A data-centric approach to anomaly detection in layer-based additive manufacturing

Zeiser, A.; Özcan, B.; Kracke, C.; Stein, N. van; Bäck, T.H.W.

Application

Alexander Zeiser*, Bekir Özcan, Christoph Kracke, Bas van Stein and Thomas Bäck

A data-centric approach to anomaly detection in layer-based additive manufacturing

Ein datenzentrierter Ansatz für Anomaliedetektion in schichtbasierten additiven Fertigungsverfahren

https://doi.org/10.1515/auto-2022-0104
Received August 31, 2022; accepted December 6, 2022

Abstract: Anomaly detection describes methods of finding abnormal states, instances or data points that differ from a normal value space. Industrial processes are a domain where predicitve models are needed for finding anomalous data instances for quality enhancement. A main challenge, however, is absence of labels in this environment. This paper contributes to a data-centric way of approaching artificial intelligence in industrial production. With a use case from additive manufacturing for automotive components we present a deep-learning-based image processing pipeline. We integrate the concept of domain randomisation and synthetic data in the loop that shows promising results for bridging advances in deep learning and its application to real-world, industrial production processes.

Keywords: additive manufacturing; anomaly detection; domain randomisation; infrared imaging.

1 Introduction

Predictive models approach industrial manufacturing complexity by facilitating the assessment of multidimensional relationships. They enable optimisation of process and product quality by unveiling machining dependencies and root-causes of defects. Typically, industrial processes generate heterogeneous datasets consisting of process measurements, quality inspection feedback and maintenance events that only in combination expose causal interactions. One main challenge in industrial production is data quality and pre-processing so that reliable pipelines for modelling can be built [1–3]. Quickly changing conditions and drift, missing labels, highly imbalanced datasets and noise are some examples of impediments that need to be handled [4]. Due to that, open issues still exist in integrating and exploiting potentials of machine learning in actual series production as literature on industrial applications reveal [5].

A main challenge we like to highlight in this paper is absence of labels in industrial production data, important for the development of automated anomaly detection pipelines. We present a deep-learning-based image processing pipeline, applied to a use case from additive manufacturing for automotive components. Our main contribution is a data-centric approach that integrates the concept of domain randomisation and synthetic data in the loop for deep learning (DL) model development. It shows promising first results for bridging advances in DL and its application to real-world, industrial production processes.

The remainder of this paper is structured as follows: Sections 2 and 3 introduce the problem environment, from a broader perspective of data processing and analysis in industrial production to the more focused aspect of anomaly detection for improving process quality. As a use case from production, we present an additive manufacturing process to analyse open issues in the industrial application of an automated process monitoring based on DL. Section 4 presents pipeline components and an approach of synthetic data in the loop for developing anomaly detection models for real-world production data. Finally, in Section 5, we suggest future research directions.

2 Implications of industrial data

The industrial production system, opposed to a laboratory set-up, is characterised by many environmental and uncontrollable influences, leading to noise and disturbance in

***Corresponding author: Alexander Zeiser**, BMW Group, Ohmstraße 2, 84030 Landshut, Germany; and Leiden Institute of Advanced Computer Science, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands, E-mail: alexander.az.zeiser@bmw.de
Bekir Özcan and Christoph Kracke, BMW Group, Ohmstraße 2, 84030 Landshut, Germany
Bas van Stein and Thomas Bäck, Leiden Institute of Advanced Computer Science, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

data acquisition and mediocre data quality. As a result, any analysis is reliable only after careful data pre-processing [1]. Output targets and production key performance indicators, however, pursue high efficiency with low scrap rates and machine down times. Advanced process monitoring as well as predictive methods need to be adaptive and overcome such limitations, like robustness, to be applicable in series production. As stated in [6] less than 30% of potential in application of data analytics methods, especially machine learning, is exploited in manufacturing. It illustrates that the gap between advances in research and application in real-world is still high. Highlighted reasons are needed expert knowledge, very problem-specific solutions and uni-dimensional concentration on optimising model output rather than input data.

