An *a contrario* approach for plant disease detection

Rebecca Leygonie
rebecca.leygonie@etu.u-paris.fr

Sylvain Lobry
sylvain.lobry@u-paris.fr,sylvainlobry.com

Laurent Wendling
laurent.wendling@u-paris.fr

Université Paris Cité, LIPADE
F-75006 Paris, France

Abstract

Detecting plant diseases or abnormalities is not a trivial task, as they can be caused by multiple factors such as environmental conditions, genetics, pathogens, etc.

Because there is a need to help farmers make decisions to maximize crop yields, many studies have emerged in recent years using deep learning on agricultural images to detect plant diseases, which can be considered as an anomaly detection task. However, these approaches are often limited by the availability of annotated data or prior knowledge of the existence of an anomaly.

We propose an approach that can detect part of the anomalies without prior knowledge of their existence, thus overcoming some of these limitations. To this end, we train a model on an auxiliary prediction task (plants' age regression). We then use an explicability model to retrieve heatmaps whose distributions are studied. For each new observation, we propose to study how closely its heatmap follows the desired distribution and we derive a score indicating potential anomalies. Experiments on the GrowliFlower dataset indicate how our proposed method can help potential end-user to automatically find anomalies.

1 Introduction

Agriculture, and in particular the management of plant diseases [3, 30] are crucial issues for the economy and survival of civilizations [19] as they result in the maintenance of agricultural productivity, food security and environmental sustainability.

The use of deep learning, and in particular computer vision models, is becoming increasingly common for the detection of plant pathologies, but comes up against a major problem: the lack of annotated data [14]. Indeed, as in all anomaly detection use cases, "abnormal" data, in this case diseased plants, represent a minority in the dataset. It is therefore difficult to train supervised models because of potential over-fitting. Moreover, anomalies may be hardly perceptible visually, which complicates the classification task. Figure 1 shows an example of two cauliflowers that look similar, yet the one on the left has many defects (3 which

Anomalies :
1.   dead plant
2.   limp
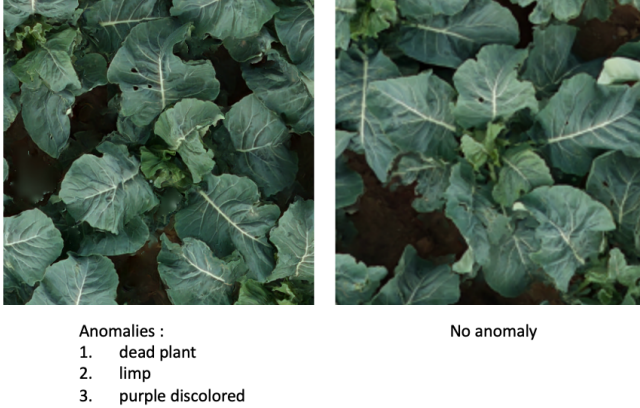3.   purple discolored

No anomaly

Figure 1: Two cauliflowers at the same age. The one on the left is abnormal and the one on the right is normal. Our objective is to automatically identify abnormal cauliflowers through the proposed *a contrario* based approach.

are identified by an expert) and the one on the right is healthy. Finally, normal and abnormal plants may present strong intra-class variations that might be hard to learn in supervised models and prevent from generalization.

The field of anomaly detection is divided into two types of approaches: visual anomaly detection at the image level, to define whether the image is normal or abnormal, and visual anomaly detection at the pixel level, used to locate abnormal areas in an image.

Approaches for detecting anomalies at image level can be divided into 4 categories [29]: **density estimation approaches**, such as kernel density estimation method [18], which estimate the distribution pattern of normal images and then calculate the probability of a new image with respect to the established distribution; **one-class classification approaches**, such as *SVDD* [26], that learn to extract features of importance from a class and then compares them with those of a new image to identify how similar or dissimilar they are; **reconstruction-based approaches**, such as *AnoGAN* [20], are based on the assumption that by training a generative model (such as VAEs [12] or GANs [4]) on normal data, it will not be able to reconstruct abnormal images; and **self-supervised models**, such as *CutPaste* [13], which learn a generalizable representation from unlabeled data by solving a supervised substitution task that is often unrelated to the target task.

