École Nationale Supérieure d'Électrotechnique, d'Électronique, d'Informatique, d'Hydraulique et des Télécommunications

Projet Long IATI-SIA
Rapport

Enhancing Poplar Plantation Classification through Super Resolution

---

Pierre Ribardière - Mohamed Elkairouh - Auriane Bonnot - Maroua Agrouaz

March 11, 2024

Contents

1 Introduction

In this project, our team collaborated with **DYNAFOR**, a laboratory specialized in satellite image processing for monitoring agroforestry landscapes. Our main objective was to identify and keep track of poplar plantations using temporal series of images of France from Sentinel-2 captured in 2018. To do so, poplars need to be detected using a novel remote sensing index called Poplar Index 2 (**PI2**)[**4**], developed by Dr. Yousra Hamrouni, our project tutor. However, this operation is challenging because poplars under 5 to 7 years old are hard to detect due to their small size. To address this, a Super Resolution approach involving a deep learning model trained on VENµS's images can be used to increase the resolution of Sentinel-2 images from 10m to 5m. We then train Random Forest classifiers on each set of data (with and without Super Resolution) to assess whether the detection of poplars is improved. Our focus is on optimizing image preprocessing and the classifiers' hyperparameters, as well as highlighting the features that have the most impact on both the classification and Super Resolution performance.

2 Details of Research

2.1 Super Resolution

The data we are processing is composed of multi-spectral time series of images from Sentinel-2 covering four different areas in France, capturing 36 dates during 2018. To compute the PI2, the B5 (705nm), B11(1610nm) and B12 (2190nm) spectral bands are required. Therefore, Super Resolution is applied to these three bands for each date and area. As the temporal information is crucial, a Single Image Super resolution approach was utilized through an ESRGAN model[**6**] trained on SEN2VENµS [**2**] dataset, with B11 and B12 images generated using the Wald protocol.[**3**] Processing remote sensing images can be computationally intensive, so the laboratory's computer was employed to generate the high-resolution images.

2.2 Data Preprocessing for Classification

2.2.1 Trainset

The objective was to train Random Forest classifiers on three tiled satellite images (T30TYT, T31TCJ, T31UEQ) and validate on one tile (T30TYQ), each tile representing different areas of France. Initially, the tiles underwent processing in QGIS for cleaning, Sample Selection and Sample Extraction. Subsequently, a Python IDE was utilized for further processing. Initially, all non poplars were grouped into the same class and rebalanced for binary classification. By plotting the PI2 medians and interquartiles, a distinction between poplars and non-poplars was observed between approximately the dates 4/26 to 03/09. The separation is even clearer when plotting for each tile.

This suggested potential advantages of conducting feature engineering to enhance classification quality. Principal Component Analysis (PCA) was tested on the features to determine the optimal number of principal components[**8**]. Additionally, a feature reduction approach was performed using Sequential Forward Feature Selection (SFFS) on raw data. The five best features were selected based on the "recall" criterion rather than

