An Agricultural Precision Sprayer Deposit Identification System

Harry Rogers[1], Beatriz De La Iglesia[1], Tahmina Zebin[2], Grzegorz Cielniak[3], and Ben Magri[4]

***Abstract*— Data-driven Artificial Intelligence systems are playing an increasingly significant role in the advancement of precision agriculture. Currently, precision sprayers lack fully automated methods to evaluate effectiveness of their operation, e.g. whether spray has landed on target weeds. In this paper, using an agricultural spot spraying system images were collected from an RGB camera to locate spray deposits on weeds or lettuces. We present an interpretable deep learning pipeline to identify spray deposits on lettuces and weeds without using existing methods such as tracers or water sensitive papers. We implement a novel stratification and sampling methodology to improve results from a baseline. Using a binary classification head after transfer learning networks, spray deposits are identified with over 90% Area Under the Receiver Operating Characteristic (AUROC). This work offers a data-driven approach for an automated evaluation methodology for the effectiveness of precision sprayers.**

I. Introduction

Precision spraying systems are designed to apply agricultural chemicals with a high degree of accuracy and control, to reduce waste, and protect the environment. Evaluation of their deposits is therefore important. According to the 2019 European Union (EU) Green Deal [1], modern precision sprayers will have to undergo further regulatory assessment to ensure sprayers can achieve suitable accuracy to minimize the usage of chemicals. Despite widespread use of precision spraying systems, current methods available for evaluating their effectiveness is limited. Two common methods used are tracers and water sensitive papers [2] which are manual assessments. However, there is a big interest in the precision agriculture community to move away from these methods and to introduce automated assessments by processing data captured by a camera directly mounted on spraying systems. The task of identifying spray deposits from images is difficult, primarily due to the scarcity of useful data and the challenge of identifying transparent spray deposits in images.

There is some research within the community which makes use of current advances in deep learning. For example, some systems have been developed using Convolutional Neural Networks (CNNs) to look at spraying effectiveness [3], [4]. However, they require human intervention and are not fully automated. Our goal is to address this issue.

This work is supported by the Engineering and Physical Sciences Research Council [EP/S023917/1]. This work is also supported by Syngenta as the Industrial partner.

[1]School of Computing Science, University of East Anglia, UK. [2]School of Computer Science, Brunel University London, UK. [3]School of Computer Science, University of Lincoln, UK. [4]Syngenta ltd, UK. Email: `Harry.Rogers@uea.ac.uk; b.iglesia@uea.ac.uk;tahmina.zebin@brunel.ac.uk; gcielniak@lincoln.ac.uk;ben.magri@syngenta.com`

To overcome the lack of publicly available data, our own dataset has been created. Spraying was completed with an expert human controller using a spot spraying system called the XY sprayer. Added to the XY sprayer was a Canon 500D camera which took images before and after spraying. Based on the images provided, a classification scenario has been developed to distinguish between sprayed or dry lettuces and weeds. A binary classification scenario is an ideal formulation as the identification of sprayed or dry lettuces and weeds is the only requirement.

An eXplainable Artificial Intelligence (XAI) pipeline has been implemented using multiple pretrained CNNs to identify spray deposits. Using transfer learning [5] with a reshaped classification head sprayed images are differentiated from dry images. To improve performance, a stratification method and an additional sampling method has been developed which produces improvements in this scenario.

In order to assess the effectiveness of the proposed methodologies, classification metrics and interpretable visual indicators are going to be used. Visual indicators are used to differentiate between similar classification scores. To implement this, Class Activation Maps (CAMs) are generated by visualizing the final layers of CNNs, which allow for the computation of an importance map of the final feature extraction layer. This process not only provides valuable insight into the spatial locations of features used to correctly classify images, but also offers a means of identifying differences between similar results. These CAMs will be also evaluated to ensure their effectiveness.

The remainder of this paper is organized as follows. Section II introduces related work on precision spraying systems and state-of-the-art XAI methods for generating and evaluating CAMs. Section III introduces the experimental setup for the XY sprayer, and the associated data collection and pre-processing stages. Section IV provides details of the XAI pipeline workflow and network architectures and other implementation details. The performance of the developed networks for spray identification and CAM explanations are reported and evaluated in Section V. Finally, the paper's conclusions and future work are presented in Section VI.