The importance of a context-based data processing was highlighted also by Andrew Ng with the 'Data-Centric AI Competition' and adopted by Motamedi et al. [2], focusing on actions of dataset preparation and data quality enhancement before training and fine-tuning a predictive model (data-centric before model-centric). Ng shows that an increase in data quality can exert far greater influence on the prediction accuracy than hyperparameter optimisation of a machine learning model alone [7]. In the context of real-world industrial applications this is an encouraging and advisable approach as well to develop maturity levels further (from descriptive towards self-optimising/prescriptive). Main success factor for achieving promising results is representative, sufficiently large and high-quality data.

Another aspect of industrial processes is the amount of data that is created and that is available for knowledge discovery. However, the process from merging of data bases to actual analysis becomes a difficult task if unique identifiers, timestamps and labels (e.g. from quality inspection) are missing. Unsupervised approaches are favourable in these environments to get insights from data, nevertheless. Another approach is domain randomisation or adaptation, hence, utilising synthetic data for model development, either simulated or abstracted from the real process. Major benefits are creation of labelled data and the high amount of data at low cost, especially for under-represented classes as in anomaly detection cases. In principle, the concept intends to reach high generalisation for the real-world data by transfer learning with models trained solely on synthetic data [8, 9].

3 Anomaly detection in industrial processes

Anomaly detection describes methods of finding abnormal states, instances or data points that differ from a normal value space [10]. The three categories of supervised, semi-supervised and unsupervised each summarise different techniques and methods from statistics (e.g. z-score) or machine learning (e.g. One-Class SVM, Autoencoders, LSTM). The term is predominantly used for highly unbalanced problems. Often no labels are available for learning or classes for different states are not known [11]. The unsupervised approach is applied in several domains, such as medicine (e.g. detection of critical cardiac arrhythmia, tumor detection with computed tomography), banking (e.g. fraudulent financial transactions, payments with stolen credit cards), security (e.g. surveillance, document forgery, network intrusion) but also engineering (e.g. critical state detection) [10]. Besides point anomalies (one data instance lies out of the normal data region) Chandola et al. define two other types of anomalies: contextual anomalies (instance of data is anomalous only in a certain context but not in another) and collective anomalies (individual values lie within normal data region but as a collection of related data instances they form an anomaly) [12].

3.1 Use case: binder jetting additive manufacturing

Additive manufacturing (AM) is a widely used technology that comprises many sub-technologies. What most have in common is a layer-wise construction with deposition of new material. However, the actual way of how these layers are created differentiates sub-technologies and is standardised by ISO 17296 [13]. One major benefit over classic manufacturing techniques is free form design of complex geometries. Another advantage is a tool-independent production of different shapes and therefore a quick adaptation to new designs or product updates. Typically, AM is used in rapid prototyping, small series and small dimensions. However, advances in technology enable also medium to high scale production of specific components. Also larger components can be manufactured with AM technologies, especially binder jetting. Due to comparably low working temperature nearly no heat induced shrinkage, cracks or porosities appear that typically hinder other AM technologies from manufacturing bigger dimensions [14].

In series production binder jetting is often combined in a multi-stage process with other manufacturing techniques, like casting. One example are automotive cylinder heads of BMW straight-four and straight-six engines. Advantages of both conventional and modern manufacturing are combined in this way. In the following current developments are described.

3.2 Related work

A majority of research on optical process monitoring and computer vision related to additive manufacturing is concentrating on powder bed fusion. Quality and defect predictions are built especially on (image) data from the melt pool. Nevertheless, also other AM processes are adduced for image-based process monitoring and anomaly detection for quality enhancement. A comprehensive summary can be found in [15].

With a focus on image processing, various algorithms are discussed in [16]. In an application to a small sample of image data from an extrusion process they are compared based on accuracy for defect prediction. The trained data model is constantly fed with a stream of new images and classifies in real time. As soon as a production error is discovered, the process is automatically stopped. However, issues like generalisation and processing of large datasets hinder application in series production. Günther et al. [17] describe requirements for condition monitoring for binder jetting and propose an image-based defect detection. Research is focused on nozzle failures that lead to work-piece defects. In a series of steps the work-piece shape is extracted from the recorded image and defect analysis is performed based on the transformed binary image. The distribution density of black and white pixels along the printing direction indicate a printing failure [17].

