59. Monitoring the nitrogen status of a cereal in intercropping by semantic segmentation of RGB images

Z. Yao[1,*], E. Denimal[2] and C. Gée[1]
[1]*Agroécologie, INRAE, Institut Agro, Univ. Bourgogne Franche-Comté, Dijon, France ; * zexing.yao@inrae.fr*
[2]*Institut Agro Dijon, direction scientifique, Cellule d'appui à la Recherche en Science des Données, 26 Bd Dr Petitjean, F-21000 Dijon, France*

Abstract

A scalable and non-invasive solution for monitoring the nitrogen status of a cereal crop in intercropping (triticale-faba bean) using low-cost digital tools and visible imagery was presented in this study. First, to accurately isolate triticale from the scene, semantic segmentation of RGB images was performed using the DeepLabV3+ model. Two preprocessing methods, resize and cut, were evaluated to optimize the model performance. The cut method yielded better results (mean intersection over union (mIoU)=90.64±3.66). Then, using the triticale images, a new agronomic index, the normalized dark green colour index (nDGCI), was tested as an alternative to chlorophyll meter measurements, enabling large-scale nitrogen monitoring. These findings highlighted the potential of digital tools and image-based methods to drive sustainable agriculture by simplifying nitrogen monitoring techniques.

Keywords: DGCI, faba bean, machine learning, RGB image, SPAD

Introduction

Contemporary agriculture faces significant challenges related to environmental pollution, primarily driven by the excessive use of chemical inputs such as fertilizers and pesticides. While these inputs may be effective in the short term, their prolonged use poses a threat to the sustainability of agricultural ecosystems. This underscored the necessity of transitioning toward more environmentally friendly farming systems based on the principles of agroecology, which integrated ecological, social, and agronomic perspectives (Wezel *et al.*, 2009).
Intercropping is fully in line with the principles of agroecology and is currently enjoying renewed attention in Europe, particularly in low-input farming systems. This approach enhances yields, increases nitrogen supply, and provides better control of pests and weeds. In addition, legumes can fix atmospheric nitrogen and redistribute it in the soil to improve its availability for companion plants. However, to optimize cereal crop yield and fully capitalize on intercropping benefits, accurate and continuous monitoring of crop status is essential (Ditzler *et al.*, 2021).
This study first explored the potential of semantic segmentation models, such as DeepLabV3+ with ResNet50, for monitoring cereal crop performance in intercropping systems (Chen *et al.*, 2020). Semantic segmentation assigned a class label to each pixel in an image, enabling the extraction of detailed agronomic information, such as leaf area, flower counts, and crop coverage. Unlike object detection models (e.g., YOLO, Faster R-CNN), which classified objects using bounding boxes, semantic segmentation captures pixel-level information, making it particularly useful in complex intercropping scenarios, where crops had irregular growth patterns and morphological diversity. By leveraging RGB images captured with smartphones, this study evaluated the effectiveness of different preprocessing methods (image resizing and cutting) to optimize the segmentation model's performance.

Accurate nitrogen monitoring was another critical aspect of crop management. Traditional tools, such as SPAD chlorophyll meters, estimated leaf chlorophyll content, which was often used as a proxy for nitrogen status. While these tools provided non-destructive measurements, they required physical leaf contact and were not scalable for large field applications (Dong *et al*., 2019).
To overcome these limitations, the dark green colour index (DGCI) had been developed as an image-based alternative for nitrogen assessment. Unlike SPAD meters, DGCI enabled non-invasive, large-scale monitoring by extracting colour features from canopy images. In this study, a normalized version of DGCI (nDGCI), originally developed for wheat (Gée *et al*., 2023), was tested for triticale-faba bean intercropping systems. By integrating semantic segmentation models and AI-driven analysis, this research provided an efficient, low-cost nitrogen monitoring solution, advancing precision agriculture practices.

Materials and methods

Materials and datasets

This study was conducted on microplots at the INRAE experimental unit in Bretenière (Eastern France, 47°13′25″ N; 5°5′57″ E, altitude 206 m). The varieties used were Ramdam for triticale and Irena for faba bean. To achieve this, three cropping modalities were implemented (Figure 1):

1. Triticale monoculture, without nitrogen input (TTH_0N).
2. Triticale monoculture, with nitrogen over-fertilization (TTH_TS).
3. Triticale-faba bean intercropping, including mixed cropping and in-row intercropping.

