Wheat leaves diseases segmentation : a comprehensible pipeline for phenotypic analysis

Titouan Le Gourrierec[1]**, Léna Connesson**[1]**, Axel Prouvost**[1]**, Baptiste Magnier**[2]**, and Clément Plessis**[3]

[1]Department of Computer Science & Artificial Intelligence, IMT Mines Ales, Ales, France
[2]EuroMov Digital Health in Motion, Univ Montpellier, IMT Mines Ales, Ales, France
[3]INRAE, team Ge2pop - UMR AGAP 1334, Montpellier, France

This study presents a novel high-throughput phenotyping method using artificial intelligence to quantify the presence of rust and oidium on wheat leaves within complex varietal mixes. Addressing the challenge of efficiently identifying multiple foliar diseases in mixed-species agricultural plots, our approach utilizes a custom-designed Python package interfaced with the ilastik annotation environment. We developed a comprehensive pipeline involving color space transformation to enhance disease segmentation on high-resolution leaf images. Utilizing a traditional random forest algorithm for pixel classification, the model demonstrates an accuracy of 99% on isolated disease patches, with average surface error rates of 14% for rust and 5% for oidium, reflecting small actual surface errors. The performance on real leaf samples, with more nuanced textures, is modest but improved through outlier exclusion, achieving practical accuracies of 62% for rust and 52% for oidium. This study illustrates the potential of AI-enhanced high-throughput methods in reducing pesticide use through better understanding and management of crop diseases.

Wheat phenotyping | Machine learning | Rust | Oidium

Introduction

Inter- and intraspecific mixed cropping in the same agricultural plot allows for better control of various bioaggressors : fungal diseases, insects, weeds... (1)(2). In the perspective of minimizing pesticides in field crops, it is necessary to optimize varietal mixtures for bioaggressor control. The study of the development of different pathogens in the plant cover is an important point in improving these mixtures. In complex mixtures (more than 2 varieties), it is difficult to efficiently evaluate the symptoms of several diseases at the same time in the same plot. We propose here a high-throughput phenotyping method in complex mixture for rust and oidium based on a machine learning model.

The purpose of this study is to develop a tool of this type, to measure in particular the surfaces infected by two diseases on wheat leaves.

The literature includes a large number of plant segmentation experiences (3)(4) or foliar disease detection (5). But examples of foliar disease segmentation are rarer (6). However, the problem arises when we are interested in comparing the resistance of plant varieties to certain diseases (7) in particular by noting the diseased surfaces in order to determine the most resistant variety.

Our study will focus on the segmentation of oidium and rust on wheat leaves. It is part of the INRAE project MoBiDiv on the transition to a pesticide-free agriculture model (8). In this study, brown and yellow rust will be confused to limit the complexity of the model.

The core of the model uses the features and the annotation environment of the ilastik (9) software that we have run on Python via a package that we have designed. The preprocessing of the wheat leaf images to be analyzed consists of a change in color space ; several variants have been compared. These two steps constitute our analysis pipeline, available on github.

Material and Methods

The dataset consists of 1600 scanned A4 sheets like in Figure 1, each containing 5 wheat leaves from different varietal mixtures.

FIGURE 1 – Five wheat leaves scanned on a white A4 sheet, each identified by a unique label that indicates the plot number and the sample number

These leaves are taken from 304 micro-plots, each combining 8 genotypes among the 100 present, under different treatment conditions. For each micro-plot, 40 leaves were collected to best represent the mixture. Samples were also taken from 364 micro-plots in monoculture where 5 leaves were collected per plot. A label appears on each of the sheets of paper to identify the origin of each wheat leaf.