When the problem of anomaly detection is such that we do not have labels on the presence of anomalies, it is difficult to use learning-based approaches trained on any set of images. For this reason, *a contrario* approaches [2, 16] can be effective. The principle of *a contrario* approaches is to design a reference model, considered as a null hypothesis representing normality. We then look at the probability of a new observation belonging to the reference model. After thresholding this probability, we can take a decision on whether this sample is normal or abnormal. This approach has proved its effectiveness, particularly for localizing anomalies in images [25].

Our approach aims to use an *a contrario* approach by working with the heatmaps obtained by an explicability model, revealing the features of importance in an image. To do this, we exploit the ability of a model to extract the representative characteristics of a plant,

by training it on a pretext task, the prediction of age, and then recovering the heatmaps linked to the model's predictions. We then propose an *a contrario* approach to determine whether a new image is abnormal or not, by comparing it with the reference distribution. This idea is based on the assumption that an abnormal plant should not have the same characteristics of importance as a normal plant.

Many agricultural-related projects incorporate explicability models in their pipelines. The authors of [10] detect and localize onion blight symptoms using a weakly supervised model trained from image-level annotations, and localize symptoms by thresholding the model's activation map obtained by CAM [32]. In [9], the authors detect plant diseases with an unsupervised model combining observations with normalization flows, a visual saliency map and position encodings. The authors of [1] propose CountNet, a weakly supervised model that counts fruits and flowers in unstructured environments. The model learns from image-level annotations with the number of objects as input, without explicitly specifying the nature and location of the object. The authors use two explicability models, Score-CAM [28] and Guided Backpropagation [24], to demonstrate that the network does indeed examine flower/fruit features to count. In [11], the authors train a classification model to classify 3 crop zones, then use class activation maps (CAMs) [32] to segment the image and retrieve this information as annotation to automate a harvester. Finally, the authors of [7] use Grad-CAM [23], OSM [31], and LIME [17] to retrieve heatmaps in inference from a model trained to detect whether a cauliflower is ready to be harvested or not, they then cluster on the heatmaps to understand the import characteristics for each class. This approach is used to improve model performance and as a decision aid for farmers harvesting cauliflowers.

Our work aims to use heatmaps as new data, assumed to be less noisy and to indicate relevant areas, to apply an *a contrario* approach in order to determine if a cauliflower should be considered abnormal. We calculate a score based on the probability that the heatmaps of cauliflowers at different ages belong to the distributions of normal cauliflowers heatmaps at the same ages and compare it to a predefined threshold (the different stages of our approach are shown in Figure 2). Our approach requires a sample of images that are assumed to be normal, without any assumptions about possible anomalies in the rest of the data.

2   Methodology

2.1   Problem definition

Let $X = [x_1, x_2, ..., x_T]$ be a temporal series of T images of size $(L \times L)$ pixels. Our goal is to detect whether X contains an anomaly using an *a contrario* approach based on heatmaps obtained by an explainability model. To achieve this, we train a predictive model on an auxiliary task (see subsection 3.2). We propose to apply an explainability model, Grad-CAM [23] that highlights the important areas of the image through heatmaps, indicating the regions that contributed to the model's prediction. We obtain the temporal series of heatmaps $H = [h_1, h_2, ..., h_T]$ where each heatmap is a matrix $h_{t_{(i,j)}}$ for $i, j \in [1, L]^2$ and $t \in [1, T]$ corresponding to the image x_t. These heatmaps provide information about the areas of importance in images for the auxiliary task. Our assumptions is that these areas are different for normal and abnormal observations.