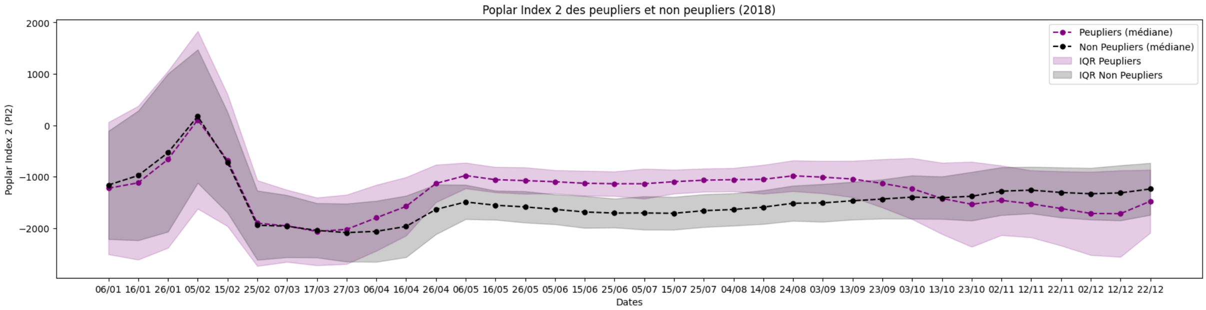

Figure 1: Poplar Index 2 medians and interquartiles from the trainset

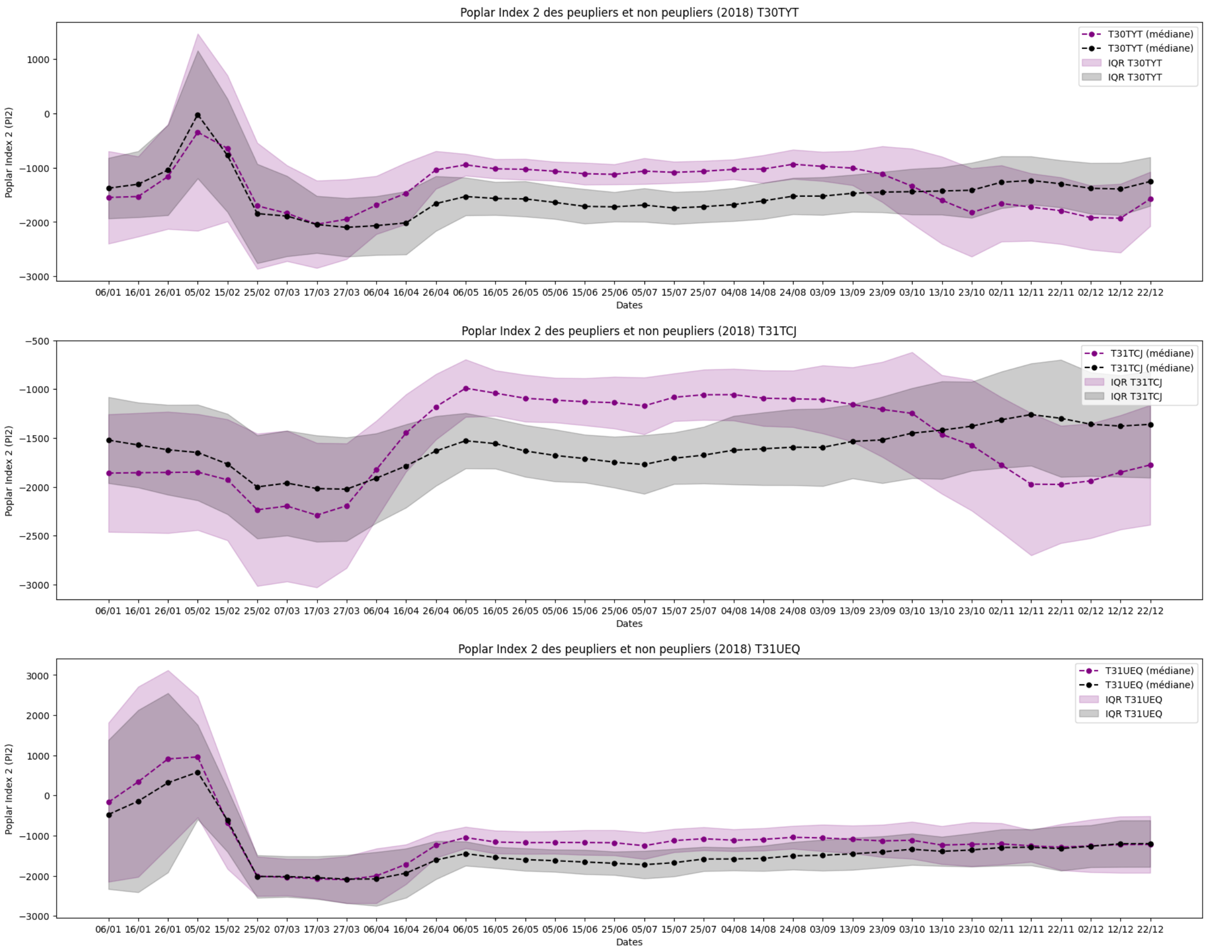

Figure 2: Poplar Index 2 medians and interquartiles for each tile

the "accuracy" criterion, as these features were chosen by Dr. Hamrouni for generating super-resolution images. Finally, PCA was also applied to the SFFS-selected features.

