Anomaly detection: detection of fruit defects

Grzegorz Cielniak & Adrian Salazar

GCielniak@lincoln.ac.uk & asalazargomez@lincoln.ac.uk

UNIVERSITY OF LINCOLN

Department of Computer Sciences & L-CAS & LAR
University of Lincoln

Introduction

"*Anomaly detection is the process of identifying data instances that significantly deviate from the majority of data instances*"

In agriculture we can use anomaly detection tools to detect pest and diseases....

"*In agriculture, anomaly detection is the process of identifying* **crops** *that significantly deviate from the* **normal and healthy crops**"

Why: Applications

- **Important for phenotyping breed disease resistant crops**
- **Anomalies mean significant losses to farmers**
  - Global yield losses projected to increase by 10 to 25% due to the increase pests in a warming climate[1]
  - Currently, pest and diseases reduce 10% to 30% the yield of the most important crops [2]
- **Traditional methods to early detect anomalies**
  - Visual assessment of expert agronomists
  - Prone to human errors, unattainable in large plantations, expensive

Problems: Identify anomalies

Anomaly data is:

- **Poorly sampled** → Hard to obtain in comparison with normal data

Most of the samples are normal

- **Hard to define** → Unlimited number of anomaly manifestations

Same anomaly? → Yes

- **Requires experts to label the data** → Not always available

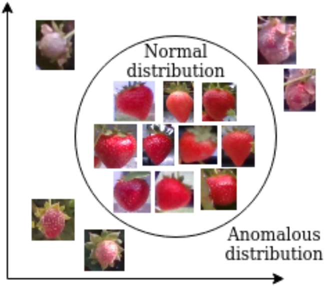
Is it an anomaly? → Yes, What kind of anomaly? → ???

Problems: Traditional methods are not made for anomaly detection

- **Supervised methods** do not work well to identify anomalies
  - Supervised models require a substantial number of training samples
  - Supervised require accurate labels
  - Supervised method are sensitive to unbalanced dataset
- **Pure unsupervised methods** do not work well to identify anomalies
  - Challenging to learn commonalities within data in a complex and high dimensional space
  - Unsupervised techniques techniques are very sensitive to noise
  - Often requires hyper-parameter tuning for optimal results

Solution → One class classification

- Recognise instances of a concept (normality) by only using examples of the same concept



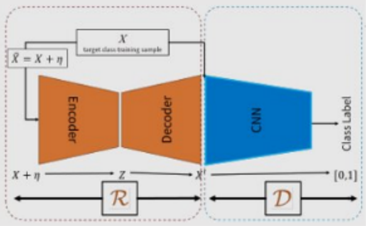
(a) One class classifier for anomaly detection

(b) Training data

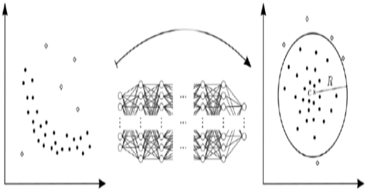
(c) Test data

Experiments

Two main frameworks to perform One-class classification:



(a) **Generic OC methods**:
Learn normality from an objective function that is **not focused on** anomaly detection
e.g. Autoencoers & GANs



(b) **Anomaly based OC methods**:
Learn normality from an objective function **focused on** anomaly detection
e.g. Deep SVDD & DOC

$$\Theta^*, \mathbf{W}^* = \underset{\Theta,\, W}{arg\ min} \sum_{x \,\epsilon\, \mathcal{X}} \ell\left(\psi(\phi(x;\Theta);\mathbf{W})\right),$$
$$s_x = f(x, \phi_{\Theta^*}, \psi_{\mathbf{W}^*})$$

$$\Theta^*, \mathbf{W}^* = \underset{\Theta,\, W}{arg\ min} \sum_{x \,\epsilon\, \mathcal{X}} \ell\left(f(\phi(x;\Theta);\mathbf{W})\right),$$
$$s_x = f(\phi(x;\Theta);\mathbf{W})$$

Where s_x is the anomaly score for image x, $f()$ is a function that maps a feature space into s_x , $\phi(,\Theta)$ in an encoding function, $\psi(,\mathbf{W})$ is a surrogate function that operates in the encoded feature space, and $\ell()$ is the objective function

Dataset

- Data from Riseholme poly tunnels **(2018)**
- **3 Training (Normality) sets:**
  - **normal: ripe**: 462 ripe strawberries
  - **normal: ripe + ripe occluded**: 462 ripe strawberries + 462 ripe occluded strawberries
  - **normal: ripe + unripe**: 462 ripe strawberries + 462 unripe strawberries
- **Test set** 110 anomalous strawberries + 110 instances from training set