Current literature presents and discusses approaches for process monitoring and defect detection with computer vision. Even though the need for anomaly detection in processes with un-labelled data is stated research concentrates on process engineering parameter setting and is often related to detection of material-specific and mechanical defects [15, 18]. The development of methods for bigger datasets, series production and the implication associated with it is still at an early stage.

4 Data processing pipeline for the AM binder jetting *in-situ* monitoring

In order to bring anomaly detection into deployment, first, process historic data is analysed and used for model development. The pipeline components are described in the following.

4.1 Data acquisition

Data from manufacturing at BMW plant Landshut consists of processing measurements, machining parameters, ambient conditions and quality inspection. However, in focus of our work is image data coming from an infrared (IR) camera mounted in the inside of the printing room. During the process of sand and binder application an infrared light activates the binder that bonds loose sand particles. With a normal camera setup only a flat sandy surface can be seen. However, as the process heats up the whole powder bed, shapes of the printed part become visual by infrared imaging. Another benefit of IR-imaging by visualising temperatures is to make information of energy deposition onto the print bed available for process monitoring. Since binder and energy deposition have a major influence on dimensional accuracy a systematic image data analysis is needed to support process optimisation decisively. In terms of a live monitoring system the objective is to detect temperature and geometric anomalies and provide information to a worker about location, layer and severity. The setup in place is triggered automatically by the machine control unit (PLC) and generates an image per each layer. Depending on the part produced a complete image stack consists of 600–800 images. An example is shown in Figure 1. Images are recorded in grayscale where the pixel values (0-black, 255-white) correspond to a fixed range of temperatures to assure comparability between print jobs and over time.

4.2 Data cleaning

Data cleaning is performed to have a dataset free of incorrectly acquired records and an important step for improving data quality. With the goal of learning patterns it is necessary, especially for training machine learning models,

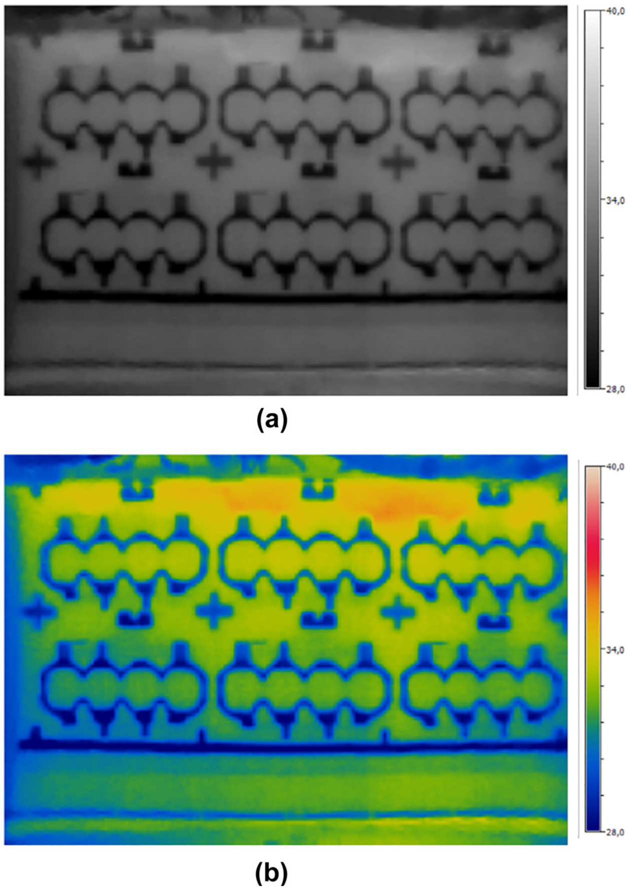

Figure 1: Example thermal image from an AM print job (layer 500 of 700) with different visualisations. (a) Original grayscale image. (b) False colour transformed image for an improved visualisation.

to not detract learning from erroneous data that falsifies underlying dependencies. However, at the same time the cleaned dataset must still accurately represent the distribution of the real sample to generalise well [19].

For the binder jetting process, two types of anomalies must be distinguished. First type (Type 1 anomaly) is related to possible sensor malfunction or human intervention that leads to incorrect data acquisition. These anomalies can also be referred to as actual process flow outliers that lead to data quality issues. Examples are shown in Figure 2, where the camera triggered in a wrong process instance. The anomalous or outlier image, in this context, does not present a clear view onto the print bed.

