Resolving Issues with Wheat Head Detection: A Use Case of XAI in Agriculture Scenario

Abhi Lad[1][0000−0002−6585−0634] and Mehul S Raval[2][0000−0002−3895−1448]

[1] Pandit Deendayal Energy University,India
[2] Ahmedabad University,India
abhi.lce16@sot.pdpu.ac.in
mehul.raval@ahduni.edu.in

Abstract. The application of machine learning and artificial intelligence in agriculture scenario is increasing at a tremendous pace. With increasing availability of data, the scope of application of AI increases, but in turn also introduces novel challenges. In this work, we explore the explainability aspect for application in wheat head detection. The wheat head data is varied in nature, collected from multiple sources, during different growth phase and under different lighting conditions. Such multi-domain data is difficult to adapt to and results in poor generalization of the model. We perform ablation on colors, affine characteristics and using shap-value visualization of the model prediction values examine occlusion scenarios to identify how the performance of model is affected. Through this we identify the major cause of performance variation which can be used to improve the results by targeting that particular issue.

Keywords: Agriculture · Explainability · Object detection · Shapely-values

1 Introduction

With the current strides in machine learning and computer vision algorithms and availability of high volume data, the application of such techniques is increasing in agricultural practices around the globe. One such application is detection of wheat heads from images captured using either aerial drones or by vehicle mounts. Global Wheat Challenge (GWC'21) is the most recent challenge in this task. The associated GWHD'21 [1] is the bench marking dataset for wheat head detection which contains 6,515 images each with dimensions of 1024 x 1024 and 2,75,187 unique labeled wheat heads. The images are captured in individual sessions (domain), with a total of 47 sessions. A *domain* is a set of images acquired at the same location, during a coherent timestamp (usually a few hours), with a specific sensor, and corresponds to a particular development stage. The wheat can be from any of the four development stages in a particular session, namely: *Post-flowering, Filling, Filling-Ripening, Ripening.* Figure 1 represents the different development stages observed in the dataset.

The wheat heads being affected by natural factors like development stage, wind, and inherently dense growth, creates a challenging environment for object

Fig. 1. Images corresponding to development stages: Post-flowering, Filling, Filling-Ripening, Ripening respectively.

detection. Following are the major factors which contribute to the difficulty in wheat head detection.

– Densely packed and overlapping wheat heads.
– Orientation of wheat heads.
– Change in colour of wheat and surroundings.

Generally there are hundreds of wheat heads per image which results in wheat heads partially or fully overlapping other wheat heads. The mode of image capture and factor like wind change the orientation of wheat heads, which can cause difficulty in wheat head identification. The type of image sensor, time of day and growth stage of wheat determine the colour of wheat heads and the background. Figure 2 represents cases which are challenging based on perception of wheat heads.

Fig. 2. Images where wheat heads are difficult to detect due to blending background or high density.

In order to identify how each of these factor affect the performance of the object detection model, we measure drop in detection performance and observe the change in visualization of shapely values. We perform ablation on color channels, affine variations and random occlusion scenarios to observe how region of interest is affect in predicting the bounding boxes of wheat heads. We use average domain accuracy as the metric to evaluate the performance.The ADA

(Average Domain Accuracy) is calculated as follows:

$$ADA = \frac{1}{D}\sum_{d=1}^{D}\frac{1}{n_d} \times \sum_{i=1}^{n_d} AI_{di} \quad (1)$$

Where D represents the total number of domains (47 in GWHD 2021), n_d represents the number of images belonging to domain d, $d_i$represents the i^{th} image in the domain d and AI represents the accuracy for an image.

We use our custom pretrained YOLOv5 [3] model which is the GWC'21 state of the art (SOTA) and use SHAP [2] library for visualizing the region of interest. SHAP [2] (SHapley Additive exPlanations) is an approach at machine learning model explainability which relies on game theory. It gives an estimate importance to each feature based on the change in model performance in absence of that feature. In case of wheat head detection, the cases of overlap are prominent and impact the detection performance. Using SHAP, we can identify if a partially overlapped wheat head has significant enough contribution to be detected by the model. The following section explores the proposed approaches for explaining the impact of these factors and the insights derived from them.

2 Proposed Work

2.1 Ablation on Colour

The overall hue of the images vary drastically among the sessions. Wheat which are early in growth cycle i.e. post-flowering generally are greenish in hue and the ones at he end of growth cycle are yellow in hue. In order to understand the importance of colour, we perform evaluation on images where we turn off color RGB channels and measure change in performance. Figure 3 represents the resultant images of change in colour channels.

Table 1. Wheat head detection performance for different colour modes.

Color Space	Average Domain Accuracy	Performance Delta	Color Space	Average Domain Accuracy	Performance Delta
RGB	0.715	-	Gray	0.672	0.043
GB	0.546	0.169	GB-Gray	0.677	0.038
RB	0.028	0.689	RB-Gray	0.636	0.079
RG	0.679	0.036	RG-Gray	0.672	0.043

It is evident from Table 1 that in case of RGB images, green colour has a disproportionate effect on the model performance. The color green is predominant in wheat head images and thus generally will result in poor performance in its absence. The model is trained on RGB images but manages to perform very well on grayscale images. This can be attributed to model's ability to identify