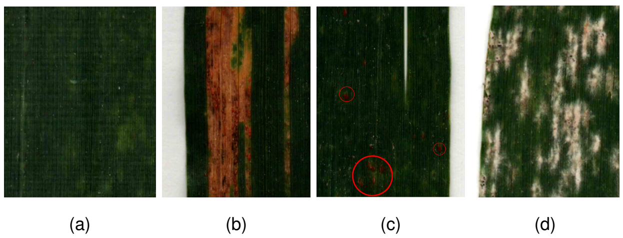

FIGURE 2 – Leaf disease and main pattern detection, (a) Healthy wheat leaf with no pattern of disease, (b) Yellow rust (groups of longilinear pustules), (c) Brown rust (sparsed dark pustules), (d) Oidium (white spots with black points, downy)

The scanner that was used is the IRIScan™ Book 5 and can scan A4 paper sheets to .jpeg format at a resolution of 1200 dpi, enabling to identify clearly signs of diseases Figure 2. Sheets that had scanning defects are taken care of by our pipeline but are discarded from usable data. Others had physiological defects (senescence, drought, necrosis, other diseases) but we kept them, as well as for sheets that were attacked by insects.

The leaves were isolated and then renamed according to their label using Python.

Six color spaces were tested during training : RGB, HSV, YUV, LAB, Eigen color, HSL. Each transformation is illustrated by Figure 3. These are often used in the context of plant segmentation to improve predictions (10)(11).

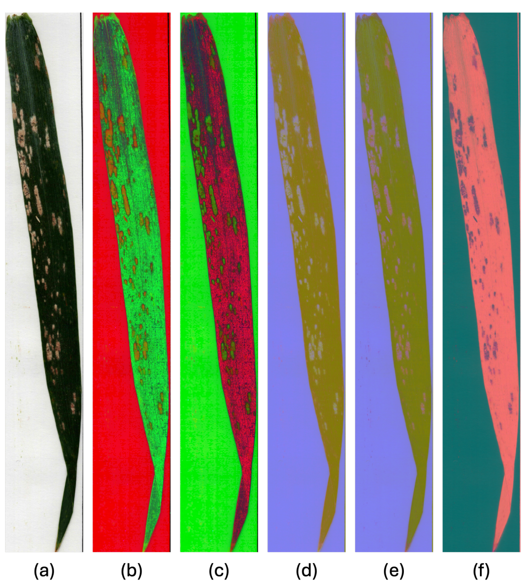

FIGURE 3 – Annotated leaves with *ilastik* on 4 classes : background (yellow), healthy leaf (blue), oidium (red), rust (cyan), brush thickness is changeable

Leaf annotation for training was performed using the *ilastik* 1.4.0 software (Figure 4). The annotations were divided into four labels : healthy leaf, background, oidium, and rust.

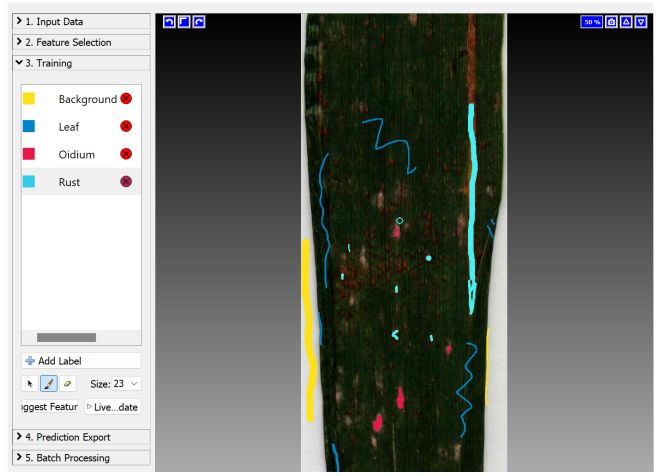

FIGURE 4 – Annotated leaves with *ilastik* on 4 classes : background (yellow), healthy leaf (blue), oidium (red), rust (cyan), brush thickness is changeable

A classical random forest algorithm was used as the pixel classification algorithm (5). Additionally, *ilastik* can suggest the most relevant filters to apply, among the 37 available. Training and the pixel classification algorithm were later run on a MacBook Pro 16" 2021 - M1 Pro Chip - 3.2 GHz - APPLE GPU 16 - 10 Cores - 16 GB DDR5 RAM - 2TB SSD - MacOS 14.4.1 Sonoma.

