euspen's 22nd International Conference & Exhibition, Geneva, CH, May/June 2022
www.euspen.eu

Augmenting image datasets for quality control models using CycleGANs

Alexander Moriz[1], Dominik Wolfschläger[1], Benjamin Montavon[1], Robert H. Schmitt[1,2]

[1]*WZL | RWTH Aachen University;* [2]*Fraunhofer Institute for Production Technology IPT*

A.Moriz@wzl.rwth-aachen.de

Abstract
Deep learning (DL) has proven to be a powerful tool for solving common machine vision tasks, such as image classification, defect segmentation and defect recognition. Usually, training DL models requires significant amounts of annotated data samples, which are generally sparse or of inadequate quality in many quality assurance applications in the engineering domain. Especially the thorough annotation of data yields a major obstacle for the generation of industrial datasets, since it is a complex, time-consuming task requiring expert knowledge of the process under examination. Further, the rareness of defects in rather stable production processes can lead to highly unbalanced datasets, hampering the training process. Combined with the seldom distribution of industrial data due to privacy concerns, the lack of data often hinders the adoption of DL approaches for quality assurance. Recently, network structures following the design of Generative Adversarial Networks (GANs) show astonishing results in the field of image synthesis and neural style transfer. Given a set of unpaired images from two domains, cycle-consistent GANs (CycleGANs) learn how to translate a given image from one domain to the other and vice-versa. This capability can be exploited to augment datasets in a controllable manner in order to alleviate the problems arising in the application of DL for realizing vision-based quality control. This work investigates the employment of CycleGANs to extend the image datasets for two use cases, the detection of pores in computed tomography data and the detection of surface defects on sheared edges of fine blanked parts. Given randomly generated binary masks, the trained CycleGANs are capable of generating an arbitrary amount of synthetic yet realistic images in the desired domains, alleviating the problems of both the data amount and the necessary annotations and demonstrating the great potential of image synthesis using GANs.

Keywords: CycleGAN, deep learning, image synthesis, computed tomography, defect detection, quality assurance

1. Introduction

Data-driven automation and optimization of manufacturing systems within the fourth industrial revolution requires the ability to monitor, record and link any relevant process data and quality information to analyze them with respect to a subsequent decision process. Already today vision-based quality control (QC) systems are capable of monitoring many manufacturing processes and products yielding reliable results for an in-line quality control [1]. However, these systems face increasingly demanding conditions such as small batch sizes and complex geometries for additively manufactured (AM) components [2] or changing environmental conditions during mass production processes like industrial stamping [3]. Deep Learning (DL) methods promise to be a viable solution to cope with these conditions. One particular strength is the ability to extract features in a data-driven manner without the need of human intervention [4] and adapt themselves to the hidden structure of the data. For this reason, DL methods are considered a key technology in order to realize the vision of Industry 4.0 and have proven to exceed the capabilities of classical methods in common machine vision tasks such as defect classification, defect recognition or segmentation. However, DL methods usually require large, annotated datasets to reach a critical generalization effect to ensure that the complexity and variation of the targeted task have been captured. Only if the resulting (trained) model shows this capability, it can be used for a successful deployment in an industrial quality control application.
Often, this prerequisite is not fulfilled in the industrial context, due to the high costs, necessary time or a lack of expertise for the proper annotation of large amounts of data. In order to overcome the resulting issues of insufficient data, multiple approaches have been proposed in the literature.
This work investigates the employability of CycleGANs to augment industrial datasets to build reliable QC systems and implements CycleGAN models for two industrial QC applications. In general, generative adversarial networks (GANs) allow a semi-supervised learning of the underlying distribution of a given training dataset and can be used subsequently to synthesize realistic samples from the learned distribution. In particular, the CycleGAN architecture allows to map inputs from one domain to another, without requiring paired samples from image- and annotation domains for training. Our results indicate that CycleGANs can learn to generate realistic samples even without tuning and large datasets for both use cases. Given the different nature of the considered domains, the results show further that our approach is not limited to either of the domains and could be applied to other datasets in a straightforward way to generate synthetic data as well.
This article continues in Section 2 with a revision of approaches that are used to augment datasets for QC use cases and introduces the concept of CycleGANs. Section 3 presents the use cases of a vision-based defect recognition of fine blanked sheet metal workpieces and of a Computed Tomography (CT) inspection process for detecting pores in AM parts. Section 4 presents the implementation and results of the augmented datasets using CycleGANs. Sections 5 and 6 conclude the article with a summary and future research directions.