Dataset Riseholme 2018

Some Results

Used models:

- Autoencoder for AD [3]
- Deep One-class transfer learning (DOC) [4]
- Deep Support Vector Data Description (Deep SVDD) [5]

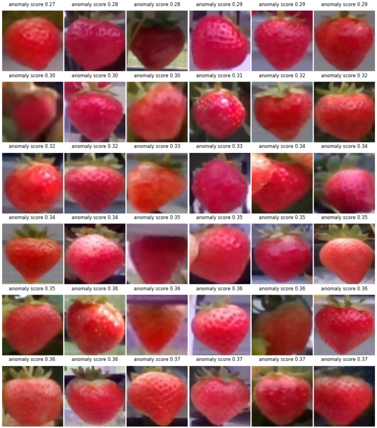
	Autoencoder	**Deep SVDD**	**DOC**
AUC (normal: ripe)	0.67	**0.95**	0.79
AUC (normal: ripe + ripe occluded)	0.59	**0.80**	0.63
AUC (normal: ripe + unripe)	0.58	**0.75**	0.52

Table 1: Results implemented methods

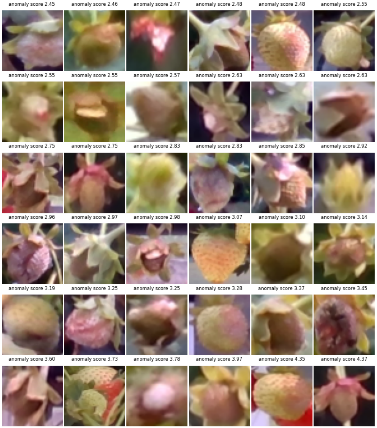
- Performance drops as normal data (training) is less concise e.g. ripe + unripe

Some Results:

Training set: Ripe $\rightarrow$ 0.95 AUC



(a) Deep SVDD, Most normal instances using ripe strawberries for training



(b) Deep SVDD, Most anomalous instances using ripe strawberries for training

Some Results: Deep SVDD

Training set: Ripe + Ripe Occluded $\rightarrow$ 0.75 AUC

(a) Deep SVDD, Most normal instances using ripe and unripe strawberries for training

(b) Deep SVDD, Most anomalous instances using ripe and unripe strawberries for training

Some Results: Deep SVDD

- **Model**: Autoencoder for AD [3]
- **Performance:** 0.67 AUC (Sensitive to parameters)

(a) **Normal** Better reconstruction Recons Loss: 0.63

(b) **Normal** Better reconstruction Recons Loss: 0.60

(c) **Anomaly** Worse reconstruction Recons Loss: 0.71

(d) **Anomaly** Worse reconstruction Recons Loss: 0.79

- **Improve current methods:**
  - Make One-class anomaly detectors robust to poorly defined normally class (multi-modal one-class?, ensembles?)
  - Try different approaches/structures
  - Purify training sets & better datasets
- **Increase anomaly types**: e.g. insects and leaves

(a) Insect, (Riseholme 2020)

(b) Pest, (Riseholme 2020)

(c) Pest, (Riseholme 2020)

(d) Pest, (Riseholme 2020)

Future plans

- **Implement** alongside the object detectors created at L-CAS e.g.

Future plans: Object detector + Anomaly detector

- **Explore Agri-robot-human interaction** Use it to raise flags . If anomaly score is high call agronomist (RHI, Active learning)
- **New and detailed anomaly dataset in process**

THANK YOU

Bibliography

[1] C. A. Deutsch, J. J. Tewksbury, M. Tigchelaar, D. S. Battisti, S. C. Merrill, R. B. Huey, and R. L. Naylor, “Increase in crop losses to insect pests in a warming climate,” *Science*, vol. 361, no. 6405, pp. 916–919, 2018.

[2] S. Savary, L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson, “The global burden of pathogens and pests on major food crops,” *Nature ecology & evolution*, vol. 3, no. 3, pp. 430–439, 2019.

[3] A. Makhzani and B. J. Frey, “Winner-take-all autoencoders,” in *Advances in neural information processing systems*, 2015, pp. 2791–2799.

[4] P. Perera and V. M. Patel, “Learning deep features for one-class classification,” *IEEE Transactions on Image Processing*, vol. 28, no. 11, pp. 5450–5463, 2019.

[5] L. Ruff, R. Vandermeulen, N. Goernitz, L. Deecke, S. A. Siddiqui, A. Binder, E. Müller, and M. Kloft, “Deep one-class classification,” in *International conference on machine learning*, 2018, pp. 4393–4402.