2.2.2 Testset

The testset represented by the tile T30TYQ reveals a superior complexity of the data than the trainset, as the PI2 plot demonstrates. The PI2 medians and interquartiles are more overlapping and there is almost no distinct separation between poplars and non poplars at any period. The structural difference between the trainset and the testset can already predict the challenging aspect of the training process.

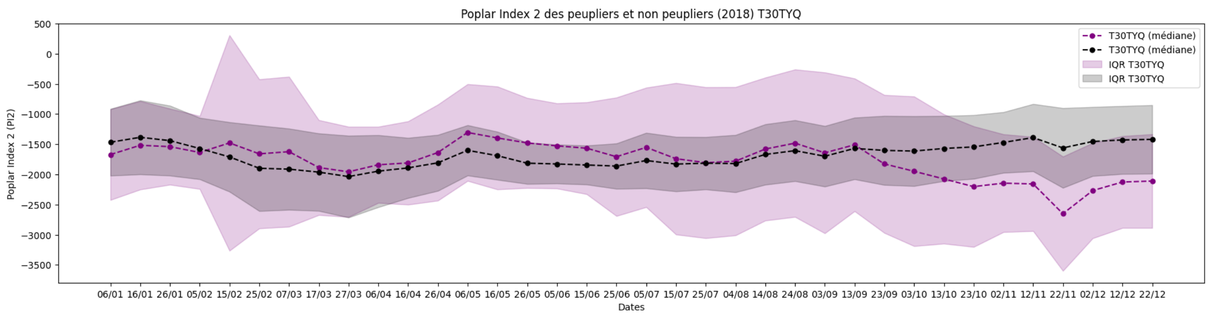

Figure 3: Poplar Index 2 medians and interquartiles from T30TYQ

2.3 Classifiers Training

The classifiers underwent training using a 10-fold cross-validation approach on the three tiles for training. Subsequently, the best model was trained on the entire trainset and evaluated on the testset. Evaluation metrics included accuracy, the confusion matrix, precision, recall, and the F1-score. The process was repeated for raw data and for each feature engineering approach, applied to both low and high resolution images. The models' hyperparameters were optimized through a RandomizedSearchCV algorithm with 40 iterations, utilizing optimization intervals derived from Dr. Hamrouni's research paper [1]: "max depth", "max features" and number of estimators. Specifically, here are the variation intervals for each parameter:

```
param_dist = {
    'n_estimators': randint(100, 151),
    'max_depth': [None] + list(range(5, 51, 10)),
    'max_features': ['sqrt'] + list(range(1, 21, 2))
}
```

Figure 4: Hyperparameters variation intervals

3 Results and Conclusions

3.1 Low Resolution Images

Applying SFFS on **"recall"** accuracy selected the following features: '04/16', '06/05', '07/25', '08/24', '10/23'. While SFFS with **"accuracy"** criterion selected these dates: '02/05', '05/06', '07/05', '07/25', '11/12'.

accuracy: ('05/02', '06/05', '05/07', ***'25/07'***, '12/11')

recall : ('16/04','05/06', ***'25/07'***, '24/08', '23/10')

Figure 5: Features selected by SFFS based on accuracy and recall

Moreover, applying PCA on raw features best performed with 3 principal components (PC), while PCA on SFFS "recall" excelled with 5 PC.