2. Related work

Comparably large annotated datasets, as they are used for benchmarking DL methods in ML sciences, are often not available with the necessary quality in the industrial context, either because annotating large amounts of data is an expensive task or because small batch sizes hinder the acquisition of a sufficient number of data points. In addition, industrial datasets are often unbalanced, e.g. since imperfect parts are usually produced less often than good parts, leading to a dangerous bias of the trained models.

2.1 Data augmentation for Deep Learning

One common way to reduce the impacts of an insufficient data basis (such as overfitting) is to add data augmentation to the model training process [5]. While not altering the data basis itself, classic augmentation techniques such as random flipping, mirroring, rotating or cropping/resizing of the input image add an additional amount of variation during the training process [5]. While preventing models from overfitting to a certain degree, these augmentation techniques are not capable of covering cases that are not present in the dataset at all. Especially in cases of strong class imbalances, as they are often present for the detection of rarely occurring defects, such augmentation techniques provide only limited remedy.

Another possibility is to use a simulation of a process under investigation, acquiring simulated data. Especially in the field of computed tomography, simulations are a common way to investigate sources of influence on the image acquisition process like the parts orientation [6]. In general, these simulations are subject to a trade-off: Extremely precise simulations, such as Monte-Carlo simulations, can generate authentic images but suffer from long computation times, making their application in practice unattractive [7]. Fast simulations like ray casting, on the other hand, can generate images almost in real time but sacrifice a significant amount of precision. Although simulated data can be used to leverage data size issues, as shown in [8] for CT images of aluminium casts, require even simplified physics-based simulations the configuration of many parameters that do not necessary correspond to ones of real CT systems, resulting in a tedious and error-prone setup process [9].

2.2 Cycle-consistent GANs

An alternative to a complete, physics-based image simulation are Generative Adversarial Networks (GANs). GANs are composed of two neural networks, a generator network ***G*** and a discriminator network ***D*** that are trained adversarially in a zero-sum-game. During training, ***D*** is exposed to samples from the training dataset as well as samples synthesized by ***G*** and learns to distinguish so-called 'real' from 'fake' samples. Upon that, ***G*** learns a mapping from a tensor from a different domain to synthesize samples that resemble samples from the training distribution. ***G*** can be used subsequently as a neural sampler for data augmentation. Neufeld et al. showed for the use case of automotive pistons that GANs can generate realistic CT slice images [10]. CycleGANs [11] are an extension of the GAN framework that comes with multiple advantages for data augmentation. The vanilla GAN architecture is extended by an inversely directed GAN resulting in the structure depicted in Figure 1. CycleGAN's training procedure relies on cycle consistency, transforming a given image from its domain A to the desired domain B and back. After training, the individual parts of the CycleGAN can be used independently, e.g. by using one of the generators to transform images from domain A to domain B and vice versa.

In this study, this domain transformation capability shall be exploited to investigate the plausibility of generated synthetic images for two industrial use cases. As demonstrated in [12] for the case of biomedical particles, using images and binary masks as datasets enables the CycleGAN model to relate both domains to each other in a semi-supervised manner. The respective generator parts can then be used to either segment elements in the image (particles, pores, defects, etc.) or to synthesize images given a binary mask. Hence, by forwarding a random-generated (yet realistic) binary mask through the respective generator part of the CycleGAN, it is possible to acquire an arbitrary amount of new images to enrich the initial (real) image dataset.

