treatment exists** once the wood is infected.
- To study the behavior of pathogens, which is still **poorly understood**, we make **inoculation** experiments and use **fluorescence microscopy**.



Getting this type of image is challenging, time-consuming, and our collection is limited.

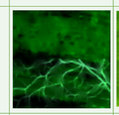
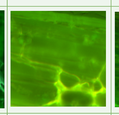
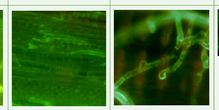
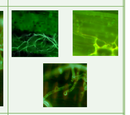
SYNTHETIC IMAGE GENERATION

Our objective: offset the deficiency of images by creating realistic images containing the desired filamentary pattern and variable blur effect and employing a data augmentation approach.

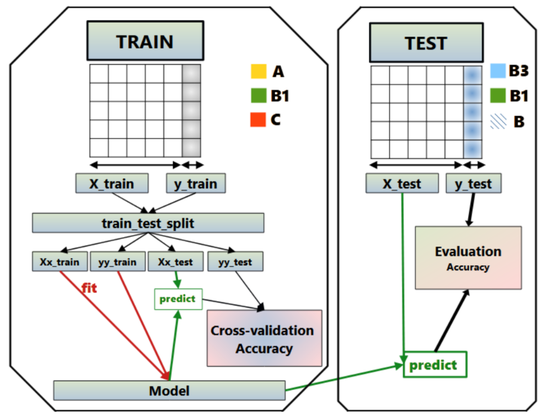


AVAILABLE DATASETS

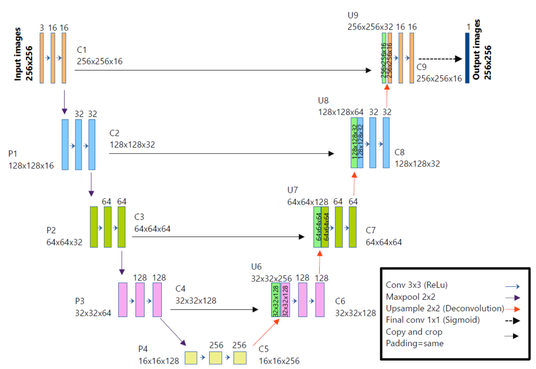
B : real images

dataset name	A	B1	B2	B3	C
description	synthetic images	with fungi in lower quality image	without fungi	with fungi	mix A, B1 and B3
number of images	427	247	312	569	128
images					
segmentation					

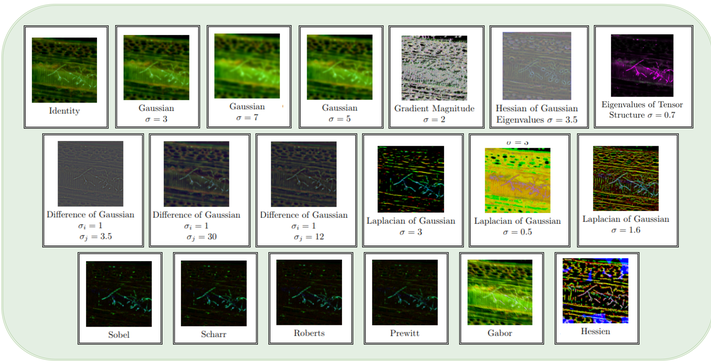
TRAINING & TEST



LEARNING WITH U-NET [2]



LEARNING WITH RANDOM FOREST [3]

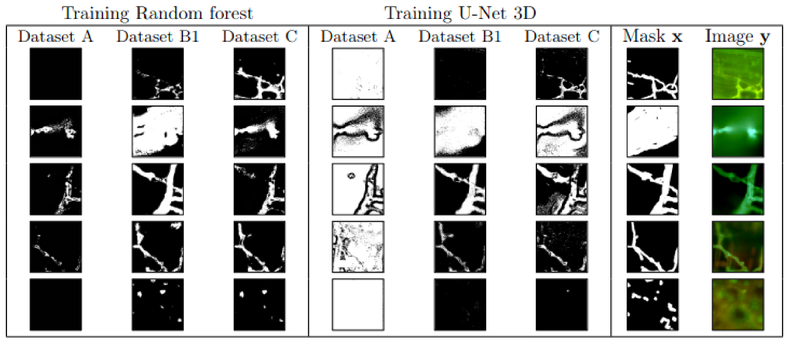


RESULTS

Segmentation accuracies, as a function of the segmentation method, the training database, and the testing database

		Random forest			U-Net 3D		
	Training database:	Dataset A	Dataset B1	Dataset C	Dataset A	Dataset B1	Dataset C
Test	CV accuracy	92.96%	96.16%	97.06%	89.68%	93.84%	90.96%
	Dataset B1	2.69%	94.13%	**95.70%**	18.98%	90.10%	83.25%
	Dataset B	41.00%	**95.70%**	92.45%	10.47%	94.17%	85.73%
	Dataset B3	6.58%	92.14%	88.08%	10.83%	**93.79%**	89.00%

Example of results on real images



- This data augmentation based on image formation helps segmenting segmenting lower-quality images.
- This method could be generalized to other microscopy imaging techniques.

REFERENCES

[1] Lecomte, P. et al. (2018). Esca of grapevine and training practices in France. *Phytopathologia mediterranea*, 57(3), 472-487.
[2] Ronneberger, O. et al. (2015). U-net: Convolutional networks for biomedical image segmentation. In *MICCAI 2015*, pp. 234-241.
[3] Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32