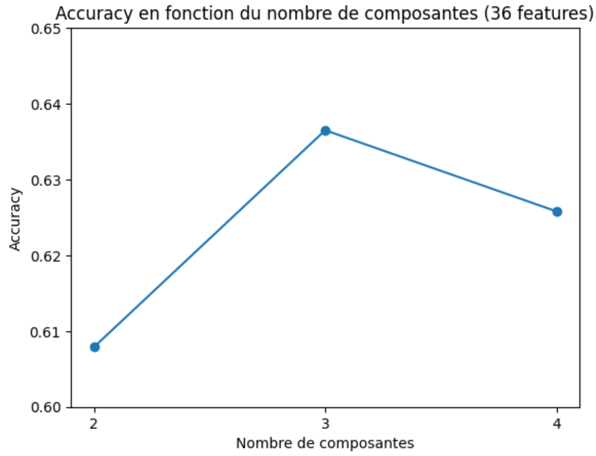

Figure 6: Variation in accuracy based on the number of raw feature components used

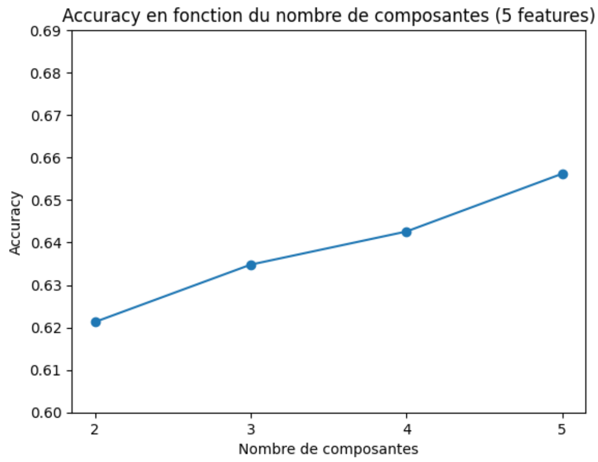

Figure 7: Variation in accuracy based on the number of the selected SFFS "recall" components used

Overall, the best performance was achieved with SFFS on the "accuracy" criterion with 74% of accuracy, followed with 73% by training on raw features and 70% by training with recall-based SFFS.
Focusing on the F1-scores, accuracy-based SFFS outweigh other methods with 0.79 and 0.66, respectively non poplars and poplars' score.
Despite PCA being not as efficient as hoped, it still enhanced the F1-scores compared to the raw features. PCA on SFFSs outputs the lowest accuracy while still giving better F1-scores than raw data. The SFFS with "accuracy" wins on all criteria, however, for comparison motives with training on high-resolution images, SFFS with "recall" is selected for further studies.
The best model obtained with RandomizedSearchCV is set with the following hyperparameters: N_estimators = 125, Max_features = 17, Max_depth = 45. Hyperparameter optimization operated on recall-based SFFS data did not enhance the accuracy but increased the F1-score from 0.70 to 0.77 for non poplars, and 0.49 to 0.59 for poplars. This is due to a better precision and recall.

Features	Test Accuracy	Train Accuracy
Raw features	0.73	**0.98**
SFFS Recall	0.70	0.96
SFFS Accuracy	**0.74**	0.97
PCA	0.64	0.95
PCA SFFS Recall	0.62	0.96
PCA SFFS Accuracy	0.56	0.97
SFFS "Recall" Optimized	0.70	0.96
PCA SFFS "Recall" Optimized	0.62	0.92
SFFS "Recall" Optimized SR	0.62	0.91
PCA SFFS "Recall" Optimized SR	0.32	0.82

Table 1: Accuracy

	Non Poplars | Poplars		
Features	**Precision**	**Recall**	**F1-score**
Raw features	0.51/0.51	0.49/0.53	0.50/0.52
SFFS Recall	0.58/0.76	0.88/0.36	0.70/0.49
SFFS Accuracy	**0.66**/**0.95**	**0.97**/0.51	**0.79**/**0.66**
PCA	0.64/0.63	0.63/**0.65**	0.63/0.64
PCA SFFS Recall	0.62/0.63	0.64/0.61	0.63/0.62
PCA SFFS Accuracy	0.56/0.56	0.56/0.57	0.56/0.56
SFFS Recall Opt	**0.63**/**0.94**	**0.97**/0.43	**0.77**/0.59
PCA SFFS Recall Opt	0.62/0.63	0.64/**0.61**	0.63/**0.62**
SFFS Recall Opt SR	0.58/0.76	0.88/0.36	0.70/0.49
PCA SFFS Recall Opt SR	0.26/0.36	0.20/0.44	0.23/0.40

Table 2: Precision, Recall and F1-scores

3.2 High Resolution Images

Super Resolution did not enhance the optimized classifiers' performance. With recall-based-SFFS features, the accuracy achieved was 62% with F1-scores of 0.70 and 0.49 (non poplars and poplars respectively).

