

# Advancing Anomaly Detection: The IDW dataset and MC algorithm

In this work we present a novel anomaly detection dataset, Industrial Defects in the Wild (IDW). IDW contains images of various industrial and household inspection processes. It features real images with complex and varied perspectives from freely moving cameras. We show this is more challenging than the well-known MVtec dataset. We also present MultiCore (MC), a novel algorithm that achieves state-of-the-art results on the introduced IDW dataset and popular MVtec dataset. The MC algorithm trains multiple nearest neighbour predictors, each with different hyperparameters. We propose that an ensemble is more powerful than any individual model. Synthetic anomalies are created using a novel schema intended to systematically cover as many variations as possible. The ensemble output is fed into a heatmap fusion module, which is trained in a supervised fashion using the synthetic anomalies and a perimeter-based loss function. On the popular MVtec dataset, the MC algorithm achieves P-AUC score of 0.986. On the introduced and more challenging IDW dataset, the MC algorithm achieves a P-AUC of 0.935. We verify that these results are state-of-the-art by trialing the existing top fourteen anomaly detection algorithms which have code available. The IDW dataset can be found at: <https://github.com/alex1995/IDW>, and MultiCore code can be found at: <https://github.com/alex1995/MultiCore>.

Alexander D.J. Taylor (1)  
adjt20@bath.ac.uk

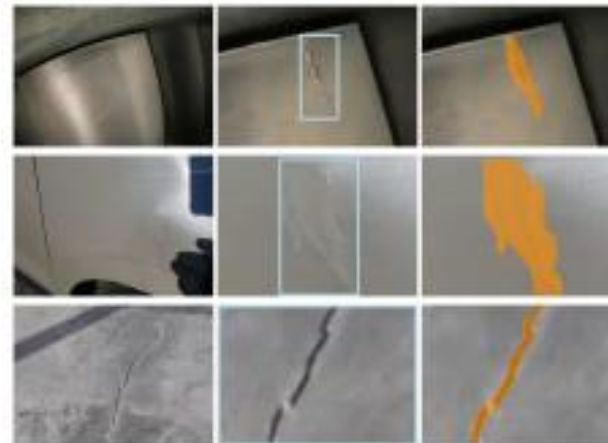
(1) ART-AI Department of  
Computer Science  
University of Bath

Jon Morrison (2)  
jonathan.morrison2@rolls-royce.com

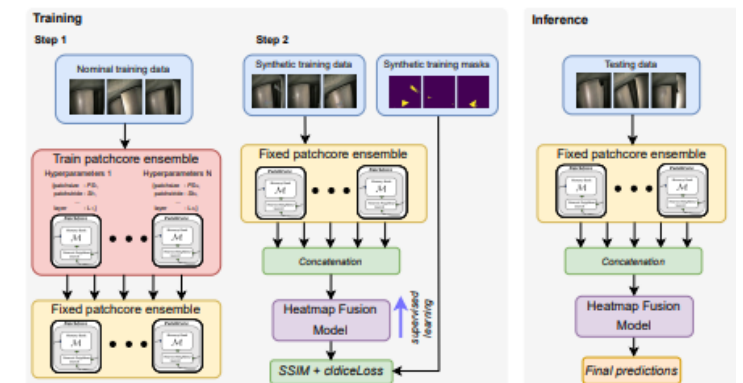
(2) Defence Operations  
Rolls-Royce

Phillip Tregidgo (2)  
phil.tregidgo@rolls-royce.com

Neill D.F. Campbell (1)  
nc537@bath.ac.uk



IDW dataset; from left to right: image, zoomed-in image on anomaly, zoomed-in image with pixel-wise label



Schema of the MultiCore algorithm.