Each cropping modality was replicated twice on micro-plots (9 m×1.5 m). The experiments were conducted at three phenological stages of the cereal: tillering, the one-centimeter ear stage, and the first node stage. At each stage, RGB images were captured using smartphones with the white balance set to manual. These top-down images had a resolution of 3000×4000 pixels and covered a surface area of 40 cm×70 cm per image.
Two image datasets were developed for the study (Table 1):

1. The first dataset, created in 2023, consisted of 948 images.
2. The second dataset, collected in 2024, included 1100 images, providing additional diversity in terms of growth conditions and environmental variability.

a)

b)

Figure 1. Example of RGB images. (a) Image of triticale monoculture; (b) image of triticale/faba bean intercropping.

Table 1. Number of images per modality and year for triticale monoculture and triticale-faba bean intercropping.

Year	Culture		Total
	Triticale monoculture	Triticale/faba bean intercropping	
2023	388	560	948
2024	550	550	1100

To automate the traditionally time-consuming task of image annotation, two open-source artificial intelligence models were employed: YOLO v5 and SAM (Segment Anything Model). In the absence of ground truth data, these tools facilitated labelling of various regions of interest within the images, particularly faba bean crops (25 618 labelled regions). Then, in a second pass, a visual analysis of the results was carried out to exclude problematic images. Finally, 520 images were retained among the 948 images (45%).

After creating three classes – triticale, faba bean and soil – 80% of the samples were randomly selected for training and 20% for validating the DeepLabV3+ semantic segmentation model (Chen *et al.*, 2018). This model assigns a class label to each pixel, enabling precise identification of triticale zones within complex agricultural systems. Semantic segmentation is central to this process, as it provides pixel-level precision to distinguish triticale from other classes, such as faba bean and soil. This distinction allows for the extraction of key agronomic features, including leaf density, which are directly linked to nitrogen status. Such an approach is fundamental for detailed monitoring and optimized nitrogen management within the framework of precision agriculture.

Furthermore, to compare the image-based results with a traditional method, a SPAD meter was used to obtain reference measurements for triticale by pinching the leaves. For each plot, average SPAD was calculated from 2 sets of 30 measurements, taken from the last fully expanded leaf of the triticale plants. The SPAD values ranged from 35 to 57, with higher values indicating a higher nitrogen status. To facilitate comparisons across plots, the SPAD values were normalized against those of an over-fertilized reference plot. These normalized measurements were referred to as nSPAD values.

Image processing

Semantic segmentation of the images was carried out using the DeepLabV3+ model, combined with the ResNet50 backbone. This backbone was selected for its ability to extract complex image features while maintaining a favorable balance between depth and computational efficiency (Zhu *et al.*, 2019). ResNet50 is particularly well-suited for image segmentation tasks where precision in object differentiation is critical, such as in agricultural crop monitoring (Asad & Bais, 2019).

Data preprocessing

To prepare the images for compatibility with the segmentation model, two distinct preprocessing methods were applied prior to model training:

1. Resize: this method involved resizing the original images, which had dimensions of 3000×4000 pixels, into square images of 512×512 pixels. Resizing was a simple and quick approach; however, it had a major drawback: it altered the proportions of objects within the images, potentially leading to visual distortions. This could negatively impact the quality of segmentation, as the plants and their visual characteristics were modified, thereby affecting the model's ability to accurately identify and differentiate features.

2. Cut: to address the limitations of resizing, a second method was employed, inspired by the SAHI approach for detecting small objects (Akyon *et al.*, 2023). This method involved cropping the images into multiple square patches of 512×512 pixels. Unlike resizing, this approach preserved the original proportions of objects within the images, enhancing segmentation accuracy. However, this cropping process significantly increased the dataset size, resulting in 21 347 patches generated from the original images, compared to only 520 resized images in the first approach.

Model training

The DeepLabV3+ model was used in this study is an advanced semantic segmentation architecture introduced by Chen *et al.* (2018). It combines an encoder-decoder framework with dilated convolutions (or atrous convolutions), designed to effectively capture global context while preserving fine spatial details. This architecture is particularly well-suited for segmentation tasks requiring pixel-level precision, especially in complex environments where boundaries between classes are often blurred.