Data validation was performed using disease annotation on VGG Anotator 2.0.12 (12).

Metrics were also evaluated on real texture patches from the leaves as shown in Figure 5. Their simple geometry allows for a better qualitative comparison of the results with the annotated images (13)(14). Since disease tasks are small in size, the textures used were reproduced identically and mirrored to avoid edge discontinuities. The metrics were then calculated on the model's predictions on normal leaves.

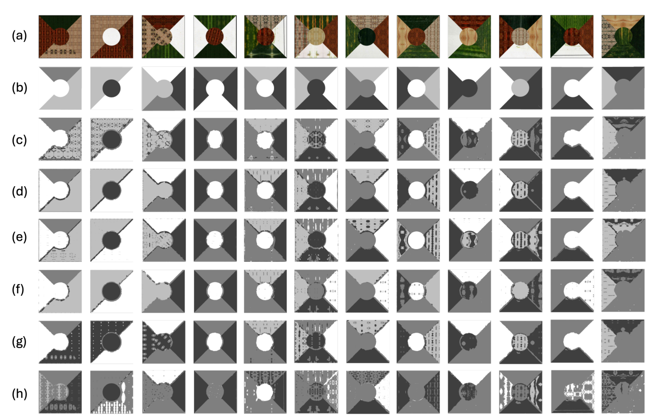

FIGURE 5 – Efficacy Testing of Different Color Spaces on Texture Patches : (a) RGB texture patches, (b) Validation mask, (c) RGB predictions, (d) HSV predictions, (e) HSL predictions, (f) LAB predictions, (g) YUV predictions, (h) Eigen color predictions

Results

Comparing the metrics on patches and leaves allowed us to identify LAB as the most performing color space.

