Deep Learning in Anomaly Detection for Manufacturing Processes

Master's thesis in Mathematics • Anna-Maria Gleißner
Supervised by: Dr. Filippo Riccio (evopro systems engineering AG) • Prof. Dr. Jürgen Frikel (OTH Regensburg)

Overview

The fundamental question of this work is, how one can develop efficient and robust methods for detecting anomalies on objects in automated visual inspections.
Main task of an inspection is the classification of objects into class "**O.K.**", which means that the requirements for the object were met, or class "**n.O.K.**", meaning that an anomaly was detected. Depending on the underlying task, the anomalies should additionally be localized.
In order to solve the detection problem, one applies Deep Learning tools since traditional image processing approaches quickly reach their limits.
As the demands in today's world are becoming more and more complex and manifold, it is important that the developed methods immediately adapt to new requirements.

Deep Learning systems including Neural Networks as their main component therefor provide efficient solutions.

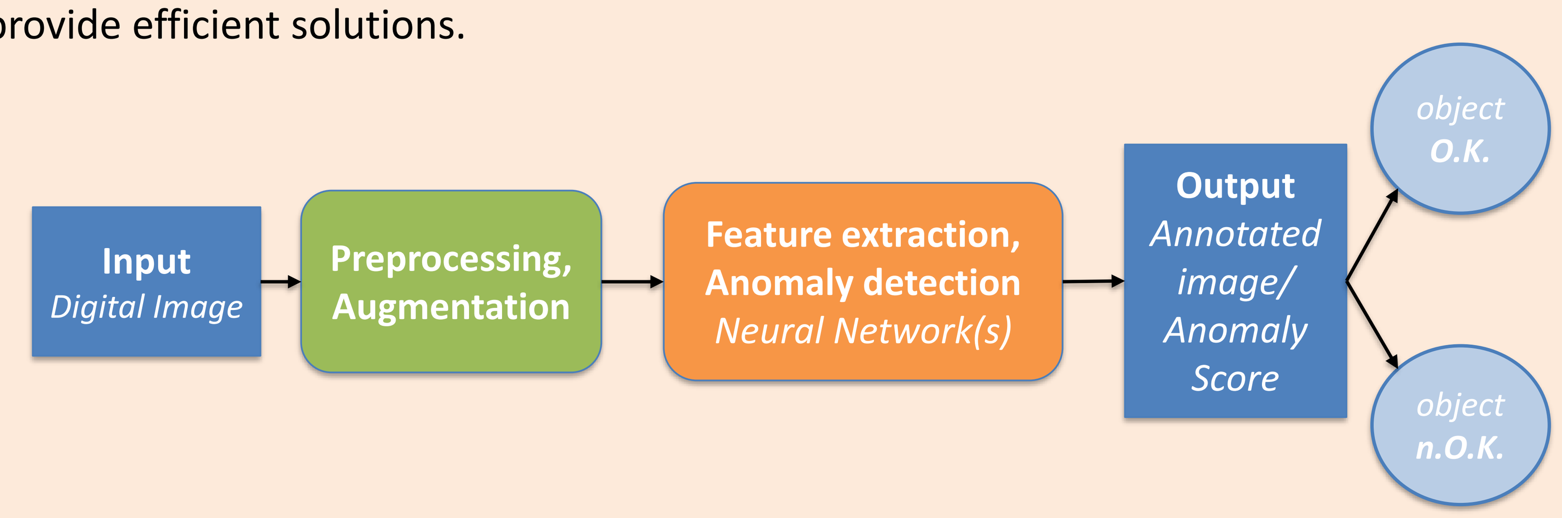

Fig. 1: Workflow anomaly detection with Deep Learning.

Basis for image classification tasks are digital images which show the object to be classified. After passing the Deep Learning element, one receives an annotated image containing the desired information.
From a mathematical point of view, we are aiming to construct a decision function

$$\boxed{d: N \times N \times C \to \{0,1\}}$$

where $N \times N$ describes the size of the square input image, C is the number of channels and $\{0,1\}$ is the set of possible outputs, where class n.O.K. is indicated as 0 and O.K. as 1 [2].

The following work compares two common approaches to proceed anomaly detection, YOLOv3 and fast-AnoGAN.

YOLOv3 – Supervised Learning

YOLOv3 is an example of Supervised Learning working on balanced datasets which include an equal number of both abnormal and normal image samples for training and validation. It is constructed out of a single Neural Network [3].