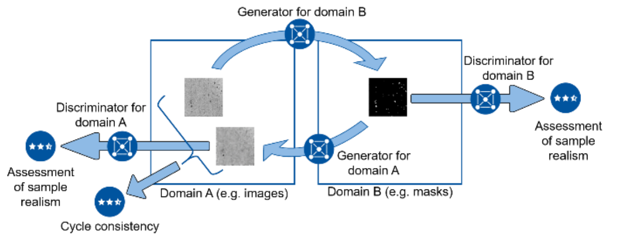

Figure 1: Outline of the CycleGAN architecture, consisting of two discriminator- and two generator pathways. The generators transform a given input (image) from one domain to the other, while the discriminators assess the plausibility of the given input. For the sake of clarity, the cycle consistency is shown for one pathway only.

3. Experimental setup and data preparation

This section introduces the investigated use cases and the respective datasets. Further, the implementation and training procedure of the CycleGAN models is outlined.

3.1. Porosity determination of AM parts using CT

AM parts are considered in an increasing number of applications due to their high degree of flexibility. However, QC for AM parts constitutes a challenge due to the small batch sizes and complex manufacturing process. An important characteristic influencing (among others) the quality of the manufactured parts is the degree of porosity. To determine the effective part quality, an accurate, but non-destructive assessment of the pore density is required. One common non-destructive method is the CT-based pore analysis. Wong et al. applied a 3D U-Net for pore segmentation in CT images of AM parts with promising results. However, they state that for a deployment of the approach more data is required to ensure a proper model generalization [13]. To investigate, how well pores of AM parts can be synthesized using a CycleGAN model to enrich a respective dataset, a CT scan of a Laser Powder Bed Fusion (LPBF) generated specimen was conducted for this study. Figure 2 shows the reconstructed CT volume on the left and an exemplary slice of the volume on the right. The pores formed during the AM process of the specimen are clearly visible as dark, mostly circular spots in the slice. The resulting stack of high-resolution slice images (3322 x 3325 px) was transformed into a more homogenous patch-dataset to focus the investigation on pores. Each slice of the specimen was cropped into quadratic, non-overlapping patches of size 300 x 300 px, explicitly excluding patches including transition areas between material and background, resulting in a patch-dataset of over 8000 image patches in total, where each patch was stored individually as a distinct image. To generate ground truth masks of the present pores for each patch, an automated, locally adaptive gray value based approach was chosen to segment pores in the image patches. In contrast to a global threshold, this approach accounts the severe effect of beam hardening better, which can be observed in Figure 2 on the right. An example of such a generated ground truth mask is shown in Figure 3. Finally, the resulting image (patch) dataset is divided in a random, but fixed manner, in a train- and a test split, where 90% of the data is kept as the training set for the CycleGAN model.

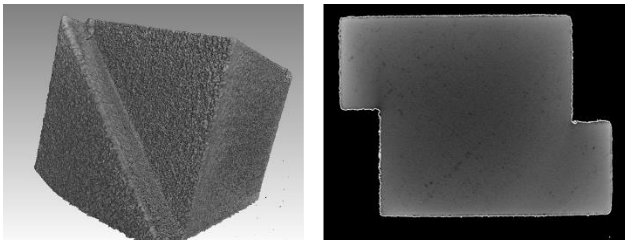

Figure 2: Overview of the used specimen for pore segmentation in CT data. Left: 3D-view of the specimen. Right: 2D-slice.

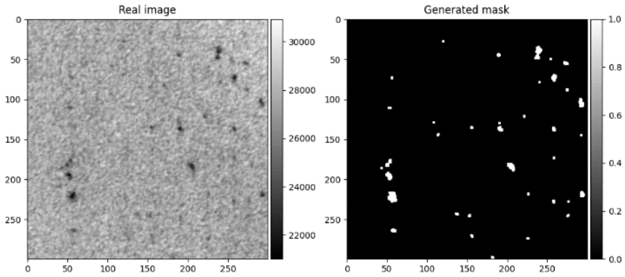

Figure 3: Example of a constructed binary mask for a real CT patch as ground truth.