3.3 Discussion

The Random Forest classifiers' best accuracy was achieved with accuracy based SFFS with 74% on validation and 98% on cross-validation, showing a potential overfitting. However, results may confirm the superior complexity of the testset in comparison to the trainset, as perceived when plotting the PI2 graph of the testset. Super resolution with 5m precision was not sufficient to improve the accuracy. However, zooming in the accuracy per age, one-year-old and three-years-old poplars gained respectively 3% and 0.1% in accuracy, while age groups from 4 to 8 years old, including 2 years-old, all lost performance by a range of 4 to 24%.
Besides, the plot shows that accuracy increases with age. This behavior is expected as

older poplars generally correlates with greater growth, thus easier to detect. However, four-years-old stands out with the highest accuracy (83% and 76%, respectively on low resolution and high resolution images).

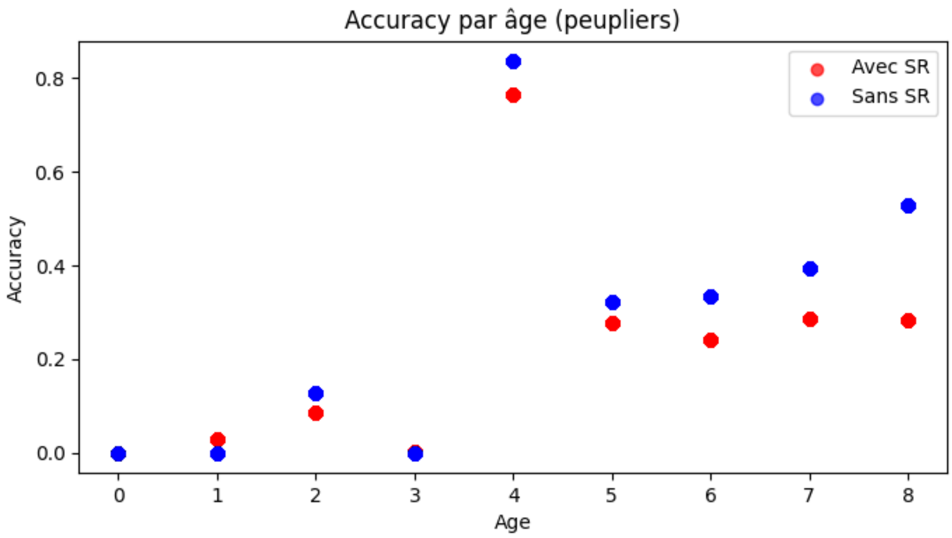

Figure 8: Variation in accuracy based on the age group

This is explained especially by the type of cultivars differing from amongst and between age groups. Certain type of cultivars such as Soligo and Lena have faster juvenile growth and growth speed than other cultivars which implies a higher accuracy.