We aim to calculate a score to determine if h_t is abnormal by estimating the probability that each pixel is drawn from the pixel distribution (see subsection 2.2) obtained from heatmaps of normal images at time t of the image time series.
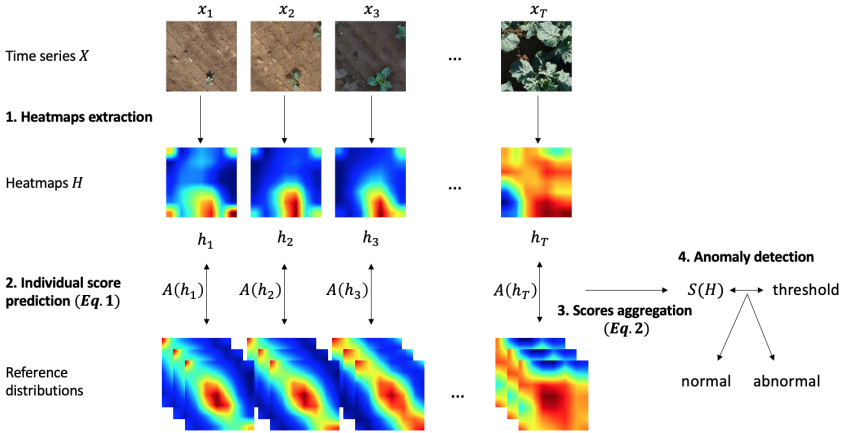
Figure 2: Pipeline of the proposed anomaly detection approach. For each new observation, consisting of a time series of images of a single plant at different ages, we retrieve the heatmaps associated with the different time steps, by applying Grad-CAM (step 1) to a model trained on an auxiliary task, cauliflower age prediction. We then compare each heatmap to the reference distribution of heatmaps of healthy plants of the same age. We calculate an individual score $A(h)$ for each image (see Eq. 1, step 2). These individual scores are aggregated on the time series to obtain the $S(H)$ score (see Eq. 2, step 3). We then compare $S(H)$ with a predefined threshold to determine whether the cauliflower can be considered abnormal (step 4).

In other words, for an image of timestep t, for each pixel $h_{t_{(i,j)}}$, we calculate the probability that it is drawn from the reference distribution denoted

$$\tilde{H}_{t_{(i,j)}} = [\tilde{h}_{t,1_{(i,j)}}, \tilde{h}_{t,2_{(i,j)}}, \tilde{h}_{t,3_{(i,j)}}, ..., \tilde{h}_{t,n_{(i,j)}}],$$

which represents the pixel values $\tilde{h}_{t_{(i,j)}}$ of n heatmaps obtained from normal images at timestep t. We compute $A(h_t)$, the average of the probabilities that each pixel is drawn from the reference distribution of pixels i, j at timestep t:

$$A(h_t) = \frac{\sum_{i,j=1}^{L} P(h_{t_{(i,j)}} | \tilde{H}_{t_{(i,j)}})}{L^2}. \tag{1}$$

For each time series H of T heatmaps, we calculate a score $S(H)$:

$$S(H) = \frac{\sum_{t=1}^{T} A(h_t)}{T}, \tag{2}$$

which represents the normality score of a heatmap time series. Please note that the score is not a probability, but rather an arithmetic mean of probabilities, as we cannot assume that the pixels are independent. This $S(H)$ score can then be thresholded to determine if the time series X is abnormal.
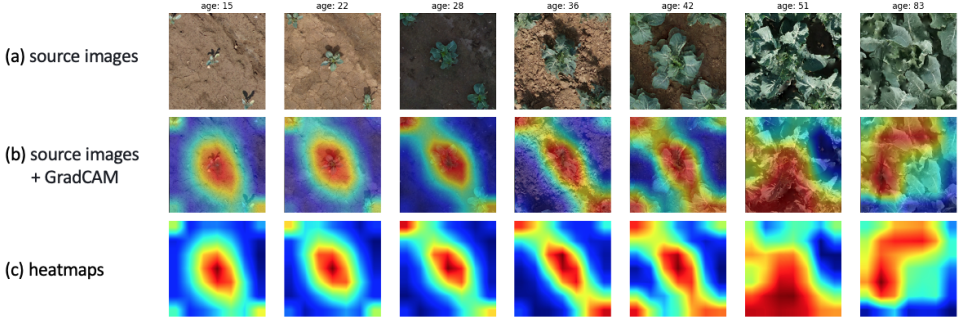
Figure 3: Transformations of an image considered in this study (a) Original Images, (b) Grad-CAM Application after Model Training, (c) Heatmaps Obtained by Grad-CAM