II. Related Work

A. Precision Sprayer Systems

Precision spraying systems for the detection of weeds and crops are being developed with advancements from computer vision. One such system [3] adapted an all-terrain vehicle to be used as a precision spraying system. The study found that using CNNs was effective at identifying weeds and crops to then spray. Using visual observations, the authors define classes to evaluate the effectiveness of spraying. The classes are: target is fully sprayed; target is partially sprayed; target is identified but spray missed the target; target is not identified (and not sprayed); a non-target is sprayed. However, evaluation under this multi-class scenario is complex with classes that could be considered arbitrary and require human intervention to classify.

A more recent study, [6], uses a boom spraying system with CNNs to identify sugar beet and weeds as real time targets and spray them. The authors, however, only estimate a weed coverage rate given the spraying area from each nozzle and do not actually record spray deposits or evaluate what actually happened.

The studies mentioned so far do not use classical agricultural methods. There are some studies that combine CNNs and conventional agricultural methodologies to evaluate precision sprayers. An example of this is an Unmanned Aerial Vehicle (UAV) that sprays into a box that then detects the deposition of spray deposits with water sensitive papers [7]. After the UAV sprays onto the water sensitive paper, illumination is used to make the spray actuation deposition clear for a CNN. Other similar agricultural methods use tracers in precision spraying with a UAV. For example, Gao et al. [8] use water soluble Allura red food dye as a tracer. The method proposed is very effective at calculating the deposition of spray deposits and it is reported that the use of the dye does not alter the physico-chemical properties of pesticides. However, to analyze the deposition, the target must be harvested and then tested.

Ground based systems also use agricultural techniques combined with computer vision developments. Liu et al. [4] use CNNs to identify weeds within a real-world environment and included further tests in a laboratory setting. However, the evaluation of spraying deposits is completed by manually observing a red tracer added to the system.

As useful as agricultural methods are, they still fall short. The texture of water sensitive papers is different to those of crops and weeds, therefore deposits will act differently when applied to crops or weeds. Moreover, tracers can cause difficulties when mixing with chemicals that may be used when spraying. Both methods are also usually applied for specific regions of fields, to create estimates, whereas we are proposing a method that could be used throughout an entire field. There needs to be a concerted effort to move away from agricultural methodologies with their shortcomings to automated assessments.

B. *eXplainable Artificial Intelligence*

Within this paper feature extraction from CNNs is used to increase interpretability and to differentiate between similar scoring CNNs. CAMs are generated to visualize features used within CNNs giving spatial locations of features that are used in images as a heatmap. Evaluation metrics for CAMs are also needed to prove their effectiveness.

GradCAM [9] is a recent development in CAM generation. The method takes an average of the gradients after a target class is passed into the CNN from the last Convolutional layer. This creates a CAM that is adequate at locating important regions in images for predicting the target class. Another development, GradCAM++ [10], supplies better visual explanations by localizing around multiple objects in images. The method uses a weighted combination of positive partial derivatives of the last Convolutional layer with respect to target class. This allows for better visualizations, especially when multiple objects need to be located. GradCAM++ performs the best for multiple objects of the same class when comparing to other methods like Layer-CAM [11], HiResCAM [12], and FullGrad [13]. Therefore, as the number of spray deposits in our images is not known, it would be best to use a method that performs well with multiple objects.

The performance of CAMs has been evaluated using various metrics, including deletion, insertion, and stability [14], [15]. Deletion and insertion are two complementary metrics used to assess the quality of an explanation. Deletion measures the change in classification confidence when different regions of an image are removed, while insertion measures the change in confidence when different regions are added with surrounding noise or with no surrounding context. Stability is another metric used to compare different CAM explanations for the same class. This metric involves adding random uniform noise to an input image and generating a new CAM, and then computing the average distance between the original CAM and the noisy input CAM using L1 distance. These metrics have been used to evaluate the performance of CAMs in a variety of applications and will be used.

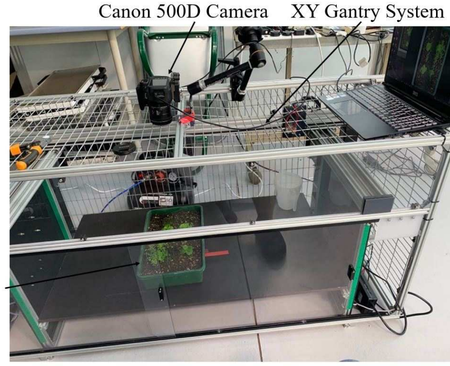


Fig. 1: XY Sprayer with tray of lettuces and weeds.

