Zero-shot Anomaly Detection with Pre-trained Segmentation Models

Matthew Baugh[1], James Batten[1] , Johanna P. Müller[2], Bernhard Kainz[1, 2]
[1] Imperial College London, [2]Friedrich–Alexander University Erlangen–Nürnberg
matthew.baugh17@imperial.ac.uk

JUNE 18-22, 2023 CVPR VANCOUVER, CANADA

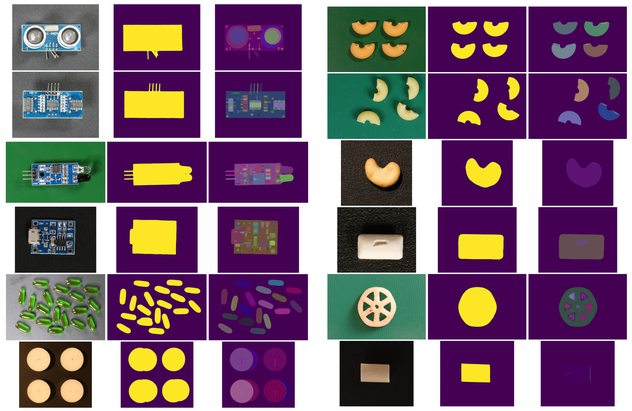
3rd Place: Direct zero-shot anomaly detection

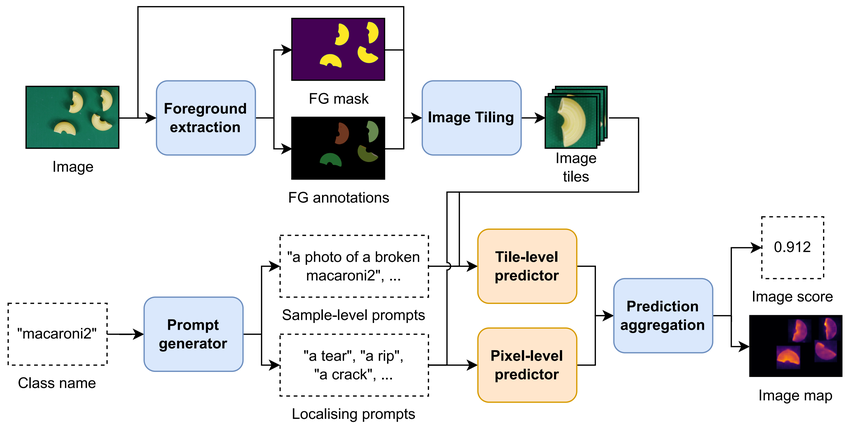
Zero-shot segmentation models in order to improve WINCLIP's [1] ability to localise anomalies.

Our method **does not change in any way** when being applied to different test sets, so can be directly applied to a new dataset **without additional information**.

Foreground extraction:

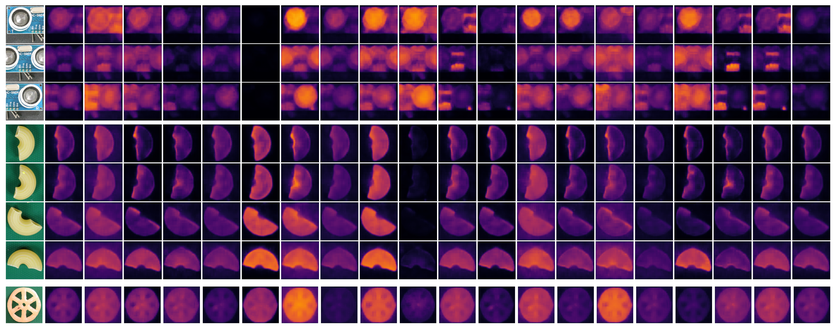
- **Dichotomous Image Segmentation** [2] produces foreground mask
  - Works well even for complex objects (pcbs)
  - Misses some instances when number of objects high (capsules)
- **SAM** [3] annotations give estimate of object complexity.
- **Tile foreground components** in order to focus on relevant image areas
  - Tile based on component size, aspect ratio and SAM complexity.





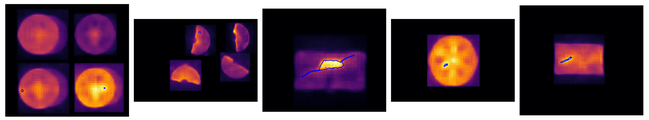
Predict on foreground component tiles

- Apply **CLIPSeg** [4] to each image tile with a variety of generic anomaly descriptions to produce multiple predictions
- **Harmonic average** across segmentations to get pixel-wise prediction per tile
- Use WINCLIP [1] prompt ensembling for tile-level anomaly score



Final prediction

- Average tile-level scores for each foreground component
- **Weight anomaly localisations** by foreground component anomaly score
- Combine tiles to produce anomaly map over image



	pcb1	pcb2	pcb3	pcb4	capsules	candle	macaroni1	macaroni2	cashew	chewinggum	fryum	pipe_fryum	Mean
WINCLIP	2.4	4.7	10.3	32.0	9.2	22.5	7.0	1.0	13.2	41.1	22.1	12.3	14.8
APRIL-GAN	12.5	23.4	21.7	31.3	48.5	39.4	35.5	13.7	22.9	78.5	29.7	30.4	32.3
Ours	29.5	11.0	4.7	21.7	31.9	20.2	24.6	7.2	24.5	63.4	31.3	19.6	24.2

Comparison of F1-max at pixel level with WINCLIP as a baseline, APRIL-GAN which won the challenge, and ours which came third.

Conclusion

- We apply **zero-shot segmentation** models to better localise unseen anomalies
- **Biggest improvement** for classes with small anomalies, due to more fine-grained predictions from applying segmentation models to tiles cropped to foreground components
- Struggles when domain gap between pre-trained models and test data is large, or anomalies are highly specialised, such as PCBs.

Zoom Room

https://bit.ly/3P7tU5x

Technical Report

https://bit.ly/43G9Pry

References

1. Jeong, J., Zou, Y., Kim, T., Zhang, D., Ravichandran, A., Dabeer, O.: Winclip: Zero-/few-shot anomaly classification and segmentation. In: CVPR (2023)
2. Qin, X., Dai, H., Hu, X., Fan, D.P., Shao, L., Gool, L.V.: Highly accurate dichotomous image segmentation. In: ECCV (2022)
3. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Doll´ar, P., Girshick, R.: Segment anything. ArXiv:2304.02643 (2023)
4. Lüddecke, T., Ecker, A.: Image segmentation using text and image prompts. In: CVPR (2022)