2.2 Calculation of the probability

For the computation of $A(h_t)$ (Eq. 1), we aim as the probability that a pixel in an image belongs to the reference distribution of the pixel at the same position in the normal images. We estimate this probability from the reference distribution $\tilde{H}_{t_{(i,j)}}$. Let h_t be a new observation. We calculate $P(h_{t_{(i,j)}}|\tilde{H}_{t_{(i,j)}})$, the estimated probability that the point $h_{t_{(i,j)}}$ is drawn from the distribution of $\tilde{H}_{t_{(i,j)}}$, using the following process:

1. **Histogram calculation:** Let b be the number of bins uniformly sampled in the interval $[0,1]$ (which is the range of values of the heatmaps). For $1 \leq z \leq b$, let B_z represent the z-th bin, i.e., $B_z = \left[\frac{z-1}{b}, \frac{z}{b}\right[$ for $1 \leq z \leq b-1$, and $B_b = \left[\frac{b-1}{b}, 1\right]$. Let d_z be the probability density in bin B_z. We have:

$$d_z = \frac{1}{n \cdot |B_z|} \sum_{i=1}^{n} I_{B_z}(\tilde{h}_{t_{(i,j)}}),$$

where $I_{B_z}(\tilde{h}_{t_{(i,j)}})$ is the indicator function that equals 1 if $\tilde{h}_{t_{(i,j)}}$ is in bin B_z, and 0 otherwise. $|B_z|$ represents the bin width, and n is the total number of observations.

2. **Probability estimation:**

   Let z be the index of the bin in which the point $h_{t_{(i,j)}}$ falls: $z = \left\lfloor h_{t_{(i,j)}} \times b \right\rfloor + 1$. The estimated probability that the point $h_{t_{(i,j)}}$ belongs to the distribution $\tilde{H}_{t_{(i,j)}}$ is then $P(h_{t_{(i,j)}}|\tilde{H}_{t_{(i,j)}}) = d_z \cdot |B_z|$.

3 Experiments and results

3.1 GrowliFlower Dataset

The GrowliFlower dataset [8] contains georeferenced time series based on drone images of two cauliflower fields acquired in 2020 and 2021. We use a subset of the dataset called *GrowliflowerR* which contains RGB orthophotos and phenotypic characteristics collected on

740 plants, including information on whether cauliflowers are affected by a disease or defect. The in-situ data propose the definition of 19 defects, but in practice the defects have variants, explained verbatim in the files corresponding to each field. For the purposes of this study, since our aim is to detect anomalies but not to characterize them, we consider all defects to be anomalies.

We manually split the dataset into *Train*, *Validation* and *Test* sets, containing 60%, 30% and 10% of the cauliflower image time series respectively. Our auxiliary task is age regression (see Figure 3(a) which represents a series of images with associated ages), defined by the difference in days between the date the photo was taken and the date the cauliflower was planted (from 1 to 93 days).

3.2 Auxiliary task: plant age prediction

In this study, we trained a *ResNet18* model [5] to predict the age of cauliflowers. The hyper-parameters were set to optimize the model's performance on the validation set. The model was trained over 200 epochs with an initial learning rate fixed at 0.1. We used stochastic gradient descent with a weight decay of 0.2 with the Huber loss [15]. We use a learning rate scheduler to divide the learning rate by a factor of 10 if the loss is not reduced for 5 consecutive epochs. The model obtains a mean square error (MSE) of 30.76 and an R^2 score of 0.95 on the test data.

3.3 Anomaly detection

Once the age prediction model has been trained, we apply Grad-CAM to the predictions made by the model, highlighting areas of importance (see Figure 3(b) superimposing the original images with the heatmaps obtained by applying Grad-CAM). We retrieve the heatmaps (see Figure 3(c) series of heatmaps obtained with Grad-CAM) to apply our anomaly detection method.