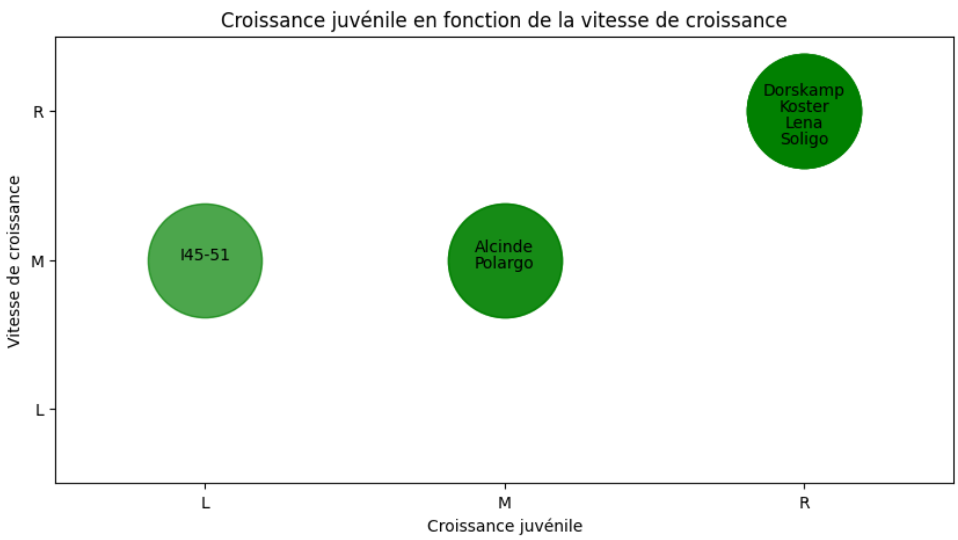

Figure 9: Growth speed as a function Juvenile Growth linked to the type of cultivars

As the Soligo's poplars are all four years old, it explains the peak of accuracy for this age group. Besides, 7 out of 24 cultivars gained performance ranging 8 to 15% with Super Resolution, while others' degraded. Therefore, not only the age groups but also the type of cultivars impacts the classification and Super Resolution performance. Plotting the accuracy per age and cultivar for both low and high resolution images supports the previous statement, as for a same age, cultivars all have different performance. However, for the most, the older the cultivars are, the accuracy is higher. Besides, although some previously performant cultivars slightly lost performance, Super Resolution has led to new small bubbles appearing on the graph (figure 11). Young cultivars aged under 3 years old, such as I45/51, Koster-I45/51 and Raspalje are now more detected than before.