Object detection

In general, YOLOv3 is an object detection algorithm outputting bounding boxes for the identified objects. It uses the idea of **feature maps** at 3 different sizes to detect either small, medium or large size objects. Additionally, one applies 3 **anchor boxes** of fixed size on each feature map cell which saves us from solving a regression problem for the box coordinates [4].

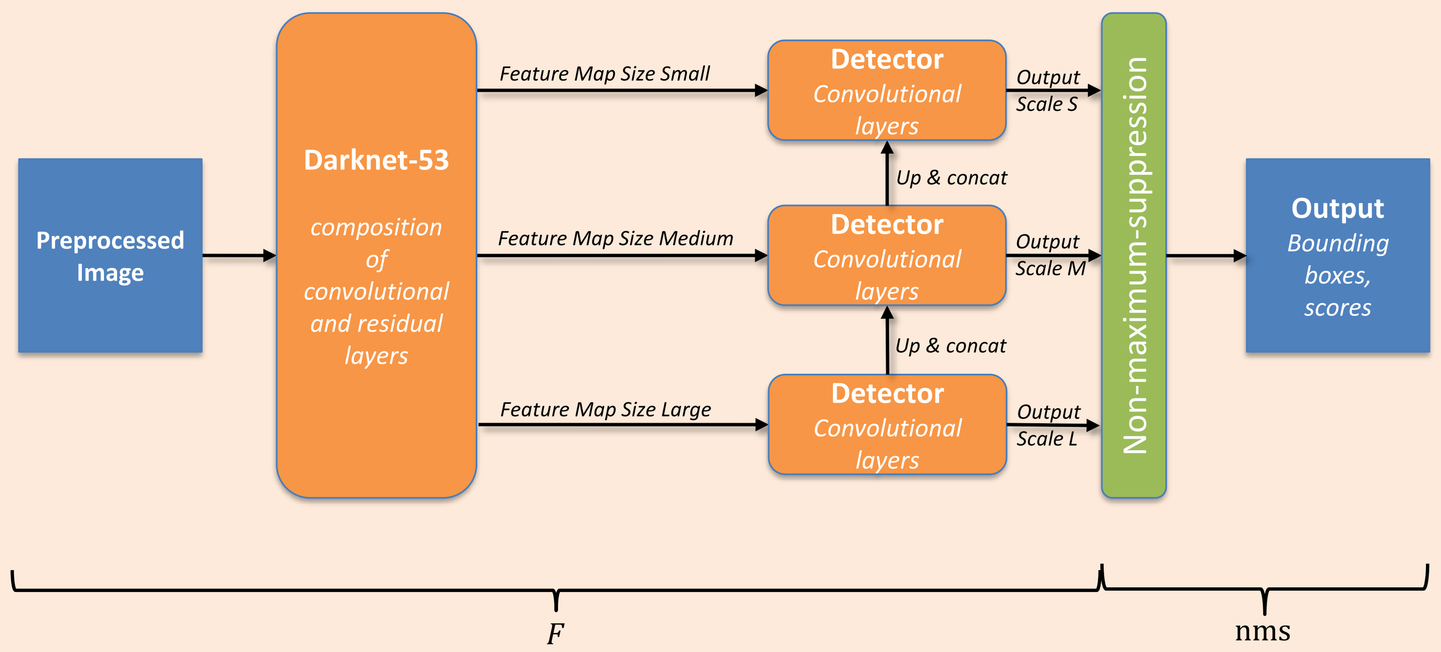

Fig. 2: Workflow object detection with YOLOv3.
Up & concat means upsampling and concatenation.

Output format

- 4 **bounding box coordinates** (x- and y-coordinates, height and width)
- 1 **confidence score**, i. e. P("bounding box contains object")
- 2 **class scores**, i. e. P("object i.O.") and P("object n.i.O.").

Decision function

$$\boxed{d = \pi \circ \mathrm{nms} \circ F \circ p}$$

- p is a **preprocessing function** conducting a set of data augmentation operations,
- nms is the **non-maximum-suppression function** suppressing output boxes with low confidence scores and boxes detecting the same object,
- π is a **projection** from the network output to 0 or 1, depending on the task.

Fast-AnoGAN – Semi-supervised Learning

Fast-AnoGAN is a representative of Semi-Supervised Learning combining an Unsupervised Generative Adversarial Network (GAN) with an Encoder network. It operates with unbalanced datasets which contain images of normal samples only for training and few normal and abnormal images for validation [3].