We collect all normal observations from the training and validation datasets and define our age-based reference distributions. The dataset on which we aim to detect anomalies includes the *Test* dataset, as well as all the abnormal data from the *Train* and *Validation* datasets, i.e. 64% abnormal cauliflowers and 36% normal cauliflowers.

3.4 Model comparison

We compare our approach with two other anomaly detection approaches: **f-AnoGAN** [21], a self-supervised reconstruction-based approach, which learns to reconstruct normal images. f-AnoGAN detects anomalies by calculating a score by combining the residual error of the discriminator features and the image reconstruction error; and **One-Class Support Vector Machines (OCSVM)** [22], an unsupervised model of the one-class classification family, which learns the boundary of normal data points and identifies data outside this boundary as anomalies.

We train the f-AnoGAN model with the same parameters as the authors of the original paper [21]. Considering OCSVM, we chose to use an "rbf" kernel, with $gamma = 0.001$ and $nu = 0.03$ after comparing different parameters. f-AnoGAN calculates a score per image, so we average the scores to obtain a score per cauliflower. The OCSVM model, on the other hand, does not give an anomaly score, but classifies data according to their position relative to a decision boundary. In order to have a class for each cauliflower, we take the class most
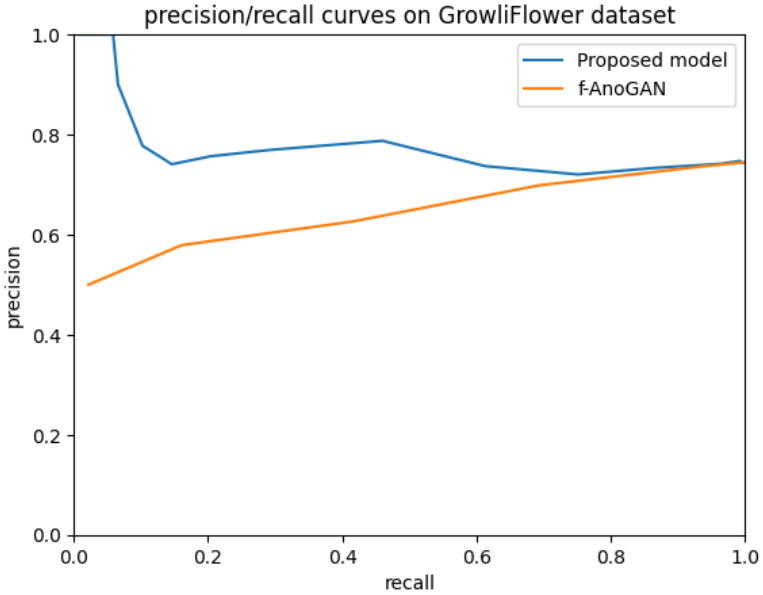
Figure 4: Precision-recall curves for the proposed model and f-AnoGAN.

Model	AUC	F1-score	precision	recall
Proposed model	**0.76**	0.85	0.74	1.0
f-AnoGAN	0.63	0.85	0.74	1.0
OCSVM	X	0.72	0.69	0.76

Table 1: Anomaly detection results. We report the Area Under Curve (AUC), the F1-score, the precision and the recall.

frequently attributed to the different time steps in the evolution of a cauliflower. Figure 4 shows the precision and recall of the different approaches tested. Note that OCSVM does not appear as the decision boundary cannot be adjusted to balance the precision and the recall. We report the different numerical results in Table 1.

4   Discussion

Auxiliary task   The core idea of our method is to use heatmaps as data for *a contrario* study of anomaly detection. For this, the prerequisite is to have a prediction model powerful enough to exploit its ability to extract features of importance in an image. We work with image time series dataset. Hence, we first tried to train a hybrid model (ResNet18 followed by an LSTM [6]) on the age prediction task so that the recurrent model could take into account the temporality of cauliflower evolution from the feature vectors obtained from the ResNet18 model. However, despite explicitly taking into account sequential data, it underperforms in our context. The age sequences are almost always identical, which leads the model to learn on the basis of these sequences rather than the images themselves, resulting in very poor performance (MSE of 316, which is a very high value given that the age to be predicted is in

the interval $[1, 93]$).