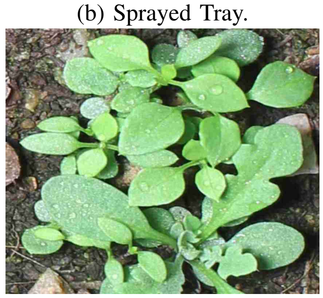
III. Experimental Setup and Dataset Description

The XY Sprayer is an experimental spraying system. Figure 1 shows the system, it uses a gantry XY system with a removable floor to change the spraying height. A Canon 500D camera is attached to the system to capture images. The spraying height from the spray plate to the tray bed is 30 cm whilst the distance from the camera lens to the tray bed is 45 cm. Spray deposits were completed with a pressure of 3 bar with a spray time of 8 ms which was recommended by Syngenta.

(a) Dry Tray.

(b) Sprayed Tray.

(c) Dry Tray region.



(d) Sprayed Tray region.

Fig. 2: Comparison of dry tray, Figure 2a, and sprayed tray, Figure 2b, as well as the chickweed region in both dry, Figure 2c, and sprayed Figure 2d.

A. Lettuce Trays

To simulate realistic data, Syngenta transplanted partially grown lettuces into trays with an even spacing and sowed commonly found weeds randomly among the trays. In Table I is the Biologische Bundesanstalt, Bundessortenamt und CHemische Industrie (BBCH) breakdown.

TABLE I: BBCH Scale for plants used.

Latin Name	EPPO Code	Common Name	BBCH
Poa annua	POAAN	Annual Meadowgrass	10-13
Stellaria media	STEME	Chickweed	10-22
Lactuca sativa	LACSA	Lettuce	19

B. Data Collection Procedure

Trays were placed into the system so that the Canon 500D camera could capture each tray including the corners of the tray. The system then was used by an expert human controller to target all weeds to be sprayed once using the software provided. In total 89 images were taken for both dry and sprayed trays. Therefore, the overall dataset is 178 images. Shown in Figure 2 is a comparison of the same tray before and after spraying with a specific region zoomed in. As shown by this example, it is a visually complex recognition problem given variability of background and very small deposits which are hardly visible.

C. Data Pre-processing: Augmentation

After data collection, images were labelled and split using a 70%, 20%, and 10% ratio for training, validation, and test, respectively. To achieve robust and generalized CNNs, large amounts of data are needed. However, sprayed lettuces and weeds data are not available to our knowledge from public repositories. Therefore, multiple augmentations are applied to our dataset to increase the data available for training CNNs. Augmentations include horizontal and vertical flips, rotations from 15 to 45 degrees, blurring the image with a box blur with a factor of 3 times, increasing and decreasing the brightness by 30%, and increasing the contrast by 30%. Thus, our training split has increased from 122 images to 854.

IV. Pipeline

An XAI pipeline for a binary classification task has been implemented, a number of modular stages were employed in the proposed pipeline. Each CNN was customized for the classification task by changing the classification head. The pipeline has an interpretability module using GradCAM++ [10]. Figure 3 outlines all the major components and in the next subsections implementation details are provided.

A. Stratification and Sampling

When considering the overall problem of identifying sprayed weeds or lettuces, it could be assumed that this problem is essentially pixel-wise change detection. Therefore, to ensure a CNN can learn these pixel-wise differences a strategy has been devised for this as well as an additional sampling method to observe and learn those differences. We stratify our dataset by ensuring that if image N is a dry version of the image available to the train sample, then image N', the sprayed version, is also present in the sample. Intuitively, this is how a human may distinguish between the two types of image, much like spotting a difference.

As our dataset is stratified, images have the dry and sprayed instances for the same tray in each split. However, as CNNs are trained with batches of images, without sampling we cannot ensure that the dry and sprayed instances are used within the same batch. Therefore, to ensure in each

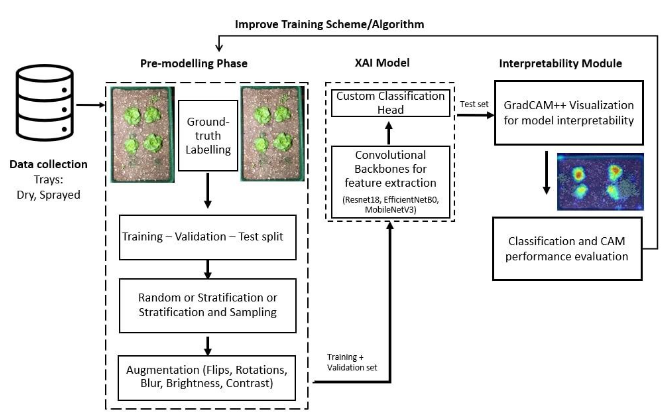


Fig. 3: XAI Pipeline for Classification and CAM evaluation.

batch images are loaded as matches. This includes loading the match even if it is augmented. This has been applied this to the training split of our data only.