Anomaly Score

In fast-AnoGAN, anomaly detection is conducted by comparing a testing image with its "normal" representation. Out of this, an anomaly score is calculated, which is used to classify the image.
To set up the model, one first trains a GAN consisting of a **Generator** and a **Discriminator** Network. This specializes the Generator on producing images of normal objects out of a latent space $\mathbb{R}^n$, $n \in \mathbb{N}$. For the classification process, a trained **Encoder** Network is added to complete the calculation of the Anomaly Score [5].

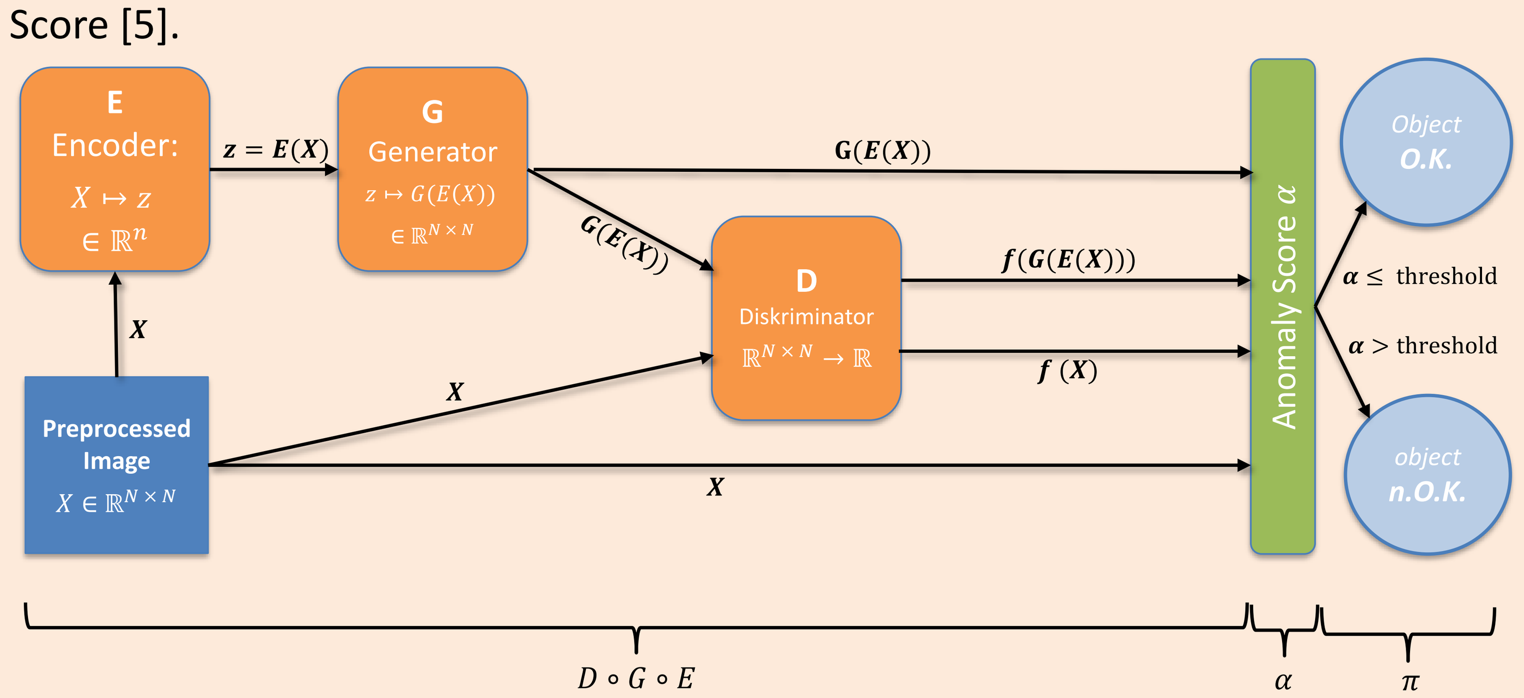

Fig. 3: Workflow anomaly detection with fast-AnoGAN.

Decision function

$$\boxed{d = \pi \circ \alpha \circ D \circ G \circ E \circ p}$$

- p and π as before in YOLOv3,
- $\alpha: [0,1] \to \mathbb{R}$ is the **Anomaly Score function**, given by

$$\boxed{\alpha(X, G(E(X))) = \frac{1}{N^2} \cdot \|X - G(E(X))\| + \frac{\kappa}{n_d} \cdot \left\| f(X) - f\left(G(E(X))\right) \right\|^2}$$

where f is a mapping to the output of a middle layer in the Discriminator, n_d its output dimension and κ a weighting factor.