Comparison with other methods   For comparison, we evaluate f-AnoGAN, a model based on a reconstruction strategy. In this model, an image is assigned a high score if it diverges significantly from the expected reconstruction, suggesting that it is an anomaly. Observations are then classified as anomalous if their scores exceed a defined threshold. Figure 4 indicates that our model achieves a perfect precision for recall values up to 10%, enabling accurate detection of a fraction of anomalies. Conversely, the f-AnoGAN model assigns consistently low scores. It should be noted that this distribution of scores does not clearly distinguish normal from abnormal images, resulting in a decrease in precision and recall with increasing threshold. Furthermore, Table 1 shows that the $F1$-score is identical for our model and f-AnoGAN. This $F1$-score is obtained at a recall value of 1. However, our approach presents a superior value of AUC, a performance attributable to the high precision obtained at low recall. OCSVM shows lower performances in terms of precision and recall. These observations underline the advantage of an adjustable threshold, particularly in contexts where anomalies are hard to visually distinguish.

Limitations of the proposed approach   Our study is based on the hypothesis that the characteristics of importance are not the same for normal and abnormal cauliflowers, and that this information can be found in the heatmaps obtained by Grad-CAM (see the heatmaps of a healthy cauliflower in the first series of images in the Figure 5 and those of a cauliflower with anomalies in the second series). We calculate a score for each cauliflower to determine whether it is abnormal. To do this, we calculate the average of the $A(h_t)$ scores for t in $[1, T]$, representing the extent to which the image at age t follows the reference distribution of healthy images at the same age. Thus, in an exaggerated case, if $A(h_t)$ equals 1, the image is normal, and if $A(h_t)$ equals 0, it is abnormal. Since we calculate the final score $S(H)$ as an average of the $A(h_t)$ at each time step, and it's possible for a cauliflower to present an anomaly at just one time step, the value of the score can be close to the score of a normal cauliflower. Thus, our approach is better able to detect abnormal cauliflowers that contain anomalies at several time steps. The proposed approach therefore makes it possible to detect a small proportion of anomalies without prior knowledge, with a certain level of confidence depending on the chosen threshold.

Our approach does not result in a constant number of false alarms. However, we show that it is possible to set a small threshold below which all observations are correctly classified as anomalies. This finding is important, as it highlights the possibility of using a predefined threshold to successfully identify a proportion of anomalies. Despite the fluctuating number of false alarms, such a strategy can be useful in ensuring a minimum level of anomaly detection, providing valuable insights for future research and practical applications. In practice, we show that this score can be used as an on-field indicator to be manually reviewed. In addition, the superimposition of heatmaps on the original images, combined with information such as the age predicted by the model and the actual age (see Figure 5 which represents the original images as well as the heatmaps superimposed on the original images with the predictions made by the model), can be used by farmers to analyze the various factors linked to anomalies. In fact, knowing which areas of the image have enabled the model to make its prediction, whether correct or not, can help to understand the important characteristics of a plant, and thus identify diseases more quickly.
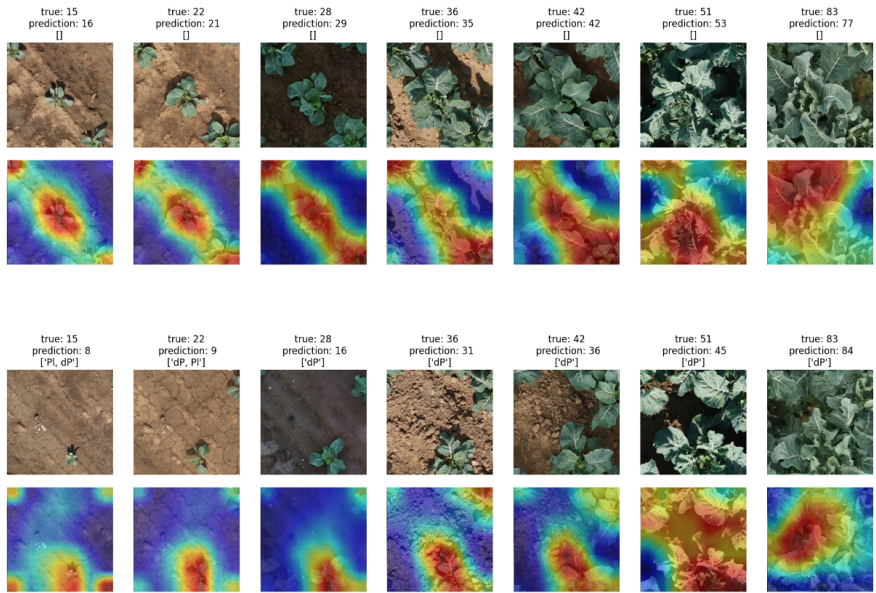
Figure 5: Time series of images of two cauliflowers, with and without heat map overlay, with, at each step, the age of the cauliflower, the age predicted by the model and associated annotations indicating the presence/nature of an anomaly. 'Pl' stands for "plant lying down" and 'dP' indicates a dead plant.