Type 2 anomalies, in contrast, are defined as deviations from the normal manufacturing procedure, potentially resulting in product related quality issues like defects or dimensional inaccuracy. Figure 3 shows an example of Type 2 anomaly. These anomalies show an explicit deviation from the print model and are mostly based on geometrical features within the print bed.

Data cleaning refers only to Type 1 anomalies since Type 2 anomalies are the actual matter of investigation in the dataset. In later model development for Type 2 anomaly detection the Type 1 anomalies must be filtered out to minimise pseudo-defects. Since it is a classification task a DL model can potentially solve the distinction of good and bad images for Type 1 anomalies.

This has been tested with a convolutional neural network (CNN) of six convolutional 2D layers, each followed by 2D max-pooling and batch normalisation layer. Finally, a flatten layer and two dense layers are used for classification (activation function: ReLU, optimiser: Adam with default values). Early stopping is applied based on validation accuracy. The starting set consists of 22,000 images (70/30% good/bad) for training, of which 20% are hold out for validation to monitor model generalisation during training. 10 learning repetitions were carried out with randomised weights initialisation. Standard data augmentation was performed on the training data within the Tensorflow model. Testing was performed with additional 5000 images (70/30% good/bad distribution) on each of the 10 models. Train and test images have been randomly sampled from different jobs and printers over a production period of 3 months.

The mean performance including 95%-confidence interval is shown in Figure 4 in terms of the confusion chart. The network is clearly able to distinguish well between 'good' images of the process working fine and 'bad' images (Type 1 anomalies) as exemplified in Figure 2. Labelling was performed manually by reviewing thumbnails of a collection of images. It must be pointed out that for process flow outlier images this is still a bearable effort since these Type 1 anomalies are simple to distinguish for the human vision. A decision about the labels can be done easily, even by looking at a collection of images at a time. Type 2 anomalies must be treated differently as defect characteristics may not be that explicit and manual labelling is not efficient.

Even though data of three months was used and 10 different models were compared, results must be examined closely for a long-term, on-line application. One main limitation is data drift and changing conditions over time. This can influence performance negatively, as discriminative features may disappear or borderline cases are not effectively detected. A regular re-training is needed for such application to be always based on current data and related variability.

Figure 2: Example of Type 1 anomalies. Print bed and components are hidden by print head due to falsely timed camera trigger. This is a data quality issue as pseudo-defects need to be avoided in the model training phase.

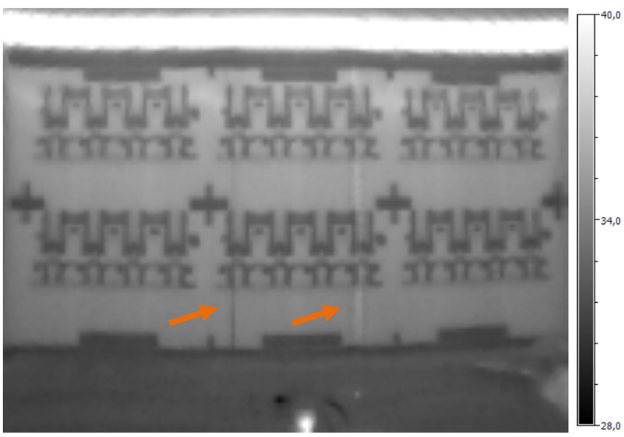

Figure 3: Example Type 2 anomalies that influence process quality which results in product defects.

Predicted Category	True Category: Bad images	True Category: Good images
Bad images	0.987 ± 0.006	0.004 ± 0.001
Good images	0.013 ± 0.006	0.996 ± 0.001

Figure 4: Normalised confusion chart for test set (5 000 images) of Type 1 anomalies. The chart shows mean and confidence interval of the test set as input to the 10 trained models. An F1 score of 0.98 could be reached.