In this study, the model was trained separately on datasets generated using the two preprocessing methods. Training was conducted over 40 epochs with a batch size of 24, employing the stochastic gradient descent (SGD) optimizer and a loss function based on Dice Loss, which has been applied in agricultural image segmentation tasks, such as weed segmentation (Ullah *et al.*, 2021). These hyperparameters were selected following preliminary tests aimed at maximizing model convergence and performance.

Validation was carried out after each epoch using metrics such as IoU and overall accuracy to assess the model's generalization and prevent overfitting. This approach ensured that the model achieved robust performance in various data configurations.

Performance evaluation

To evaluate the quality of the segmentation achieved by the model for each preprocessing method, two standard metrics were employed:

1. Mean Intersection over Union (mIoU): These metric measures segmentation accuracy by calculating the ratio between the intersection and the union of the predicted and actual regions (Everingham *et al.*, 2010). It reflects the model's ability to accurately segment all pixels belonging to each class. A high IoU indicates good spatial alignment between the model predictions and ground truth, signifying overall segmentation accuracy.
2. Mean Per-Class Accuracy (mPA): mPA calculates the accuracy for each class independently and then averages these accuracies across all classes (Shelhamer *et al.*, 2017). Unlike IoU, mPA focuses on the proportion of correctly classified pixels for each class, allowing assessment of the model's balanced performance across all classes, even when some classes are underrepresented.

DGCI and nDGCI indices for nitrogen monitoring in triticale

After segmentation and extraction of pixels corresponding to the triticale crop, the dark green colour index (DGCI) (Karcher and Richardson, 2003) was employed to estimate the nitrogen status of the crop. This colorimetric index is calculated for the pixels associated with the segmented plants and is based on the HSB (hue, saturation, brightness) colour space. A high DGCI value (close to 1) indicates better crop vigour.

For each segmented image, the DGCI is computed as the average of the DGCI values calculated for each pixel corresponding to the triticale canopy. This approach provides a robust, pixel-level assessment of the nitrogen status, enabling precise monitoring of the crop's health. The formula for calculating the dark green colour index (DGCI) is as follows:

$$\mathrm{DGCI} = \frac{(\mathrm{Hue} - 60)/60 + (1 - \mathrm{Saturation}) + (1 - \mathrm{Brightness})}{3} \tag{1}$$

To minimize the impact of lighting variations and sensor differences, a normalized index (nDGCI) was developed (Gée *et al.*, 2023). The nDGCI is calculated by dividing the DGCI of the studied plot by the DGCI of a reference plot with over-fertilization, thereby stabilizing comparisons across different images:

$$\mathrm{nDGCI} = \frac{\mathrm{DGCI_{plot}}}{\mathrm{DGCI_{reference}}} \tag{2}$$

To test this new approach for monitoring nitrogen status in triticale using visible imagery, it was compared to nSPAD values.

Results and discussion

Model performance evaluation

The performance evaluation of the DeepLabV3+ model with the ResNet50 backbone was conducted for both preprocessing methods, Resize and Cut (Figure 2a–c). This evaluation involved generating segmentation predictions and comparing them with the initial annotations to calculate performance metrics (mIoU, mPA, mPrecision, and mRecall) for the three classes: triticale, faba bean and soil.

To ensure the robustness of the results, a *K*-fold cross-validation (*K*=5) was used. This standard approach made it possible to measure performance variance in different preprocessing configurations. This method provides a comprehensive assessment of the model's reliability under varying data treatment scenarios.

The analysis of results of Table 2 highlighted a significant improvement in segmentation performance with the cut method applied to the DeepLabV3+ model. The average metrics (mIoU: 90.64±3.66, mPA: 95.12±2.03) outperformed those obtained with the Resize method (mIoU: 75.63±0.93, mPA: 86.94±2.4), confirming the importance of preserving original proportions and resolutions during training.