5   Conclusion

In this study, we propose a decision-support model for anomaly detection and demonstrate its effectiveness on a dataset of time series of cauliflower images. The proposed model leverages spatial information learned from an auxiliary task (in our case, plant age regression) and does not require prior assumptions about the presence of anomalies in the data. The only prerequisite is to have a sample of healthy data. The objective is to identify a subset of anomalies thereby assisting farmers in understanding plant phenotyping.

The results show that our approach is useful for detecting anomalies that are difficult to see visually. Indeed, it manages to accurately detect a small proportion of anomalies, unlike the f-AnoGAN and OneClass SVM models, which fail to capitalize on anomaly-related features.

In future work, we aim to test other explainability approaches such as LIME [17] or Guided Backpropagation [24] to evaluate their impact on the score computation and, subsequently, on anomaly detection.

In the dataset under study, the anomalies associated with cauliflowers do not exhibit temporal continuity. An anomaly may be present at a certain timestep t and then vanish, just as a new anomaly may surface at timestep $t + \Delta t$. Consequently, while our current work deals with time series of images, we treat each image independently due to this lack of continuous anomalies. We have not observed any continuity in the in-situ data indicating cauliflower-related anomalies. In our future research, we want to work with datasets that exhibits continuous anomalies from a certain time step t, and exploit the temporal aspect of

the images. More specifically, we wish to build on models such as BFAST [27], which would enable us to accurately determine when an anomaly manifests itself. In doing so, we aim to provide even more accurate and timely information to help farmers manage their crops more effectively.

References

[1] Uddhav Bhattarai and Manoj Karkee. A weakly-supervised approach for flower/fruit counting in apple orchards. *Computers in Industry*, 138:103635, 2022.

[2] Agnes Desolneux, Lionel Moisan, and Jean-Michel Morel. *From gestalt theory to image analysis: a probabilistic approach*, volume 34. Springer Science & Business Media, 2007.

[3] William E Fry. *Principles of plant disease management*. Academic Press, 2012.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) conference on computer vision and pattern recognition*, pages 770–778, 2016.

[6] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[7] Jana Kierdorf and Ribana Roscher. Reliability scores from saliency map clusters for improved image-based harvest-readiness prediction in cauliflower. *arXiv preprint arXiv:2305.15149*, 2023.

[8] Jana Kierdorf, Laura Verena Junker-Frohn, Mike Delaney, Mariele Donoso Olave, Andreas Burkart, Hannah Jaenicke, Onno Muller, Uwe Rascher, and Ribana Roscher. Growliflower: An image time-series dataset for growth analysis of cauliflower. *Journal of Field Robotics*, 40(2):173–192, 2023.

[9] Taejoo Kim, Hyeongjun Kim, Kyeonghoon Baik, and Yukyung Choi. Instance-aware plant disease detection by utilizing saliency map and self-supervised pre-training. *Agriculture*, 12(8):1084, 2022.

[10] Wan-Soo Kim, Dae-Hyun Lee, and Yong-Joo Kim. Machine vision-based automatic disease symptom detection of onion downy mildew. *Computers and Electronics in Agriculture*, 168:105099, 2020.