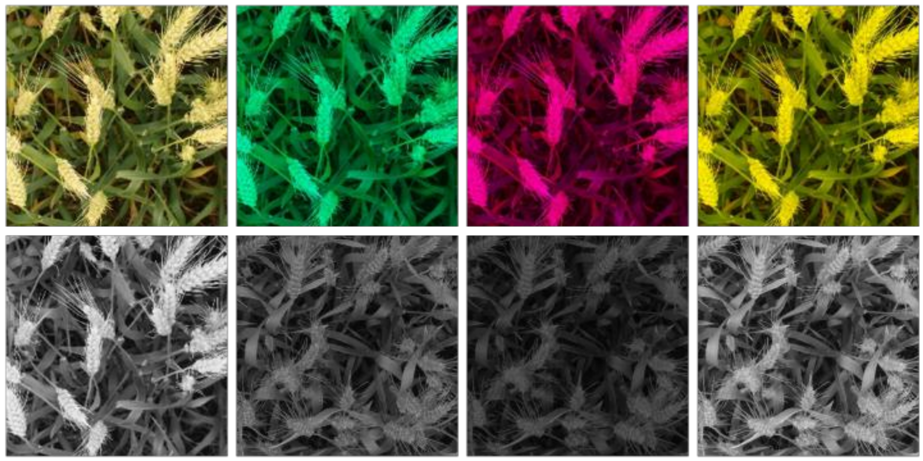

Fig. 3. Images with RGB, GB, RB, RG channels respectively and their grayscale representations.

wheat heads due to relative color difference stemming from RGB values. From this, we can infer that not the RGB colour space, but the relative difference in RGB values are crucial for model performance. Moreover, since the model is able to perform well on grayscale images, storage and training for grayscale images is far more efficient than RGB images.

2.2 Ablation on Affine Variation

There is a possibility of model being biased to a particular orientation of wheat head. We use horizontal and vertical flip of the original image to change the orientation of the wheat heads and compare the performance with original image. The size of the wheat heads is also important factor, so we reduce the size of input images to 512 x 512 and evaluate the performance.

Table 2. Wheat head detection performance for affine variation images.

Affine Variation	Average Domain Accuracy	Performance Delta
None	0.715	-
Horizontal Flip	0.710	0.005
Vertical Flip	0.709	0.006
0.5x Scale	0.471	0.244

Table 2 represents the performance of model when orientation of images are changed. The values do not change significantly when images are oriented in different direction, indicating that the model is robust to such affine changes.

However, the smaller resolution of images impacts the performance drastically. This is due to smaller object of interest which is difficult to identify.

2.3 Ablation on Occlusion Scenario

We use random occlusion to block certain part of the image and visualize the shapely values to observe the change in the region of interest. This will help to identify, if the model is able to predict wheat heads in case of partial overlap. Figure 4 shows the contribution of pixels which are important to detection of overlapped wheat head.

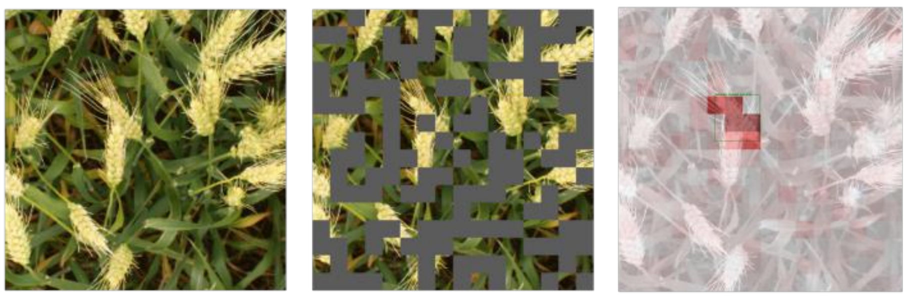

Fig. 4. Contribution of super pixels on overlapped wheat head based on SHAP values.

Based on results of SHAP visualizations, we can confirm that partially occluded wheat heads can contribute significant enough information to predict a bounding box.

3 Conclusion

In this paper, we have presented task specific explanation of object detection model performance by performing ablation on the factors of concern. These results can help improve the automated detection of wheat heads which has tremendous potential to reduce cost and time consumption in health surveillance of wheat crops. The results obtained indicate that certain colours are more important than others which supports our claim. We also establish that grayscale images are just as performant and thus usage of grayscale images can increase storage efficiency. The ablation on affine variation also indicates that the model is not biased towards a particular orientation of wheat heads, though has a small effect on the performance. But the resolution of image which indirectly represents the size of individual wheat head is very important factor. Higher resolution images, can thus further improve the performance. And finally, occlusion test confirmed that the model is able to find the wheat head even in case of occlusion, provided the wheat is not completely occluded.

References

1. David, E., Serouart, M., Smith, D., Madec, S., Velumani, K., Liu, S., ... Guo, W. (2021). Global Wheat Head Detection 2021: an improved dataset for benchmarking wheat head detection methods. Plant Phenomics, 2021.
2. Lundberg, S. M., Lee, S. I. (2017, December). A unified approach to interpreting model predictions. In Proceedings of the 31st international conference on neural information processing systems (pp. 4768-4777).
3. Jocher, G., Stoken, A., Ayush Chaurasia, Borovec, J., NanoCode012, TaoXie, Yonghye Kwon, Kalen Michael, Changyu, L., Jiacong Fang, Abhiram V, Laughing, Tkianai, YxNONG, Skalski, P., Hogan, A., Jebastin Nadar, Imyhxy, Mammana, L., . . . Wanghaoyang0106. (2021). Ultralytics/yolov5: V6. 0 - yolov5n "nano" models, roboflow integration, tensorflow export, opencv dnn support (v6.0) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.5563715