VIADUCT: Multisector dataset for Visual Industrial Anomaly Detection

Jan Lehr*
Fraunhofer IPK
Berlin, Germany

Jan Philipps
Fraunhofer IPK
Berlin, Germany

Alik Sargsyan
Ngene LLC.
Yerevan, Armenia

Maximilian Botschen
Fraunhofer IPK
Berlin, Germany

Shoghik Gevorgyan
Ngene LLC.
Yerevan, Armenia

Anna-Maria Paust
Fraunhofer IPK
Berlin, Germany

Martin Pape
Fraunhofer IPK
Berlin, Germany

Viet Nguyen Hoang
Fraunhofer IPK
Berlin, Germany

Abstract

In recent years there have been many new methods of semi-supervised learning in visual industrial anomaly detection. But there only a few new datasets available. The most popular dataset currently is the MVTec-AD dataset and many methods have already achieved AUROC scores of over 99 % on this dataset. As a result, it seems that the problem provided by this dataset has been solved. In order to address this, we have created a new dataset for visual industrial anomaly detection. The VIADUCT dataset aims to provide a new challenge and benchmark for Industrial Anomaly Detection methods. It contains 49 diverse categories with several defects of varying difficulty each. With the support of many manufacturing companies, we have included many real inspection problems in this dataset. It contains a large number of different defects, all of which are pixel-wise annotated.

Keywords: anomaly detection, industrial anomaly detection, automated optical inspection, quality control, visual inspection

Figure 1: This Figure shows one example image from each of the 49 categories.

1 Introduction

This data sheet is a supplement to the paper *AD3: Introducing a score for Anomaly Detection Dataset Difficulty assessment using VIADUCT dataset* [2] and introduces the novel VIADUCT dataset, released under the Creative Commons Attribution 4.0 (CC BY) license. Please always cite the original work from the conference proceedings.

The current scenario in industrial anomaly detection datasets presents several challenges that hinder the development of robust and reliable anomaly detection models. Many widely-used datasets for training and benchmarking Industrial Anomaly Detection (IAD) methods, like [1, 3, 4], face issues such as a lack of categories or diverse images, or presenting problems that have essentially been solved, resulting in most methods achieving nearly perfect scores. This scarcity of suitable datasets makes it difficult to compare the performance of different methods. Furthermore, the identified problems extend to defects within the datasets, which often lack the diversity, realism, and complexity needed to challenge state-of-the-art anomaly detection algorithms. This situation not only impedes the development of models that can address real industrial challenges but also increases the risk of overfitting, as current methods be-

*Correspondence: jan.lehr@ipk.fraunhofer.de
DOI: http://dx.doi.org/10.24406/fordatis/305

come tailored to the specific characteristics of these limited datasets.

To address these challenges, we submit a new and more complex anomaly detection dataset. This dataset poses a new, challenging problem, serving as a test for the effectiveness of anomaly detection methods under realistic and diverse industrial conditions. Introducing a more realistic challenge aims to fulfill the need for anomaly detection models capable of generalizing and predicting performance in real-world scenarios. The VIADUCT dataset also supports the development of anomaly detection models with greater robustness, comprehensiveness, and generalization capabilities.

2 Dataset

The multisector data set for Visual Industrial Anomaly Detection (VIADUCT) is created by Fraunhofer IPK in conjunction with Ngene LLC and the generous support of numerous production companies that supply samples of parts and defects from their production lines. It is made up of 10,888 images in total consisting of 7,181 defect-free images and 3,707 images depicting various defects. The size of the uncompressed dataset is close to 20.5 gigabytes.

2.1 Dataset structure and information

The dataset consists exclusively of PNG files. Each category is divided into the sub-folders train, val, test and ground truth. Only defect-free images are in the train and val directory. Figure 2 shows how the dataset is structured.

In the test directory are the sub-folders for each defect category and the category "good" (defect-free). The annotations (also called ground truth or labels) for the images in the defect category directories are located in the ground truth directory mentioned above. The so-called anomaly maps are binary and are black with white pixels where the defects are in the image.

In the dataset there are 49 categories belonging to 11 diverse sectors, ranging from technological to food products. There are 28 defect-classes including missing parts and materials, deformations, contamination among many others in a large variety of appearance and prominence. These defect-classes contain 7,622 individual regions with the smallest region being $6.1 \cdot 10^{-5}$ times the size of the image area ($0.0061\,\%$). A detailed breakdown of the dataset is shown in table 1.