4.3 Data preparation

Due to space constraints, wide-angle lenses with 80° opening angles are used that create a radial distortion. By printing straight checkerboard patterns, this lens distortion can be calibrated. New images are undistorted by the camera matrix and distortion coefficients in a post-processing step. Camera calibration functions of the Python package OpenCV are implemented [20]. Checkerboards of different dimensions and viewing angles are tested and examined by the re-projection error. As multiple machines work in parallel, slightly distinct camera mounting positions result in different viewing angles onto the print bed. To align images an image registration algorithm is implemented using OpenCV. Here, the enhanced correlation coefficient maximisation (ECC) algorithm, adopted from [21], outperforms other feature-based algorithms like Oriented FAST and Rotated BRIEF (ORB) with brute force matching. ECC is independent from photo-metric distortions like contrast and brightness, hence, well applicable to thermal images where temperatures actually translate into brightness or colour. Transformation parameters can be calculated by target and source image of the exact same layer and is valid for the whole stack of images of one print job. Finally, images are cropped to the region of interest, namely the print bed.

4.4 Synthetic data in the loop for DL model development

The aim of a live monitoring system is to detect anomalies in (near-)real time that become sources of product defects. This information must be passed to a worker to either stop the print job, to scrap affected parts directly after process completion without quality test or to perform a more target-oriented inspection. Additionally, process optimisation by parameter adaptation is a prospective way to approach quality enhancement in the long term.

Figure 3 shows a noticeable Type 2 anomaly, namely striations. Vertical striations from excess binder, sand or nozzle-clogging, foreign objects and porosities can lead to work-piece internal defects that remain hidden during visual inspection. Resulting layer disintegration that affects the material strength cause broken sand cores during subsequent casting steps. Other defects like layer shifts and sand agglomerates cause dimensional inaccuracies. Due to the large amount of image data in the industrial setup a manual labelling and training set preparation is very cost intensive. We therefore see huge potential in approaching the anomaly detection problem with transfer learning from synthetic data that can be generated and automatically labelled in only a fraction of time. In the AM binder jetting case this data is abstracted from the real process. Figure 5 illustrates the workflow of integrating synthetic data into model development for the real data. It follows the hypothesis that learning from synthetic data allows for generalisation to real-world data.

For the synthetic data creation process, as shown in Figure 6, we utilise the Standard Tessellation Language (stl) file generated from the work-piece CAD file. As in the preparation of the AM print job, we slice the stl file into layers and create Support Vector Graphic (svg) files that define the layer-by-layer contours of the work-piece by spatial coordinates. We adapt domain randomisation in terms of background and work-piece grayscale dispersion on texture maps, which simulates the changing temperature distribution of printbed and component. By masking the shapes characterised by the edge coordinates and the parameterised work-piece position with the textures, we obtain random layer-by-layer temperature distributions without reference to previous layers. Furthermore, the layer-to-layer temperature decay and the blurring of the sharp work-piece contours are adjusted by parameters to approximate the real AM process. The randomisation intervals of these job parameters are defined by minimum and maximum values of a random subsample of real images. This procedure is performed both layer-by-layer as well as job-wise to increase variability.

In addition, we extended the synthetic data with a set of aforementioned defects that are visually defined and placed on specific layers. These anomalous jobs have their own respective parameters such as geometry, location and duration of anomaly in terms of number of layers affected with a randomised domain to prevent bias in the data. Synthetic images, as shown in Figure 7, are labelled accordingly so that the resulting dataset of clean and anomalous jobs can be used for further research and validation of developments.

As a first experiment with the synthetic dataset, we want to evaluate supervised classification methods. Therefore, we use similar CNN architecture and parameters as for the Type 1 anomaly classification, with a softmax layer added for prediction of class membership probability and train it on 12 500 synthetic images, each 2 500 per class (healthy/no defect, agglomerates, foreign objects, porous and striation). In the next step, we evaluate the model on 684 real images, manually labelled for the anomaly type 'striation', such as shown in Figure 3. On average, the model