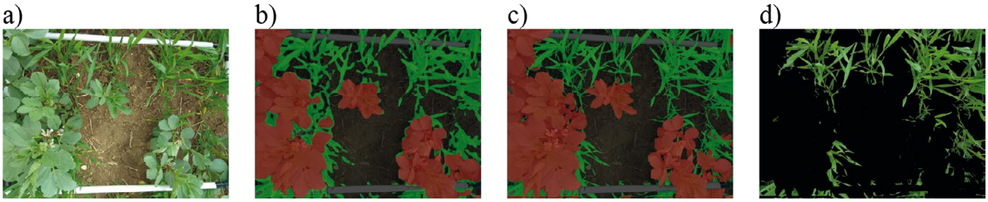

Figure 2. Examples of the results. (a) Initial image; (b) classified image (resize method); (c) classified image (cut method); (d) segmented image (final output by cut method).

Table 2. Summary of deepLabV3+ model performance according to the type of images.

Method	Class	IoU	PA	F1 score
Resize	Soil	81.49±0.49	88.2±0.86	89.80±0.60
	Faba bean	65.69±3.07	84.17±3.48	79.26±3.00
	Triticale	79.69±0.34	88.44±0.78	88.70±0.60
	Mean	75.63±0.93	86.94±2.4	86.03±1.94
Cut	Soil	93.61±0.13	97.00±0.18	96.70±0.07
	Faba bean	85.48±0.83	92.31±1.00	92.17±0.48
	Triticale	92.83±0.28	96.07±0.22	96.28±0.15
	Mean	90.64±3.66	95.12±2.03	95.05±2.04

Table 3. Relative improvement (%) of DeepLabV3+ model performance using the cut method compared to the Resize method.

	IoU amelioration (%)	PA amelioration (%)	F1 amelioration (%)
Soil	14.87	9.98	7.68
Faba bean	30.13	9.67	16.29
Triticale	16.49	8.63	8.55

The cut method significantly improved segmentation performance across all classes compared to resize, with the highest gains observed in faba bean (IoU: +30.13%), followed by triticale (+16.49%) and soil (+14.87%) (Table 3). In Table 2, the soil class achieved the best overall performance (IoU: 93.61±0.13, F1: 96.70±0.07), while faba bean benefits the most from the cut method in terms of relative improvement. The reduction in variance across 5-fold cross-validation highlighted the increased robustness of the model, essential for precision agriculture applications, where consistent segmentation accuracy is critical for crop monitoring and nitrogen management.

Nitrogen status of triticale assessed by visible imagery

After segmenting the triticale images using DeepLabV3+ and the cut preprocessing method, the nDGCI index was calculated to monitor nitrogen status by segmented image (Figure 2d). Figure 3 illustrates the relationship between nDGCI and nSPAD values for both dates and the three cropping modalities. A linear regression model is proposed, showing a correlation with a coefficient of determination (R^2) of 0.79.

This promising result encourages further studies to evaluate the robustness of the linear equation parameters (slope and bias) across different varieties, years, and pedoclimatic conditions. Such investigations will be critical to validate the applicability and consistency of the approach in diverse agricultural contexts.

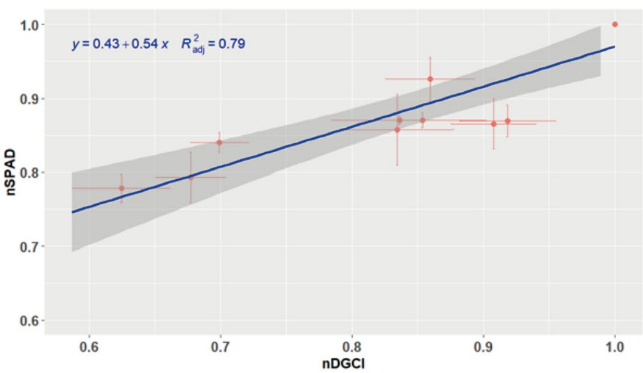

Figure 3. Linear relationship (nSPAD = 0.54 × nDGCI + 0.43, R^2 = 0.79) between nSPAD and nDGCI for three dates and the three cropping modalities with mixt and in line for triticale/faba bean.

Conclusion

This study highlighted the potential of precision technologies to enhance the sustainable management of agroecosystems, particularly in cereal-legume intercropping systems. By leveraging visible imagery and artificial intelligence models, specifically DeepLabV3+ with ResNet50, precise crop segmentation was achieved. The image cut method demonstrated notable superiority in performance (mIoU: 90.64±3.66).
The nDGCI index showed a high correlation with nSPAD measurements (=0.79) and demonstrated strong potential for proxi-sensing monitoring of cereal nitrogen through visible imagery. This study highlighted nDGCI as a scalable, fully image-based alternative to conventional chlorophyll meters, providing a cost-effective and field-adapted solution for nitrogen monitoring of a cereal crop in intercropping systems.