Figure 4: Examples of manually extracted tearings patches and masks of tearings.

3.2. QC of sheared edges of fine blanked parts

Fine blanking allows for the mass production of tolerance-compliant sheet metal workpieces. The quality of fine blanked parts is defined by the state of the smooth cut section as described in VDI 2906 [14], that is reduced by the die-roll, tearings and the cut-off zone. To allow for the quality-driven optimization of fine blanking production processes, Trauth et al. proposed the development of a 100 % capable inline machine vision system that assesses the quality of fine blanking sheared edges and acquired an image dataset of side views of fine blanked parts for that purpose [3]. In a preliminary work, it was found that Convolutional U-Nets [15] are well suited to measure the height of the cut-off zone based on the dataset from [3, 16]. However, it was observed that examples of tearings are underrepresented in the dataset, eventually hampering the recognition of tearings with the used approach. To investigate this further, [16] propose to artificially increase the number of tearings using generative data augmentation techniques [17]. Herein, we took the dataset from [16] and extracted 40 unpaired patches of tearings and their respective masks (cf. Figure 4).

3.3 Implementation and training of CycleGANs

To account for the different image domains, two different architectures for the CycleGANs were used: For the generation of pores in CT images, a custom implementation of the original proposal of Zhu et al. [11] was used, modifying the input- and output dimensions in accordance to the gray value images. For the generation of tearings on sheared edges of fine blanked parts, a pre-trained, out-of-the box implementation provided by the TensorFlow framework [18] was used without further modifications. The trainings were performed on a NVIDIA GV100 GPU with 32 GB of VRAM. For the training of the individual CycleGAN models the same approach was followed: Each model was trained on the respective training sets, splitting them further into a training- and validation part using an 80-20 ratio. The training split was augmented during training using random flipping and -cropping of the images to a size of 256 x 256 px. The resulting images were normalized in accordance with the network specifications. Finally, both models were trained for 100 epochs, lasting in the case of the larger CT patch dataset up to 40 hours. During the training, the network behaviour was monitored on the validation split to prevent overfitting.

4. Experimental results

4.1. CT pore image generation

To generate realistic synthetic CT patches of the specimen, the trained CycleGAN model is used in its inference mode. First, a simple binary mask consisting of round pores is randomly generated. The number of pores is randomly drawn from a uniform distribution between 20 and 50 pores per image. This mask is forwarded through the respective generator part of the CycleGAN model to generate a first synthetic CT-patch. In principle, this first image would be already sufficient to enrich the real dataset. However, the simple generated mask does not really correspond to a realistic pore segmentation, since real pores encounter various irregular forms. To create more realistic masks, the generated synthetic image is forwarded through the contrary generator of the CycleGAN model, whose output can be interpreted as a probability map for the occurrence of pores. After thresholding this output at a high threshold value (0.9), the resulting mask can be forwarded through the image-generator once again, generating the final synthetic CT patch. The second forwarding of the mask through the image-generator ensures that the resulting synthetic image is generated from the given mask, hence providing a distinct ground truth segmentation for the synthetic image inherently. Figure 5 shows a qualitative example of the complete generation pipeline, starting from a simple mask (top left) towards the final synthetic CT patch (bottom right). Visually, the obtained synthetic samples match the real CT patches as displayed in Figure 3 closely.