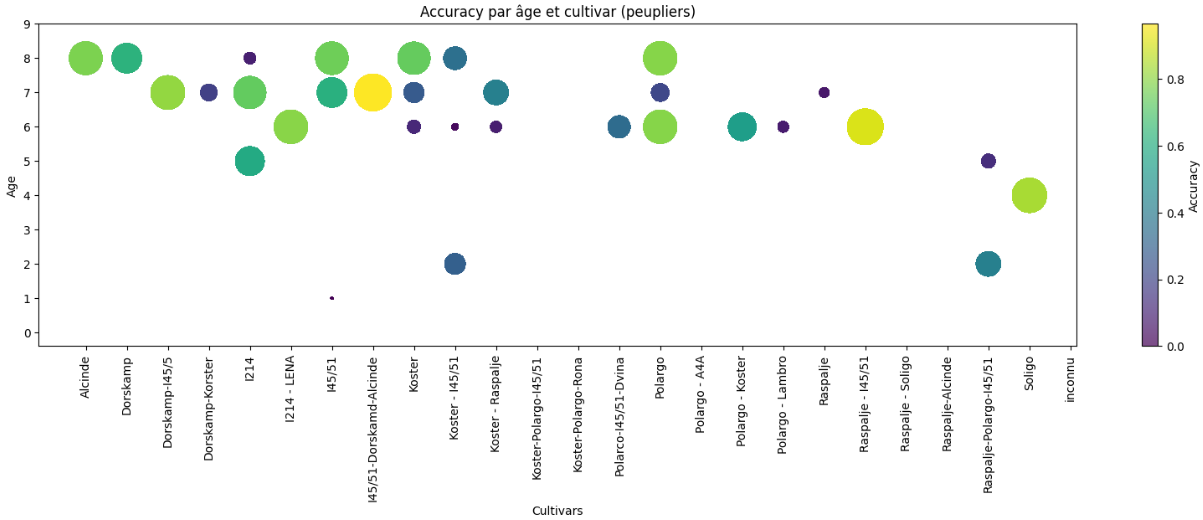

Figure 10: Accuracy per age and cultivar for low resolution images

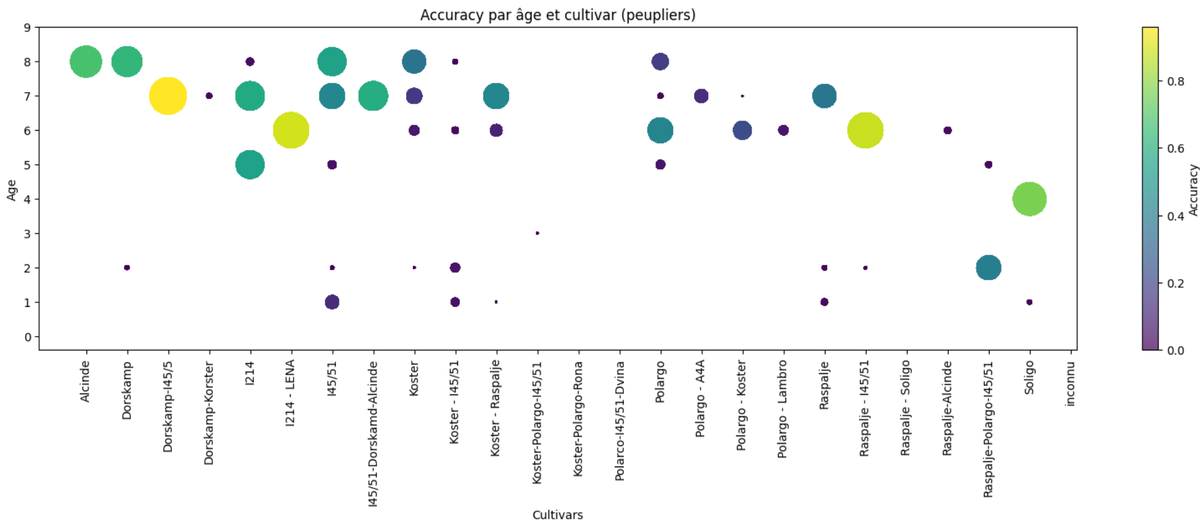

Figure 11: Accuracy per age and cultivar for high resolution images

Furthermore, it is important to notice that the data are very correlated to the area of France they belong to. While the testset samples are very different from all other tiles, we tried to train a model on two tiles amongst the trainset and validate on the third one, and repeat the process for the three possible rotations. The accuracy averaged on the three rotations was 96%, which is much better than 74% obtained earlier. This reveals once more the complexity of the testset, difficult to be captured by our models.

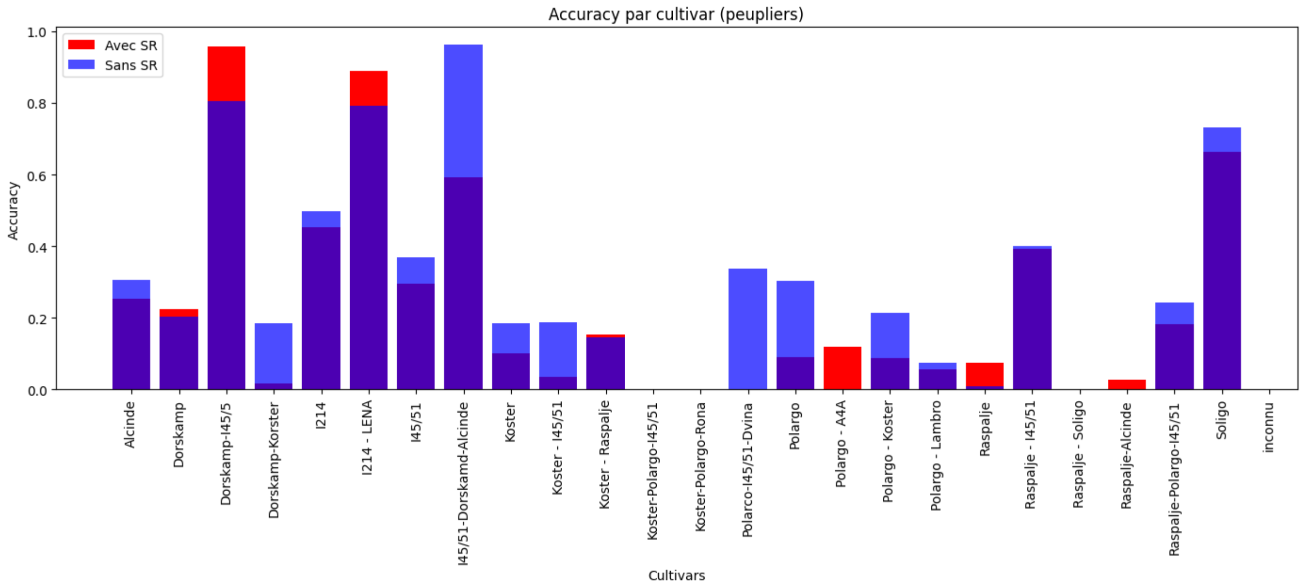

Figure 12: Comparison of the variation in accuracy based on the type of cultivars between low and high resolution images