Acknowledgements

Many thanks to Celine Colombet, Valerie Dufayet, Frédéric Lombard and Marion Mathe who were involved in this project in the design of the crop system and in the in-field N fertilization management, respectively. This work has been granted by Plant2Pro® Carnot Institute in the frame of its 2022 call for projects. Plant2Pro® is supported by ANR (agreement No. 22-CARN-0024 01 2022).

References

Ahmad, I., Yang, Y., Yue, Y., Ye, C., Hassan, M., Cheng, X., Wu, Y., & Zhang, Y. (2022). Deep learning-based detector YOLOv5 for identifying insect pests. Applied Sciences 12(12), 5403. https://doi.org/10.3390/app12125403

Akyon, F.C., Altinuc, S.O., & Temizel, A. (2023). Slicing aided hyper inference and fine-tuning for small object detection. In: Proceedings IEEE International Conference on Image (ICIP) Bordeaux, pp. 966–970. https://doi.org/10.1109/ICIP46576.2022

Asad, M.H., & Bais, A. (2019). Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. Information Processing in Agriculture 6(4), 487-499. https://doi.org/10.1016/j.inpa.2019.12.002

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. Lecture Notes in Computer Science, 11211, 801–818. https://doi.org/10.1007/978-3-030-01234-2_49

Chen, Z., Ting, D., Newbury, R., & Chen, C. (2020). Semantic segmentation for partially occluded apple trees based on deep learning. Computers and Electronics in Agriculture, 179, 105808. https://doi.org/10.1016/j.compag.2020.105808

Ditzler, L., van Apeldoorn, D.F., von Richthofen, J.-S., & Groot, J.C.J. (2021). Current research on the ecosystem service potential of legume-inclusive cropping systems in Europe: a review. Agronomy for Sustainable Development 41(6), 1-14. https://doi.org/10.1007/s13593-021-00678-z

Dong, T., Shang, J., Chen, J.M., Liu, J., Qian, B., Ma, B., Morrison, M.J., Zhang, C., Liu, Y., Shi, Y., Pan, H., & Zhou, G. (2019). Assessment of portable chlorophyll meters for measuring crop leaf chlorophyll concentration. Remote Sensing 11(22), 2706. https://doi.org/10.3390/rs11222706

Everingham, M., Van Gool, L., Williams, C.K., Winn, J., & Zisserman, A. (2010). The Pascal Visual Object Classes (VOC) Challenge. International Journal of Computer Vision, 88(2), 303–338. https://doi.org/10.1007/s11263-009-0275-4

Gée C., Denimal E., de Yparraguirre, M., Dujourdy L., & Voisin, A.S. (2023). Assessment of nitrogen nutrition index of winter wheat canopy from visible images for a dynamic monitoring of N requirements. Remote Sensing 15(10), 2510. https://doi.org/10.3390/rs15102510

Karcher, D.E., & Richardson, M.D. (2003). Quantifying turfgrass color using digital image analysis. Crop Science, 43(3), 943–951. https://doi.org/10.2135/cropsci2003.9430

Shelhamer, E., Long, J., & Darrell, T. (2017). Fully convolutional networks for semantic segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 39(4), 640–651. https://doi.org/10.1109/TPAMI.2016.2572683

Ullah, H.S., Asad, M.H., & Bais, A. (2021). End to end segmentation of canola field images using dilated U-Net. IEEE Access 9, 59741–59753. https://doi.org/10.1109/ACCESS.2021.3073715

Wezel, A., Bellon, S., Doré, T., Francis, C., Vallod, D., & David, C. (2009). Agroecology as a science, a movement and a practice. A review. Agronomy for Sustainable Development, 29(4), 503–515. https://doi.org/10.1051/agro/2009004

Zhu, Z., Xu, M., Bai, S., Huang, T., & Bai, X. (2019). Asymmetric non-local neural networks for semantic segmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 593–602. https://doi.org/10.1109/ICCV.2019.00068