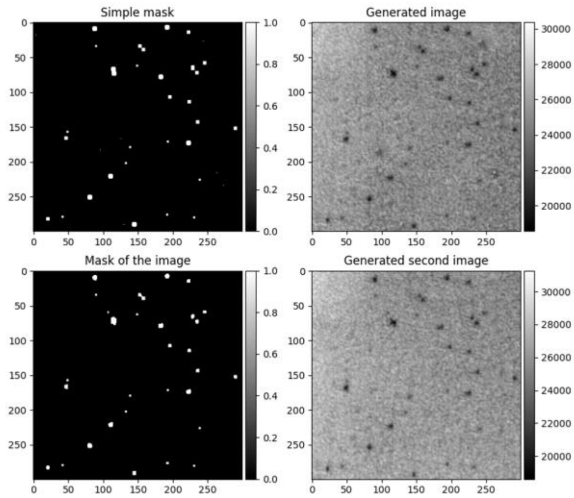

Figure 5: Example of the pipeline for CT patch generation. Starting from a simple mask, the patch and the image are refined through a series of forward-operations by the trained CycleGAN model.

4.2 Generation of tearings on sheared edges

For QC of fine blanking processes, the available dataset does not provide a suitable amount and variance of tearings on the sheared edge to build a model that is able to reliably recognize tearings even in rare forms and positions. As a consequence, it was investigated whether the trained CycleGAN model can be utilized to augment the dataset by generating synthetic tearings. The synthesized tearings can subsequently be inserted into the original data at desired positions both in the image and mask domain. An advantage of the usage of a CycleGAN is the possibility to integrate multiple tearings with arbitrary shapes and positions into the data and the ground truth masks with all relevant quality characteristics (size, number, etc.) in the same step. For this initial study, a simple mechanism generating and translating ellipses into elliptically shaped tearings using a trained CycleGAN model was tested. Given such a generated elliptical mask, the CycleGAN model generates an image patch containing a synthetically generated sheared edge surface, too. To maintain the original appearance of the target image, this mask of the synthesized ellipse can be multiplied with the synthetic patch leaving only the synthetic tearing of interest. As the fine blanked part is located almost at the same position in all images of the dataset, a region of interest was defined to paste the elliptical tearings at random positions on top of the sheared edge within that region. Figure 6 shows an example of a fine blanked part image before and after pasting multiple elliptical tearings, displaying plausible, synthetic defects.

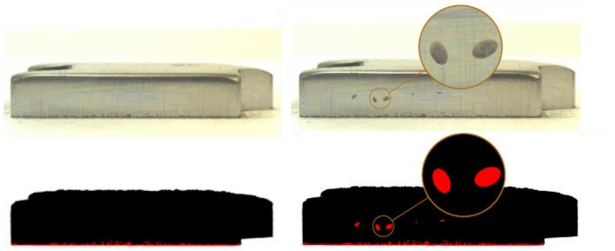

Figure 6: Original side view of a fine blanked workpiece and corresponding ground truth mask (left) and augmented with multiple synthezised elliptical tearings (right).

5. Conclusion

The main obstacle for the broader adoption of DL methods in industrial quality assurance use cases is the scarcity and quality of an available dataset. CycleGANs, being trained in a semi-supervised manner, are able to generate realistic synthetic samples suited to augment given datasets by either full images directly or by altering the appearance of existing images. In this study, it was shown how CycleGAN models can be trained and used to augment datasets for two industrial QC use cases, a CT-based pore detection application for AM parts and an image-based quality control of fine blanked workpieces. The proposed method yields promising results with visually convincing sample quality, indicating the potential benefit of CycleGANs for building more reliable QC models by largely extending sparse or unbalanced datasets with annotated synthetic images. However, the exact benefit of synthetic data for QC needs to be determined in future work. Further, both models show no particular domain dependency, being thus applicable to other industrial image datasets as well. Concluding, CycleGANs can be considered as a valid method to extend the data basis both with synthetic samples and the corresponding ground truth masks if the given amount of annotated data is not sufficient.