[11] Wan-Soo Kim, Dae-Hyun Lee, Taehyeong Kim, Hyunggun Kim, Taeyong Sim, and Yong-Joo Kim. Weakly supervised crop area segmentation for an autonomous combine harvester. *Sensors*, 21(14):4801, 2021.

[12] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

[13] Chun-Liang Li, Kihyuk Sohn, Jinsung Yoon, and Tomas Pfister. Cutpaste: Self-supervised learning for anomaly detection and localization. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9664–9674, 2021.

[14] Jiajia Li, Dong Chen, Xinda Qi, Zhaojian Li, Yanbo Huang, Daniel Morris, and Xiaobo Tan. Label-efficient learning in agriculture: A comprehensive review. *arXiv preprint arXiv:2305.14691*, 2023.

[15] Sylvain Lobry and Devis Tuia. Deep learning models to count buildings in high-resolution overhead images. In *2019 Joint Urban Remote Sensing Event (JURSE)*, pages 1–4. IEEE, 2019.

[16] David Lowe. *Perceptual organization and visual recognition*, volume 5. Springer Science & Business Media, 2012.

[17] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. " why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd The Association for Computing Machinery's Special Interest Group on Knowledge Discovery and Data Mining (ACM SIGKDD) international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.

[18] Murray Rosenblatt. Remarks on some nonparametric estimates of a density function. *The annals of mathematical statistics*, pages 832–837, 1956.

[19] Serge Savary, Andrea Ficke, Jean-Noël Aubertot, and Clayton Hollier. Crop losses due to diseases and their implications for global food production losses and food security. *Food security*, 4(4):519–537, 2012.

[20] Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International conference on information processing in medical imaging*, pages 146–157. Springer, 2017.

[21] Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Georg Langs, and Ursula Schmidt-Erfurth. f-anogan: Fast unsupervised anomaly detection with generative adversarial networks. *Medical image analysis*, 54:30–44, 2019.

[22] Bernhard Schölkopf, John C Platt, John Shawe-Taylor, Alex J Smola, and Robert C Williamson. Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7):1443–1471, 2001.

[23] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the international conference on computer vision*, pages 618–626, 2017.

[24] Jost Tobias Springenberg, Alexey Dosovitskiy, Thomas Brox, and Martin Riedmiller. Striving for simplicity: The all convolutional net. *arXiv preprint arXiv:1412.6806*, 2014.

[25] Matias Tailanian, Pablo Musé, and Álvaro Pardo. A multi-scale a contrario method for unsupervised image anomaly detection. In *2021 20th Institute of Electrical and Electronics Engineers (IEEE) International Conference on Machine Learning and Applications (ICMLA)*, pages 179–184. IEEE, 2021.

[26] David MJ Tax and Robert PW Duin. Support vector data description. *Machine learning*, 54:45–66, 2004.

[27] Jan Verbesselt, Rob Hyndman, Glenn Newnham, and Darius Culvenor. Detecting trend and seasonal changes in satellite image time series. *Remote sensing of Environment*, 114(1):106–115, 2010.

[28] Haofan Wang, Zifan Wang, Mengnan Du, Fan Yang, Zijian Zhang, Sirui Ding, Piotr Mardziel, and Xia Hu. Score-cam: Score-weighted visual explanations for convolutional neural networks. In *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE)/CVF conference on computer vision and pattern recognition workshops*, pages 24–25, 2020.

[29] Jie Yang, Ruijie Xu, Zhiquan Qi, and Yong Shi. Visual anomaly detection for images: A survey. *arXiv preprint arXiv:2109.13157*, 2021.

[30] Jan C Zadoks and Richard D et al. Schein. Epidemiology and plant disease management. *Epidemiology and plant disease management.*, 1979.

[31] Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. In *Computer Vision–European Conference on Computer Vision (ECCV) 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13*, pages 818–833. Springer, 2014.

[32] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. Learning deep features for discriminative localization. In *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE) conference on computer vision and pattern recognition*, pages 2921–2929, 2016.