B. Network Architectures

Experiments with three pre-trained CNNs, EfficientNet-B0 [16], MobileNetV3 [17], and ResNet18 [18] have been completed. All networks are pre-trained on the ImageNet [19] dataset. The choice of CNNs is informed by their successful use and deployment in the agri-robotics domain [20]. Lightweight networks are chosen, so that in the future they can be deployed onto practical spot sprayer systems.

C. Classification Metrics

To evaluate our networks, F1-score, and Area Under the Receiver Operating Characteristic (AUROC) are used. The definitions and formulas of F1-score and AUROC can be found within literature [21].

D. CAM Metrics

CAMs are generated heatmaps of spatial locations that create regions of interest to help improve interpretability within CNNs used and to differentiate similar scoring CNNs. To evaluate the CAMs generated and features identified within the CAMs Deletion, Insertion, and Stability are used [14], [15].

For Deletion and Insertion, the confidence of the CNNs prediction will be recorded with deletion and insertion increasing by 1% until the entire image is deleted or inserted. After plotting the confidence values against the amount of image deleted or inserted, the area under the curve (AUC) is calculated using the Trapezoidal Rule:

$$AUC = \frac{h}{2}\left[y_0 + 2\left(y_1 + y_2 + y_3 + \cdots + y_{n-1}\right) + y_n\right] \quad (1)$$

where y is the prediction confidence, n is equal to the number of plotted points, and h is equal to the increase in deletion or insertion change. Therefore, Deletion scores that are lower are better and Insertion scores that are higher are better. An average AUC will be taken from the entire test set and reported in Section V.

Stability measures the similarity between CAMs. Given a CAM for an image noise can be added to the original image and a new CAM can be computed. The noise added is uniform and the same throughout all images. These are then compared using L1 distance. L1 distance is:

$$D_{L1}(\mathbf{A}, \mathbf{B}) = \sum_{i=1}^{n} |A_i - B_i| \quad (2)$$

where n is the total number of pixels in the images, A(i) and B(i) are the pixel values of image A and B at position i, respectively. A lower Stability would indicate a similar CAM. This is completed for all test images and an average score is taken. The results are reported in Section V.

V. Results and Performance Evaluation

To demonstrate the usefulness of our proposed stratification and sampling methodologies a traditional random split has been implemented to be used as a baseline to compare to stratification without sampling and then stratification with sampling. Table II shows the results for each network with the F1-score, AUROC scores, Deletion scores, Insertion scores, and Stability scores. The baseline is shown as *Random* as it uses a random split for the data, *Stratified* is the stratified dataset, and *Stratified (Train Sampling)* shows the results for both stratified data split and train sampling. The best scores for each model are in bold.

The classification scores show the best overall models are the stratified train sampling EfficientNet-B0 and ResNet18. Both score an F1-Score 94.1% for dry, 93.3% F1-Score for sprayed, and 93.7% AUROC. When looking at the CAM metric scores, the best model for Deletion is the stratified MobileNetV3 with 33.6%. The best Insertion score was from the stratified train sampling MobileNetV3 with 55.0%. Finally, when considering Stability the best score was from the random split ResNet18 with 24.1%.

Comparing to the baseline, stratification and train sampling greatly improve the F1-score for both classes, the AUROC scores, and the Insertion scores for all models. Considering the F1-scores, there are increases of 9.9%, 11.8%, and 76.9% for the dry class for the EfficientNet-B0, ResNet18, and MobileNetV3, respectively. For the sprayed class, F1-scores increase by 16.4%, 13.3%, 17.6% for the EfficientNet-B0, ResNet18, and MobileNetV3, respectively. The AUROC scores for the EfficientNet-B0 and ResNet18 improve by 12.5% and the MobileNetV3 improves by 31.2%. Furthermore, Insertion improves greatly with 9% and 39.4%, for the ResNet18 and MobileNetV3, respectively. With the Deletion metric, the best model is MobileNetV3 decreasing by 13% from the baseline. However, in the case of Deletion the baseline appears better for both the EfficientNet-B0 and ResNet18. Further testing is needed to fully understand why this metric behaves differently.