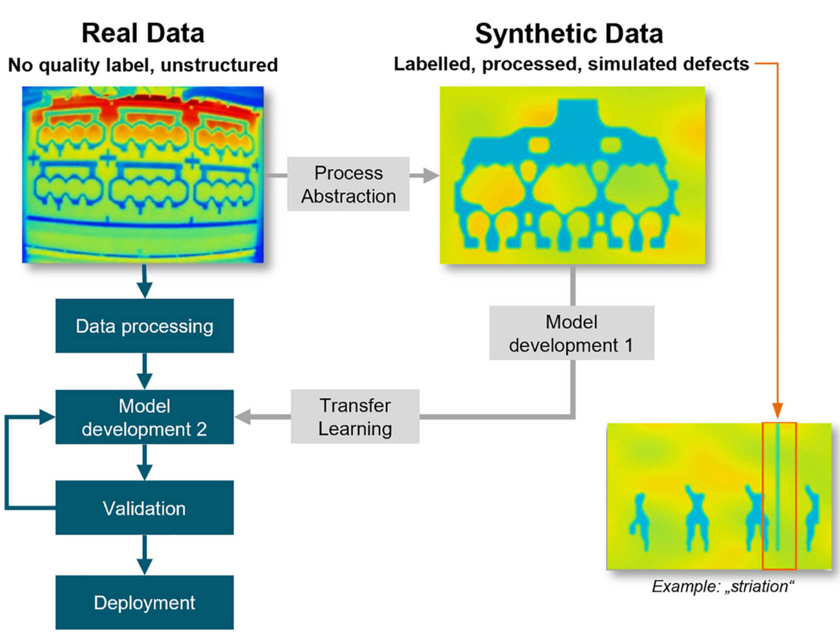

Figure 5: Synthetic data in the loop (in false-colour representation for better visualisation): data from the real process is abstracted and synthetic data is created by domain randomisation. Domain randomisation follows the hypothesis that if distributions are well projected from real to synthetic data and variability is significantly high in the synthetic dataset, then DL models, trained only on synthetic data, will generalise well on the real-world data [8, 22].

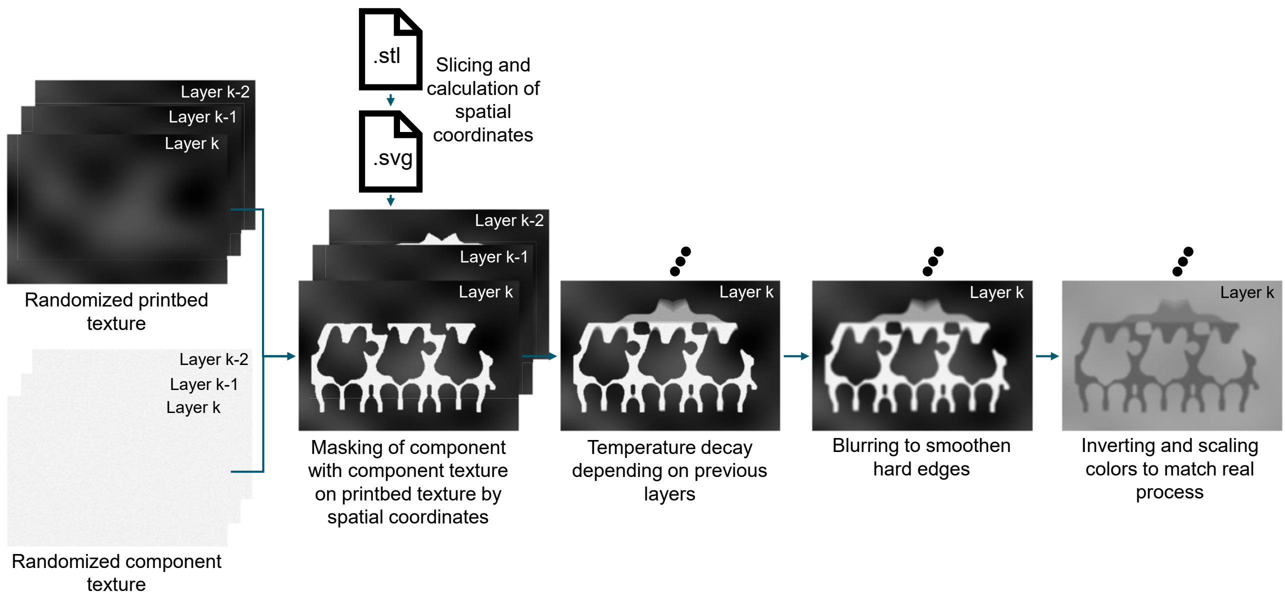

Figure 6: Synthetic data creation process: firstly, we slice a 3d-model and mask printbed and component with the created randomised textures. Then, we create a correlation of layer k to previous layers to model temperature decay over time. In a subsequent step, we blur, invert and scale the images to achieve an closer approximation to the data of the real AM process.