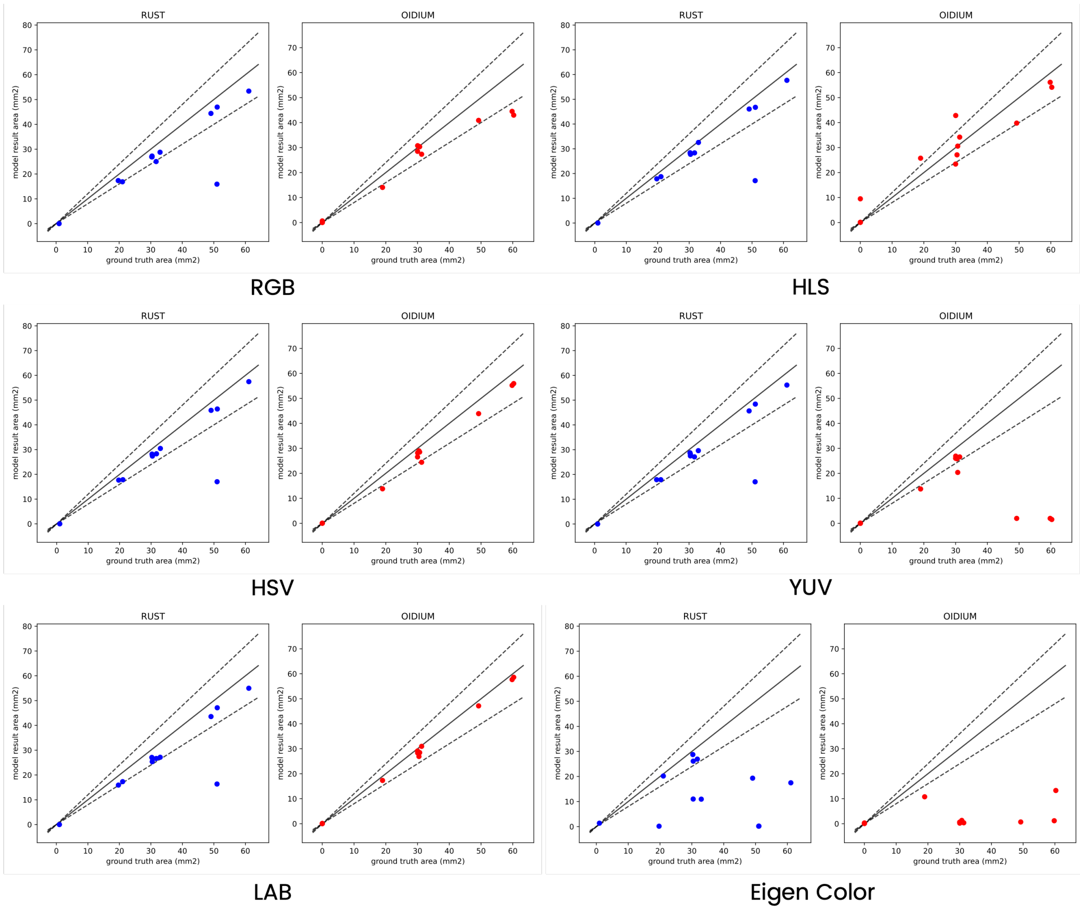

FIGURE 6 – Comparison of actual and predicted diseased areas by color spaces on patches. X-axis : Truth disease area ; Y-axis : Predicted disease area. A perfect model tends towards linearity.

Figure 6 illustrates the prediction performance for each color space. Two color spaces already stand out : LAB and HSV.

Table 1 sums up the results on patches. The segmentation seems to work well, despite some errors on the edges. The model achieves 99% accuracy for both diseases for the best case (in LAB). The error reported relative to the total disease surface area is an average of 14% for rust and 5% for oidium, which generally represents less than 0.1cm^2.

	Rust	Oidium
Precision	0.993	0.987
Recall	0.807	0.934
F1-Score	0.879	0.860
IoU	0.802	0.923
Surface error (%)	13.86	5.31

TABLE 1 – Best model metrics (LAB) on patches

	Rust	Oidium
Precision	0.62	0.52
Recall	0.53	0.61
F1-Score	0.51	0.6
IoU	0.45	0.38
Surface error (%)	25.08*	19.75*

TABLE 2 – Best model metrics (LAB) on leaves (*without outliers)

The results are logically less good for real leaves, with more nuanced textures as seen in Table 2. However, it only takes a small error on a leaf where the diseased surface is very small for the relative error reported to the real diseased surface to increase disproportionately. We have therefore chosen to remove the extreme values (or 'outliers'), which allows us to achieve an accuracy of 62% for rust and 52% for oidium. The surface errors are finally 25% and 20%, respectively for rust and oidium. An example is shown Figure 7

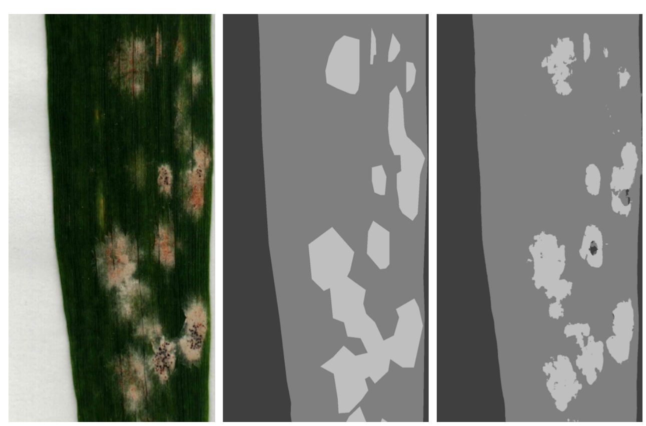

FIGURE 7 – Original sick leaf (left) - Ground truth where oidium spots were annotated with VGG anotator (middle) - ilastik's predictions about oïdium (right)

Discussion

The data collection protocol caused several problems. The first is that the wheat leaves were scanned on paper sheets of the same color as oidium. However, wheat leaves frequently have holes, which are often confused by the model with oidium.