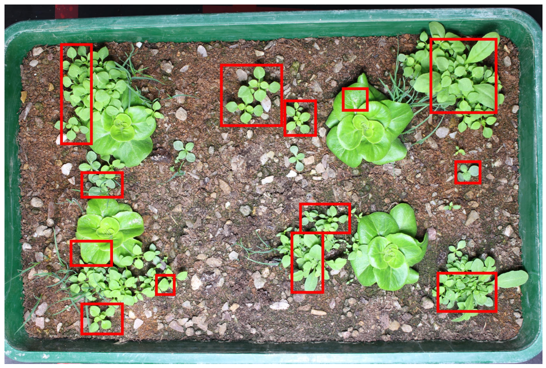
Improvements are made between stratification and the addition of train sampling when considering the insertion metric. The Insertion for the ResNet18 and MobileNetV3 increase by 39.3%, and 45.8%, respectively.

The Stability scores for all models in all methods of data splitting are very similar, showing that explanations generated are consistent. The best model Stability scores are 24.7%, 24.2%, and 24.1% for the stratified split EfficientNet-B0, random split MobileNetV3, and random split ResNet18, respectively. The Stability for all data splits and models is very similar within a range of 0.7%.

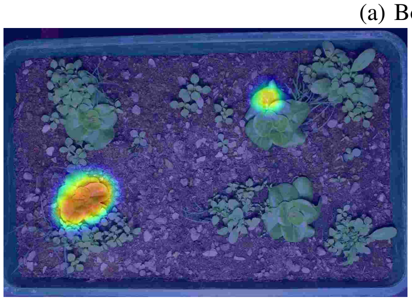
Figure 4 shows the CAMs for one of the test images within all splits from all models implemented as well as a bounding box ground truth of spray deposits. The visualization of spray deposits reinforces the results seen in Table II. When inspecting the CAMs for the EfficientNet-B0 there is improvement visually. For example, the stratified CAM in Figure 4e has an additional lower interest region when compared to the

TABLE II: Classification and CAM Results

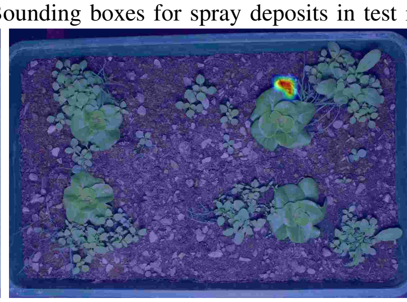
	Split	F1-Score Dry (%)	F1-Score Sprayed (%)	AUROC (%)	Deletion (%)	Insertion (%)	Stability (%)
EfficientNet-B0	*Random*	84.2	76.9	81.2	**60.0**	51.0	24.9
	Stratified	88.8	85.7	87.5	68.6	50.8	**24.7**
	Stratified (Train Sampling)	**94.1**	**93.3**	**93.7**	71.7	**51.4**	24.9
MobileNetV3	*Random*	0.0	66.6	50.0	46.6	25.6	**24.2**
	Stratified	62.5	62.5	62.5	**33.6**	9.2	24.8
	Stratified (Train Sampling)	**76.9**	**84.2**	**81.2**	63.2	**55.0**	24.9
ResNet18	*Random*	82.3	80.0	81.2	**51.8**	29.0	**24.1**
	Stratified	**94.1**	**93.3**	**93.7**	52.3	1.3	24.6
	Stratified (Train Sampling)	**94.1**	**93.3**	**93.7**	53.2	**38.0**	24.9



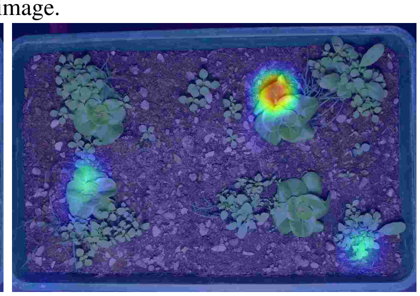
(a) Bounding boxes for spray deposits in test image.