Figure 7: Anomaly types included in the synthetic data (red: highlighted location of anomaly). Agglomerate: accumulation of unwanted loose components into a solid compound. Foreign object: unknown objects on the printbed. Porous: temperature difference in the direction of print head movement. Striation: vertical striations from excess binder, sand or nozzle-clogging.

decided with a confidence of 74% that the real images show a 'striation'. Followed by defect type 'foreign object' with 22% probability. The predicted probability for the images to fall within the 'good' class (no defect) is at only less than 2% on average, showing first proof of the concept shown in Figure 5 for transfer learning by synthetic data. Figure 8 shows the responding confusion chart.

We consider 12 500 images for training a sufficiently large amount of data for our test case. However, typically deep learning performs better when trained on big datasets.

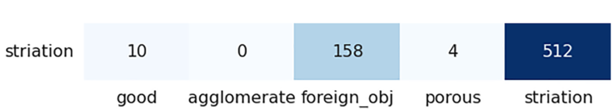

Figure 8: Confusion chart of the predicted class membership for real striation images: 512 real striation images are predicted correctly as striations, whereas 158 images are misclassified as foreign objects.

Although, the proof of concept is promising, the ML performance must be highly increased for an on-line application as false positives and negatives may disturb a smooth production flow more than it supports workers on quality decisions. Again, data drift needs to be considered additionally. In this case, not only re-training of ML models alone but also the creation of new and adapted synthetic data are essential.

5 Conclusions

One major challenge in advancing deep learning potentials for industrial processes, e.g. quality improvements, is access to clean and labelled data for model development. A more data-centric approach to artificial intelligence shifts attention towards data pre-processing and quality. With this model input optimisation deep learning, like anomaly detection techniques will likely perform better in real-world scenarios. We consider domain randomisation a valuable approach in developing anomaly detection pipelines, especially when quality labels are absent or not trustworthy. Bringing synthetic data into the loop shows considerable potential. Economically due to low cost of creation without any physical production, as well as technically due to exact labelling. First results on defect prediction with a CNN trained only on synthetic data are promising to investigate this approach further. Future work will direct towards enhancing domain randomisation in order to explore effects of higher variability levels in synthetic data. This can be implemented by:

- extending the randomisation intervals of the adjustable parameters, i.e. the range of brightness, decay and translucency of layers as well as geometric dimensions.
- including borderline cases of harmless and small synthetic anomalies in the non-anomalous class.
- combining multiple anomalies within one image.

At this stage, tests were only performed with defect type striation. In the future we will extend tests to other defects. Further anomalous images from the real dataset must be labelled manually but likely allow for more realistic synthetic defects. Temperature patterns of porosities could be further inspected and remodeled by higher resolved thermographic images, i.e. with changes in the camera set-up. Also, only object-like defects were considered as anomalies so far. Additionally, quality related trends, patterns and higher order features may be present. Deeper analysis on unknown, more implicit anomalies must be performed. Autoencoders present high potential for the kind of task. Generative models are also further to be investigated due to their photo-realistic recreation capabilities for synthetic data. GAN-based architectures like f-AnoGAN [23] or ALAD [24] are of particular interest. Additionally, supervised-unsupervised combinations are to be tested as shown by Balzategui et al. [25] and using small data for labelling big data with no prior labels. Another aspect for future research is time and spatial dependency of (thermal) images as input for anomaly detection. Approaches of combining different neural network types like CNN with Recurrent Neural Networks (e.g. LSTM) to include the aspect of time dependency are discussed in recent literature but need further elaboration for a series production case [26].

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

References

[1] J. Krauß, M. Frye, G. Teodoro, D. Beck, and R. H. Schmitt, "Selection and application of machine learning-algorithms in production quality," in *International Conference ML4CPS*, J. Beyerer, Ed., 2019, pp. 46—57.