One solution would have been to scan the wheat leaves on a blue paper sheet, for example. In addition, the images were all compressed due to their .jpeg save format. An export in .png format would have improved the results. It would even have been interesting to use scanners capable of detecting in the near infrared in order to calculate an NDVI index and thus allow earlier disease detection. (15)

The model tends to predict a few diseased pixels on healthy leaves. For performance, it would be interesting to define a threshold value from which it is considered that the leaf is really diseased, and thus not to take into account healthy leaves in the metrics. Figure 8 shows the evolution of the error as a function of the chosen threshold value.

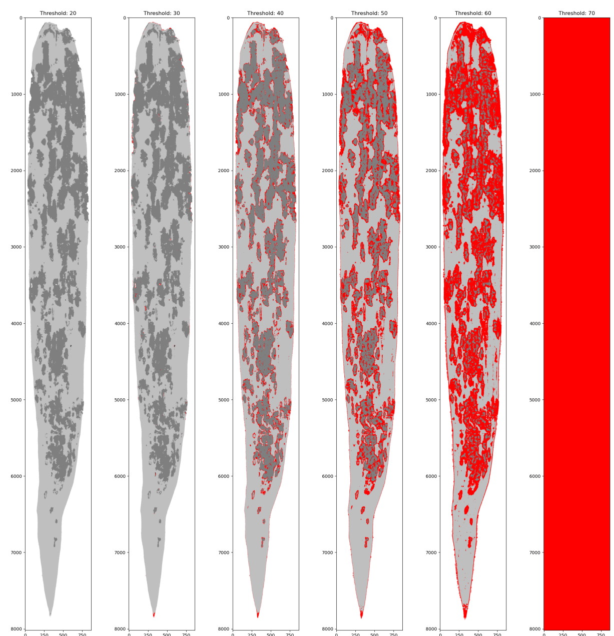

FIGURE 8 – Red color represents a pixel whose highest probability of belonging to one of the 4 classes is lower than the threshold. Thus, the highest confidence probability of a pixel is 70%.

In addition, it would undoubtedly be interesting to optimize a morphological opening-based post-processing method to improve the results by determining an appropriate kernel size, as well as threshold to determine if the leaf is ill.

Conclusions

This study presents a novel high-throughput phenotyping method using artificial intelligence to quantify the presence of rust and oidium on wheat leaves. Our approach utilizes a custom Python package interfaced with the ilastik annotation environment and develops a comprehensive pipeline involving color space transformation to enhance disease segmentation on high-resolution leaf images.

Key Points :

- A machine learning model was developed for rust and oidium segmentation on wheat leaves. The model achieved 99% accuracy on isolated texture patches and 62% for rust and 52% for oidium on real leaves.
- Our method enables the analysis of a large number of wheat leaves in a short time and at a low cost.
- The LAB color space proved to be the most efficient for disease segmentation on wheat leaves.
- A complete pipeline for high-throughput phenotyping of foliar diseases was established. The pipeline includes image preprocessing and disease segmentation.

Future Directions :

- This method can be extended to other foliar diseases of wheat and other crops.
- Several avenues for improving the model on real leaves (post-processing, optimization of the data collection protocol, etc.) remain to be explored.

To conclude our study demonstrates the potential of artificial intelligence for high-throughput phenotyping of foliar diseases. Our method has the potential to improve the accuracy, efficiency, and cost of foliar disease phenotyping, which could lead to the selection of more disease-resistant wheat varieties.