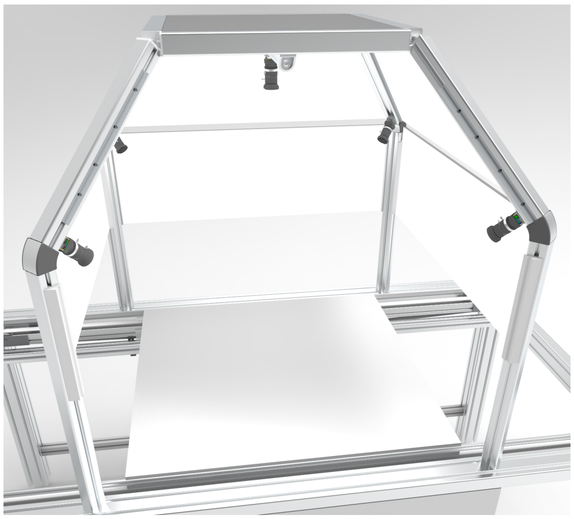
The images have resolutions between 1,080×1,080 and 2,160×2,160, with all images being square. Each image features a single object in the center with the featureless background ranging from mostly white into the light greys. The exception to this are the agricultural classes that consist of a mass of shredded plant material and cover the entire image with the background mostly obscured. Some images also feature small dust and debris particles in the background that are not labeled as defects.

```
VIADUCT/
├── 00<ObjName>/
│   ├── ground_truth/
│   │   ├── <defect_A>/
│   │   │   ├── Image_defectA_<id>_0.png
│   │   │   ├── ...
│   │   │   ├── Image_defectA_<id>_4.png
│   │   │   ├── Image_defectA_<id>_0.png
│   │   │   └── ...
│   │   └── <defect_B>/
│   │       ├── Image_defectB_<id>_0.png
│   │       └── ...
│   ├── test/
│   │   ├── good/
│   │   │   ├── Image_defect-free_<id>_0.png
│   │   │   ├── ...
│   │   │   ├── Image_defect-free_<id>_4.png
│   │   │   ├── Image_defect-free_<id>_0.png
│   │   │   └── ...
│   │   ├── <defect_A>/
│   │   │   ├── Image_defectA_<id>_0.png
│   │   │   ├── ...
│   │   │   ├── Image_defectA_<id>_4.png
│   │   │   ├── Image_defectA_<id>_0.png
│   │   │   └── ...
│   │   └── <defect_B>/
│   │       ├── Image_defectB_<id>_0.png
│   │       └── ...
│   ├── train/
│   │   └── good/
│   │       ├── Image_defect-free_<id>_0.png
│   │       └── ...
│   └── val/
│       └── good/
│           ├── Image_defect-free_<id>_0.png
│           └── ...
├── 01<ObjName>/
│   └── ...
├── ...
└── 48<ObjName>/
```

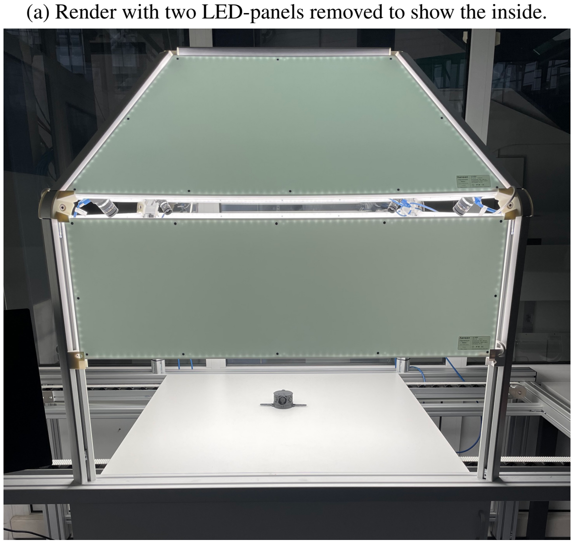
Figure 2: Structure of the dataset. The IDs, consisting of ten numerals, are unique for each image. The names of the masks in "ground_truth" are identical to the original images in "test". The last numeral (0–4) in the file name indicates which camera position the image is taken from.