Detection results

The defect detection was performed with the MVTec-AD dataset using the category "bottle", which includes images of defect-free and broken bottle bottoms.
Below, we show some output samples of both YOLOv3 and fast-AnoGAN to demonstrate the defect evaluation functionality.

Fig. 4: Detection results of **YOLOv3**. Images without bounding boxes (anomalies) are indicated as O.K., else as n.O.K. Defects are located directly by the output of the Neural Network.
For the code, see *https://github.com/ultralytics/yolov3*.

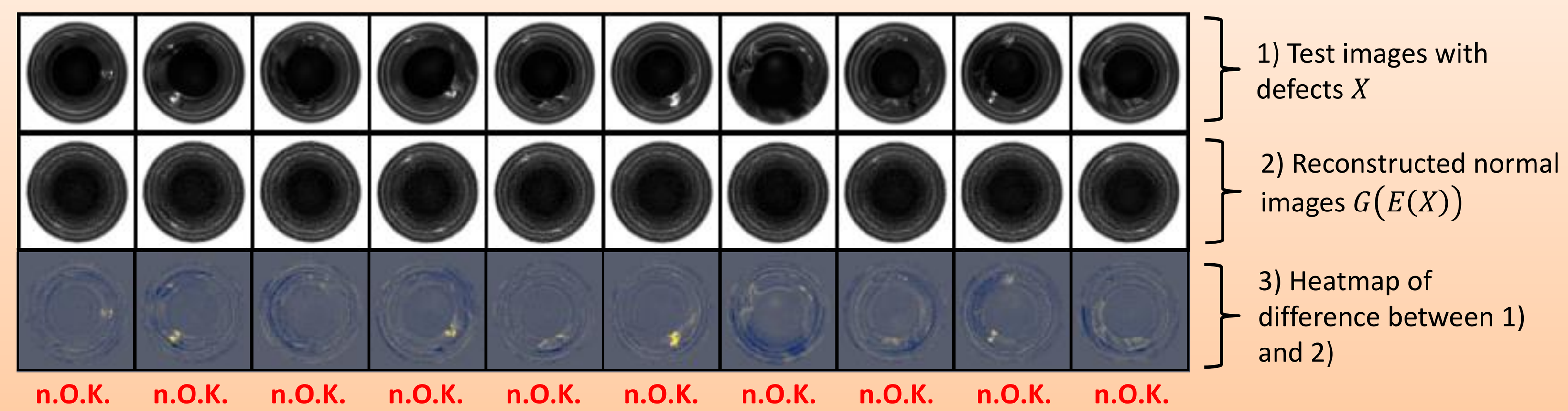

Fig. 5: Detection results of **fast-AnoGAN**. Images are converted into greyscale. The reconstructed normal images are produced by the Generator. Through comparison with the test images an Anomaly Score is calculated and the defects can be localized (see yellow spots in 3)).
For the code, see *https://github.com/A03ki/f-AnoGAN*.

Conclusions

YOLOv3

- Training of a single Neural Network → fewer hyperparameters, little computational time until convergence
- Very much developed in recent years → gain in stability and robustness

Fast-AnoGAN

- Unbalanced dataset → key advantage in industrial applications (lack of anomalous data)
- Encoder training is based on GAN training → failing GAN training implies failing Encoder training

Both methods are each suitable for different types of applications. YOLOv3 initially was developed for all kinds of object detection tasks, where anomaly detection only describes a specific application area. Anomalies must be predefined and training images need to be labelled. On the contrary, fast-AnoGAN easily reacts to new defect occurencies and therefore provides a more flexible solution.
As a next step, one would collect an image dataset of a real production scenario and evaluate the performance of both methods statistically.

References

[1] Rhagavendra, C., Chawla, S. *Deep Learning for Anomaly Detection: A Survey*. ArXiv preprint, 2019.
[2] Demant, C., Streicher-Abel, B., Springhoff, A. *Industrielle Bildverarbeitung. Wie optische Qualitätskontrolle wirklich funtkioniert*. Springer-Verlag, 2011.
[3] Jiang, Y., Wang, W., Zhao C. *A Machine Vision-based Realtime Anomaly Detection Method for Industrial Products Using Deep Learning*. Chinese Automation Congress (CAC), 2019.
[4] Redmon, J., Farhadi, A. *YOLOv3: An Incremental Improvement*. ArXiv, 2018.
[5] Schlegel, T. et al. *Medical Image Analysis. F-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks*. ELSEVIER, 2019.