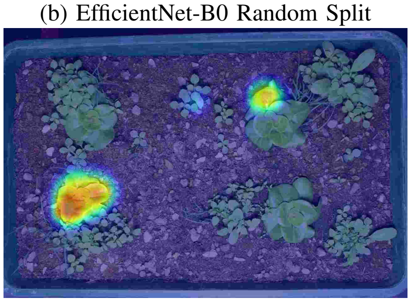
(b) EfficientNet-B0 Random Split



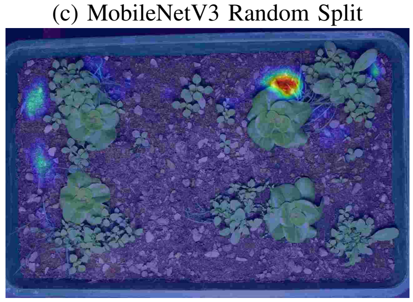
(c) MobileNetV3 Random Split



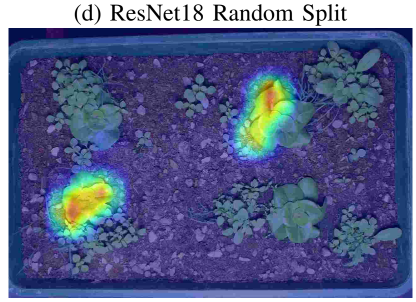
(d) ResNet18 Random Split



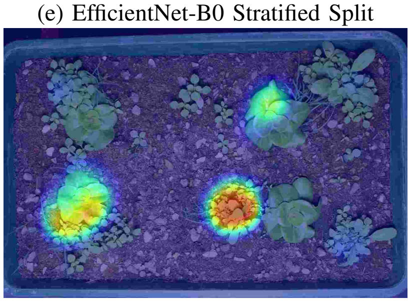
(e) EfficientNet-B0 Stratified Split



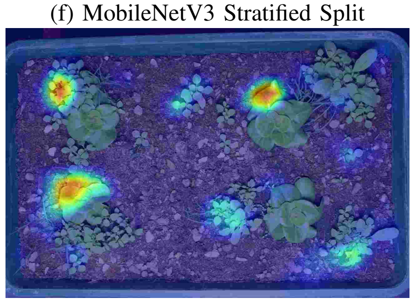
(f) MobileNetV3 Stratified Split



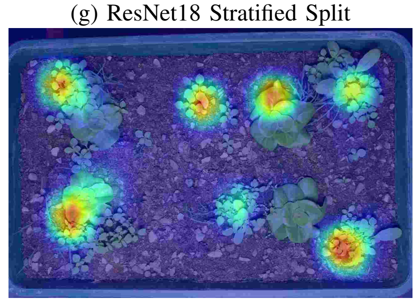
(g) ResNet18 Stratified Split



(h) EfficientNet-B0 Train Sampling Split



(i) MobileNetV3 Train Sampling Split



(j) ResNet18 Train Sampling Split

Fig. 4: Class Activation Maps for all models for sprayed test image.

baseline in Figure 4b. More detections can be found for the train sampling CAM in Figure 4h but the added 3 regions are much lower in intensity, these are difficult to see but can be found in the top left, top right and bottom right clusters of weeds. When comparing Figure 4h to the ground truth in Figure 4a each region generated contains a spray deposit. However, not all spray deposits are found.

The MobileNetV3 CAMs in Figure 4c, Figure 4f, and Figure 4i, indicate a significant increase in the detection of spray deposits compared to the baseline random split. The stratified CAM is noisier and potentially less effective than the random split CAM, whereas the train sampling includes multiple detections of spray deposits, even in the lower interest regions. When matching the regions generated to the ground truth, most of the detections for the train sampling are on the ground truth. When looking at the random split and the stratified split both have regions of interest that don't entirely overlap with spray deposits.

Finally, in the ResNet18 CAMs, improvements are made visually from the random split, Figure 4d, to higher intensity areas in the stratified split in Figure 4g. The largest improvement is seen in the train sampling CAM in Figure 4j as it has the most detections. When comparing to the ground truth all CAMs generated for the ResNet18 include regions that have spray deposits. The best CAM Figure 4j, has the most number of detections that overlap with the ground truth.

VI. Conclusion

Overall, the results are promising. Improvements have been made from a baseline using a stratification and sampling methodology by a minimum of 12.5% AUROC across multiple CNNs. Our results indicate that this methodology could be used as an automated assessment replacement for existing agricultural methods.