3.4 Paths for further enhancements

A deeper study could focus on other biases impacting both Super Resolution and classification's performance. Aspects such as artefacts, distortions and noises due to remote sensing imagery or Super Resolution could be responsible for loss of performance. It is also worth mentioning that a resolution of 5 meters may not be sufficient for the detailed detection needed, suggesting the potential for exploring further improvements in resolution. Besides, Super Resolution could be applied to the accuracy-based Sequential Forward Feature Selection (SFFS) data for a performance comparison to recall-based SFFS data. Moreover, exploring novel deep learning architectures for both Super Resolution and classification, such as those based on Convolutional Neural Networks (CNNs) or transformers, could provide enhanced methodologies for addressing the challenges in accurately detecting and classifying the target subjects.

4 Acknowledgements

We would like to express our gratitude to Dr.Yousra Hamrouni and Prof. David Sheeren for their supervision and assistance throughout the project. Their guidance and expertise were crucial in navigating the complexities of the project, and their support greatly contributed to its success. We are sincerely thankful for their dedication, patience, and encouragement during every stage of the process. Their insights and mentorship have been instrumental in shaping our understanding and approach to the subject matter.

5 Bibliography

References

[1] Yousra Hamrouni and Eric Paillassa and Véronique Chéret and Claude Monteil and David Sheeren
Sentinel-2 Poplar Index for Operational Mapping of Poplar Plantations over Large Areas, Remote Sens. 2022, 14, 3975. https://doi.org/10.3390/rs14163975

[2] Michel, J.; Vinasco-Salinas, J.; Inglada, J.; Hagolle, O.
SEN2VENµS, a Dataset for the Training of Sentinel-2 Super-Resolution Algorithms. Data 2022, 7, 96. https://doi.org/10.3390/data7070096

[3] Aurélien Lac, Julien Michel, Vincent Poulain, Nafissa Sfaksi.
Sentinel-2 Single Image Super-Resolution with the SEN2VENµS Dataset: architecture, training strategy, performances assessment and application to Water Bodies Detection. 2023. ⟨hal-04218629⟩

[4] Yousra Hamrouni.
Développement d'une approche opérationnelle pour l'identification automatique des peupleraies à large échelle par télédétection hypertemporelle. : De l'adaptation de domaine à la création d'un indice spectral dédié. Sciences agricoles. Institut National Polytechnique de Toulouse - INPT, 2021. Français. ffNNT : 2021INPT0092ff. fftel-04186362f https://theses.hal.science/tel-04186362

[5] W. Yang, X. Zhang, Y. Tian, W. Wang, J. -H. Xue and Q. Liao,
Deep Learning for Single Image Super-Resolution: A Brief Review, in IEEE Transactions on Multimedia, vol. 21, no. 12, pp. 3106-3121, Dec. 2019, doi: 10.1109/TMM.2019.2919431.

[6] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Chen Change Loy, Yu Qiao, Xiaoou Tang
ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks, 2018.

[7] Mariana Belgiu, Lucian Drăguţ,
Random forest in remote sensing: A review of applications and future directions, ISPRS Journal of Photogrammetry and Remote Sensing Volume 114, 2016, Pages 24-31, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2016.01.011.

[8] Md. Palash Uddin, Md. Al Mamun & Md. Ali Hossain
PCA-based Feature Reduction for hyperspectral Remote Sensing Image Classification, IETE Technical Review, 2021, 38:4, 377-396, DOI: 10.1080/02564602.2020.1740615