2.2 Acquisition setup

To capture the images, a dome-type setup with five cameras and nine LED panels with a total of 273 Watt is used. Figure 3 shows the CAD-model as well as the final setup. Four of the five cameras point towards the object in the center from an 45° angle, while one is positioned directly above the object and points downwards. The camera and lighting setup can be seen in Figure 3.



(a) Render with two LED-panels removed to show the inside.



(b) Real-world picture of the final setup with an object on the tray.

Figure 3: The image acquisition setup. Each of the four sides is equipped with two LED panels, an additional LED-panel makes up the top. The four side-cameras can be seen in the gap between the LED-panels in 3b.

All five cameras capture an image simultaneously, thus creating a set of five images of the same sample. Images of the same set can be identified by neighbouring IDs as well as camera position as indicated in the filename. A set starts with camera 0 (top) and ends with camera 4. Images of the same set should ideally not be separated into different splits, eg. train and test because they show the same sample.

Basler *daA3840-45uc* cameras with a resolution of $3,840\,\mathrm{px} \times 2,160\,\mathrm{px}$ are used to acquire the images. The sensor is a Sony IMX334 CMOS-Sensor with rolling shutter. Depending on object size, the cameras were equipped with lenses of 35 mm, 25 mm, 16 mm or 8 mm focal length.

Category	Sector	# Train	# Test (good)	# Test (defective)
shred. corn fine	agriculture	80	40	143
shred. corn medium	agriculture	130	65	105
shred. corn rough	agriculture	60	30	140
3-pole socket housing	electronics	80	40	144
battery holder	electronics	50	25	40
coaxial t-adapter	electronics	75	40	24
d-sub plug	electronics	100	50	48
encoder	electronics	40	20	29
ethernet plug	electronics	40	20	40
ferrule	electronics	145	75	50
joystick	electronics	50	25	51
PCB	electronics	40	20	55
ring cable lug	electronics	100	50	66
socket housing	electronics	50	25	62
terminal block A	electronics	75	40	85
terminal block B	electronics	50	25	60
vibration sensor	electronics	40	20	30
cherry	food	190	95	122
raspberry	food	120	60	94
strawberry	food	125	60	145
tea bag	food	100	50	100
cylinder screw	generic	125	65	32
device box	generic	40	20	46
pipe clamp	generic	40	20	29
reinforcement casing	generic	40	20	21
small screw	generic	95	45	110
fitting	hydraulics	40	20	29
l-fitting	hydraulics	40	20	26
faucet grip blue	mechanical	40	20	24
faucet grip red	mechanical	35	20	26
key	mechanical	65	35	40
threaded fitting	mechanical	65	35	72
bolt	medical	100	50	150
chisel pliers	medical	130	65	73
drill	medical	125	65	148
large pink tablet	medical	40	20	70
oblong orange pill	medical	75	40	146
redon needle	medical	125	65	73
retractor	medical	125	65	75
round pill	medical	40	20	92
small white tablet	medical	40	20	110
surgical mask	medical	75	40	133
aluminium plate	metal production	20	10	160
paperclip	office supplies	80	40	73
tack	office supplies	115	55	147
air muffler large	pneumatics	50	25	50
air muffler small	pneumatics	50	25	58
saw blade	tools	40	20	30
voltage tester	tools	40	20	30

Table 1: The table gives a statistical overview of all categories. Beside the sector, number of training images and test images are shown.

References

[1] Paul Bergmann, Michael Fauser, David Sattlegger, and Carsten Steger. Mvtec ad — a comprehensive real-world dataset for unsupervised anomaly detection. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9584–9592, 2019. 1

[2] Jan Lehr, Jan Philipps, Alik Sargsyan, Martin Pape, and Jörg Krüger. Ad3: Introducing a score for anomaly detection dataset difficulty assessment using viaduct dataset. In *European Conference on Computer Vision (ECCV)*, 2024. 1

[3] Pankaj Mishra, Riccardo Verk, Daniele Fornasier, Claudio Piciarelli, and Gian Luca Foresti. Vt-adl: A vision transformer network for image anomaly detection and localization. *2021 IEEE 30th International Symposium on Industrial Electronics (ISIE)*, pages 01–06, 2021. 1

[4] Yang Zou, Jongheon Jeong, Latha Pemula, Dongqing Zhang, and Onkar Dabeer. Spot-the-difference self-supervised pre-training for anomaly detection and segmentation. In *European Conference on Computer Vision*, 2022. 1