Bibliography

1. C. C. Jaworski, E. Thomine, A. Rusch, A.-V. Lavoir, S. Wang, and N. Desneux. Crop diversification to promote arthropod pest management : A review. *Agriculture Communications*, 1(1) :100004, 2023. doi : 10.1016/j.agrcom.2023.100004.
2. R. Francaviglia, M. Almagro, H. Lehtonen, R. Hüppi, and J. Rodrigo-Comino. Editorial : Agricultural diversification : Benefits and barriers for sustainable soil management. *Frontiers in Environmental Science*, 10, 2022. doi : 10.3389/fenvs.2022.1046354.
3. M. Serouart, S. Madec, E. David, K. Velumani, R. Lopez Lozano, M. Weiss, and F. Baret. SegVeg : Segmenting RGB Images into Green and Senescent Vegetation by Combining Deep and Shallow Methods. *Plant Phenomics*, 2022, 2022. doi : 10.34133/2022/9803570.
4. M. Henke, K. Neumann, T. Altmann, and E. Gladilin. Semi-Automated Ground Truth Segmentation and Phenotyping of Plant Structures Using k-Means Clustering of Eigen-Colors (kmSeg). *Agriculture*, 11(11) :1098, 2021. doi : 10.3390/agriculture11111098.
5. Manuel Fernández-Delgado, Mark D. Partridge, and Keith Stephenson. Ensemble methods for unsupervised and semi-supervised learning : A survey. *Data Mining and Knowledge Discovery*, 46 (2) :563–606, 2022. doi : https://doi.org/10.1007/s13593-021-00738-8.
6. K. Khan, R. U. Khan, W. Albattah, and A. M. Qamar. End-to-End Semantic Leaf Segmentation Framework for Plants Disease Classification. *Complexity*, 2022 :1–11, 2022. doi : 10.1155/2022/1168700.
7. G. Montazeaud, T. Flutre, E. Ballini, J. Morel, J. David, J. Girodolle, A. Rocher, A. Ducasse, C. Violle, F. Fort, and H. Fréville. From cultivar mixtures to allelic mixtures : opposite effects of allelic richness between genotypes and genotype richness in wheat. *New Phytologist*, 233(6) :2573–2584, 2022. doi : 10.1111/nph.17915.
8. Aline Fugeray-Scarbel, Jérôme Enjalbert, and Myriam Tisserand. MoBiDiv - MObilizing and Breeding Intra and inter-specific crop DIVersity for a systemic change towards pesticide-free agriculture. Quelle recherche pour répondre aux objectifs de réduction des pesticides inscrits dans le Green Deal européen, June 2022. Poster.
9. Stuart Berg, Dominik Kutra, Thorben Kroeger, Christoph N. Straehle, Bernhard X. Kausler, Carsten Haubold, Martin Schiegg, Janez Ales, Thorsten Beier, Markus Rudy, Kemal Eren, Jaime I. Cervantes, Buote Xu, Fynn Beuttenmueller, Adrian Wolny, Chong Zhang, Ullrich Koethe, Fred A. Hamprecht, and Anna Kreshuk. ilastik : interactive machine learning for (bio)image analysis. *Nature Methods*, September 2019. ISSN 1548-7105. doi : 10.1038/s41592-019-0582-9.
10. Nils Lüling, David Reiser, Alexander Stana, and H.W. Griepentrog. Using Depth Information and Colour Space Variations for Improving Outdoor Robustness for Instance Segmentation of Cabbage. *Journal of Agricultural Engineering*, 2023.
11. G. Ruiz-Ruiz, J. Gómez-Gil, and L.M. Navas-Gracia. Testing different color spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *Computers and Electronics in Agriculture*, 68 :88–96, 2009. doi : 10.1016/j.compag.2009.04.009.
12. Abhishek Dutta and Andrew Zisserman. The VIA annotation software for images, audio and video. In *Proceedings of the 27th ACM International Conference on Multimedia*, MM '19, New York, NY, USA, 2019. ACM. ISBN 978-1-4503-6889-6/19/10. doi : 10.1145/3343031.3350535.
13. Christian Ledig, Wenzhe Shi, Wenjia Bai, and Daniel Rueckert. Patch-based Evaluation of Image Segmentation. *CVPR*, 2014.
14. B. Magnier, P. Montesinos, and D. Diep. Texture Removal Preserving Edges by Diffusion. In *Image Analysis*, pages 6–8. Springer International Publishing, 2015. doi : 10.1007/978-3-319-19665-7_1.
15. Nils Lüling, David Reiser, Alexander Stana, and H. W. Griepentrog. Using depth information and colour space variations for improving outdoor robustness for instance segmentation of cabbage (version 1). 2021.