Weakly Supervised Learning for Industrial Optical Inspection

https://hci.iwr.uni-heidelberg.de/node/3616

Abstract

In the following, we present a synthetic benchmark corpus for detect detection on statistically textured surfaces.We hope that it facilitates to further develop and benchmark classification algorithms for applications of industrial optical inspection. All data is publicly available and can be downloaded from this page.

Competition at DAGM 2007 symposium

The DAGM (Deutsche Arbeitsgemeinschaft für Mustererkennung e.V., German chapter of the IAPR (International Association for Pattern Recognition))and the GNSS (German Chapter of the European Neural Network Society) offered an open competition on *Weakly Supervised Learning for Industrial Optical Inspection* held as part of the DAGM symposium in 2007.

The competition was inspired by the fact that automated optical inspection allows to reduce the cost of industrial quality control significantly. The competitors had to design a classification algorithm which:

- detects miscellaneous defects on various statistically textured backgrounds.
- learns to discern defects automatically from a weakly labelled training data.
- works on data whose exact characteristics are unknown at development time.
- adapts all parameters automatically and does not require any human intervention.
- has a moderate running time (in this competition 24 hours for training and 12 hours for the test phase).
- takes into account asymmetric costs for false positive and false negative decisions (1:20 was used for the competition).

Data description

The data is artificially generated, but similar to real world problems. The first six out of ten datasets, denoted as development datasets, are supposed to be used for algorithm development. The remaining four datasets, which are referred to as competition datasets, can be used to evaluate the performance. Researchers should consider not using or analyzing the competition datasets before the development is completed as a code of honour.
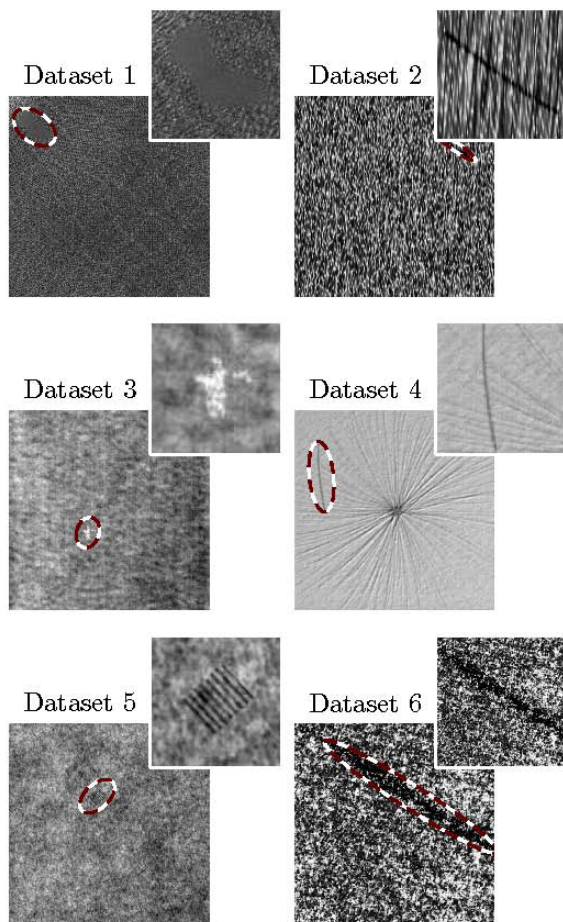In the following we provide some details about the datasets:

- Each development (competition) dataset consists of 1000 (2000) 'non-defective' and of 150 (300) 'defective' images saved in grayscale 8-bit PNG format.
- Each dataset is generated by a different texture model and defect model.
- 'Non-defective' images show the background texture without defects, 'defective' images have exactly one labelled defect on the background texture.
- All datasets has been randomly split into a training and testing sub-dataset of equal size.
- Weak labels are provided as ellipses roughly indicating the defective area. Technically, defective images are augmented with a separate grayscale 8-bit image in the PNG format located in a folder 'Label'. The values 0 and 255 denote background and defective area, respectively.

All meta-data is subsumed in a separate ASCII textfile called 'Labels.txt' which is located in the 'Label' folder. The structure is as follows:
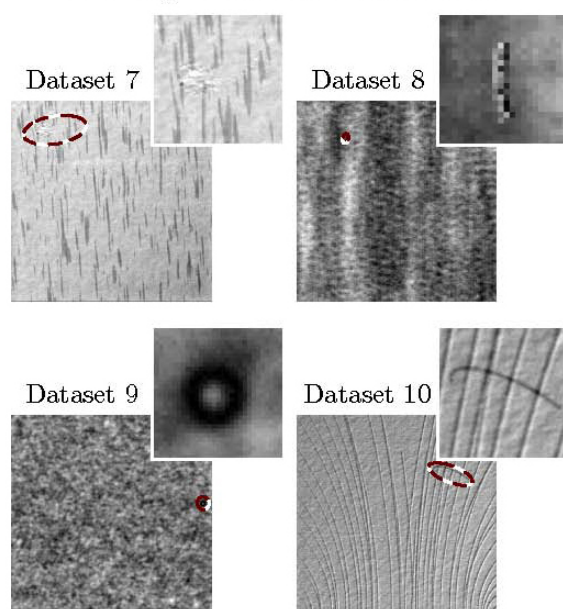1 \n
[id of item no. 1] \t [0 if non-defective, 1 if defective] \t [filename of raw image no. 1] \t 0 \t [filename of label image no. 1 if defective, 0 otherwise] \n
...
[id of item no. N] \t [0 if non-defective, 1 if defective] \t [filename of raw image no. N] \t 0 \t [filename of label image no. N if defective, 0 otherwise] \n

The following figure illustrates one defective image from each of the ten datasets.

Development Datasets

Dataset 1
Dataset 2
Dataset 3
Dataset 4
Dataset 5
Dataset 6

Competition Datasets

Dataset 7
Dataset 8
Dataset 9
Dataset 10

Acknowledgements