6. Future work

Direct subsequent work will investigate whether the addition of samples synthesized by the implemented CycleGAN models to the use case datasets leads to higher recognition rates when deriving corresponding segmentation models, such as the U-Net from [16]. Therein, the authors plan to perform a quantitative analysis of the quality of generate samples and defects. In accordance, some parameters of the initial mask generation will be optimized to align the real- and synthetic data distributions better. Instead of synthesizing elliptically shaped tearings, common data augmentation techniques such as mirroring or resizing and other affine transformation could be used to generate tearings masks that resemble the actual appearance more tightly. For the CT use case, in a first step the dataset for the image synthesis will be extended by including patches from the transition areas as well. Following, a synthesis of complete CT slices directly will be investigated for more universal applicability. Beyond this, the authors look forward to extent the augmentation approach to even more sophisticated models, such as proposed recently in [19].

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

[1] Alonso V, Dacal-Nieto A, Barreto L, Amaral A and Rivero E 2019 Industry 4.0 implications in machine vision metrology: an overview *Procedia Manufacturing* **41** 359–66

[2] Caggiano A, Zhang J, Alfieri V, Caiazzo F, Gao R and Teti R 2019 Machine learning-based image processing for on-line defect recognition in additive manufacturing *CIRP Annals* **68** 451–4

[3] Trauth D, Stanke J, Feuerhack A, Bergs T, Mattfeld P and Klocke F 2018 A characterization of quality of sheared edge in fine blanking using edge-computing approach *Procedia Manufacturing* **15** 578–83

[4] Weimer D, Scholz-Reiter B and Shpitalni M 2016 Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection *CIRP Annals* **65** 417–20

[5] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (MIT Press)

[6] Grozmani N, Chupina D and Schmitt R h 2020 Investigation of the independence of the best assembly orientations with respect to X-ray source parameters in industrial computed tomography *10th Conference on Industrial Computed Tomography, Wels, Austria (iCT 2020)* 1–6

[7] Wohlgemuth F and Hausotte T 2020 Convergence Behaviour of Numerical Measurement Uncertainty Evaluation using a Virtual Metrological Computed Tomography System *iCT*

[8] Fuchs P, Kröger T, Dierig T and Garbe S C 2019 Generating Meaningful Synthetic Ground Truth for Pore Detection in Cast Aluminum Parts *iCT*

[9] Wohlgemuth F and Hausotte T 2021 Realistic Simulation of Specific CT-Systems in aRTist 2 *DGZfP-Jahrestagung 2021*

[10] Neufeld, Deniz, Würfl, Tobias, Gruber, Roland, Schön, Maier and Andreas 2019 Realistic Image Synthesis of Imperfect Specimens using Generative Networks *iCT*

[11] Zhu J-Y, Park T, Isola P and Efros A A 2017 Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

[12] Rühle B, Krumrey J F and Hodoroaba V-D 2021 Workflow towards automated segmentation of agglomerated, non-spherical particles from electron microscopy images using artificial neural networks *Scientific reports* **11** 4942

[13] Wong V W H, Ferguson M, Law K H, Lee Y-T T and Witherell P 2021 Automatic Volumetric Segmentation of Additive Manufacturing Defects with 3D U-Net

[14] VDI-Fachbereich Produktionstechnik and Fertigungsverfahren Schnittflächenqualität beim Schneiden, Beschneiden und Lochen von Werkstücken aus Metall; Feinschneiden *(*VDI 2906 Blatt 5*)* (VDI-Gesellschaft Produktion und Logistik)

[15] Ronneberger O, Fischer P and Brox T 2015 U-Net: Convolutional Networks for Biomedical Image Segmentation

[16] Wolfschläger D, Woltersmann J-H and Schmitt R h 2021 Sheared edge defect segmentation using a convolutional U-net for quality assessment of fine blanking workpieces: Submitted *Journal of the International Societies for Precision Engineering and Nanotechnology*

[17] Sandfort V, Yan K, Pickhardt P J and Summers R M 2019 Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks *Scientific reports* **9** 16884

[18] TensorFlow Developers 2021 *TensorFlow* (Zenodo)

[19] Eastwood J, Newton L, Leach R and Piano S 2022 Generation and categorisation of surface texture data using a modified progressively growing adversarial network *Precision Engineering* **74** 1–11