Visually, CAMs can be seen to improve with the addition of stratification and then further with sampling as discussed in section V. This is reinforced by the CAM metric results as the Insertion scores for all models is the highest when using both proposed methods. Furthermore, explanations are consistent as Stability across all data splits is very close. Therefore, CAMs that are generated are also robust.

For our future work, we are going to develop a deposit deposition detection scenario for the collected dataset to detect the quantity of applied spray deposits on both weeds and lettuces. This is to give further insight to the effectiveness of spraying systems. Alongside this, the methodology will be adapted to work with moving systems.

References

[1] E. Commission, "The european green deal," in *European Commission*, 2019. [Online]. Available: https://eur-lex.europa.eu/resource.html?uri=cellar:ea0f9f73-9ab2-11ea-9d2d-01aa75ed71a1.0001.02/DOC_1&format=PDF

[2] A. Srinivasarao, T. Kumar, R. Parray, H. Kushwaha, and I. Mani, "Spray droplet deposition, collection, and analysis techniques: A review," *Pharma Innov. J.*, vol. 10, pp. 158–165, 2021.

[3] V. Partel, S. Charan Kakarla, and Y. Ampatzidis, "Development and evaluation of a low-cost and smart technology for precision weed management utilizing artificial intelligence," *Computers and Electronics in Agriculture*, vol. 157, pp. 339–350, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168169918316612

[4] J. Liu, I. Abbas, and R. S. Noor, "Development of deep learning-based variable rate agrochemical spraying system for targeted weeds control in strawberry crop," *Agronomy*, vol. 11, no. 8, p. 1480, 2021.

[5] C. Tan, F. Sun, T. Kong, W. Zhang, C. Yang, and C. Liu, "A survey on deep transfer learning," in *International conference on artificial neural networks*. Springer, 2018, pp. 270–279.

[6] A. Salazar-Gomez, M. Darbyshire, J. Gao, E. I. Sklar, and S. Parsons, "Beyond map: Towards practical object detection for weed spraying in precision agriculture," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 9232–9238.

[7] L. Wang, W. Song, Y. Lan, H. Wang, X. Yue, X. Yin, E. Luo, B. Zhang, Y. Lu, and Y. Tang, "A smart droplet detection approach with vision sensing technique for agricultural aviation application," *IEEE Sensors Journal*, vol. 21, no. 16, pp. 17 508–17 516, 2021.

[8] S. Gao, G. Wang, Y. Zhou, M. Wang, D. Yang, H. Yuan, and X. Yan, "Water-soluble food dye of allura red as a tracer to determine the spray deposition of pesticide on target crops," *Pest management science*, vol. 75, no. 10, pp. 2592–2597, 2019.

[9] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, oct 2019. [Online]. Available: https://doi.org/10.1007%2Fs11263-019-01228-7

[10] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-CAM: Generalized gradient-based visual explanations for deep convolutional networks," in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, mar 2018. [Online]. Available: https://doi.org/10.1109%2Fwacv.2018.00097

[11] P.-T. Jiang, C.-B. Zhang, Q. Hou, M.-M. Cheng, and Y. Wei, "Layercam: Exploring hierarchical class activation maps for localization," *IEEE Transactions on Image Processing*, vol. 30, pp. 5875–5888, 2021.

[12] R. L. Draelos and L. Carin, "Use hirescam instead of grad-cam for faithful explanations of convolutional neural networks," 2020. [Online]. Available: https://arxiv.org/abs/2011.08891

[13] S. Srinivas and F. Fleuret, "Full-gradient representation for neural network visualization," *Advances in neural information processing systems*, vol. 32, 2019.

[14] V. Petsiuk, A. Das, and K. Saenko, "Rise: Randomized input sampling for explanation of black-box models," 2018. [Online]. Available: https://arxiv.org/abs/1806.07421

[15] U. Bhatt, A. Weller, and J. M. F. Moura, "Evaluating and aggregating feature-based model explanations," 2020. [Online]. Available: https://arxiv.org/abs/2005.00631

[16] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.

[17] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan *et al.*, "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1314–1324.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[19] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[20] M. H. Saleem, J. Potgieter, and K. M. Arif, "Automation in agriculture by machine and deep learning techniques: A review of recent developments," *Precision Agriculture*, vol. 22, no. 6, pp. 2053–2091, 2021.

[21] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International journal of data mining & knowledge management process*, vol. 5, no. 2, p. 1, 2015.