[2] M. Motamedi, N. Sakharnykh, and T. Kaldewey, "A data-centric approach for training deep neural networks with less data," in *35th Conference on Neural Information Processing Systems*, 2021.

[3] L. Cai and Y. Zhu, "The challenges of data quality and data quality assessment in the big data era," *Data Sci. J.*, vol. 14, pp. 1—10, 2015.

[4] G. Kronberger, F. Bachinger, and M. Affenzeller, "Smart manufacturing and continuous improvement and adaptation of predictive models," *Procedia Manuf.*, vol. 42, pp. 528—531, 2020.

[5] A. Dogan and D. Birant, "Machine learning and data mining in manufacturing," *Expert Syst. Appl.*, vol. 166, p. 114060, 2021.

[6] N. Henke, J. Bughin, and M. Chui, *The Age of Analytics: Competing in a Data-Driven World*, London, McKinsey & Company, 2016.

[7] A. Ng, *MLOps: From Modelcentric to Data-Centric AI*, 2021. Available at: https://www.deeplearning.ai/wpcontent/uploads/2021/06/MLOps-From-Model-centric-to-Data-centric-AI.pdf.

[8] J. Tobin, R. Fong, and A. Ray, "Domain randomization for transferring deep neural networks from simulation to the real world," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2017, pp. 23—30.

[9] S. Z. Valtchev and J. Wu, "Domain randomization for neural network classification," *J. Big Data*, vol. 8, no. 1, p. 94, 2021.

[10] M. Goldstein and S. Uchida, "A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data," *PLoS One*, vol. 11, no. 4, pp. 1—31, 2016.

[11] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection for discrete sequences: a survey," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 5, pp. 823—839, 2012.

[12] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: a survey," *ACM Ref. Form.*, vol. 41, no. 15, p. 15:1—15:58, 2009.

[13] International Organization for Standardization, ISO/TC 261, "ISO 17296-2:2015 part 2: overview of process categories and feedstock," *Additive manufacturing - General principles*, 2015.

[14] I. Gibson, D. W. Rosen, and B. Stucker, *Additive Manufacturing Technologies*, Boston, MA, Springer US, 2021.

[15] D. Gierson and A. Rennie, "Machine learning for advanced additive manufacturing," *Encyclopedia*, vol. 3, no. 5, pp. 576—588, 2021.

[16] S. Trinks and C. Felden, "Smart Factory — konzeption und Prototyp zum Image Mining," *HMD Prax. Wirtsch.*, vol. 56, no. 5, pp. 1017—1040, 2019.

[17] D. Günther, M. Fahimi Pirehgalin, I. Weiß, and B. Vogel-Heuser, "Condition monitoring for the binder jetting AM-process with machine learning approaches," in *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, vol. 1, 2020.

[18] L. Scime, D. Siddel, and S. Baird, "Layer-wise anomaly detection and classification for powder bed additive manufacturing processes," *Addit. Manuf.*, vol. 36, no. March, p. 101453, 2020.

[19] G. Y. Lee, L. Alzamil, B. Doskenov, and A. Termehchy, "A survey on data cleaning methods for improved machine learning model performance," *arXiv*, vol. 2109.07127, 2021.

[20] OpenCV, *OpenCV Camera Calibration*, 2021. Available at: https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html.

[21] G. Evangelidis and E. Psarakis, “Parametric image alignment using enhanced correlation coefficient maximization,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 10, pp. 1858–1865, 2008.

[22] P. Fuchs, T. Kröger, and C. S. Garbe, “Defect detection in CT scans of cast aluminum parts: a machine vision perspective,” *Neurocomputing*, vol. 453, pp. 85–96, 2021.

[23] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, “f-AnoGAN: fast unsupervised anomaly detection with generative adversarial networks,” *Med. Image Anal.*, vol. 54, no. May, pp. 30–44, 2019.

[24] H. Zenati, M. Romain, C. S. Foo, B. Lecouat, and V. R. Chandrasekhar, “Adversarially learned anomaly detection,” in *Proceedings of the 20th IEEE International Conference on Data Mining*, 2018.

[25] J. Balzategui, L. Eciolaza, and D. Maestro-Watson, “Anomaly detection and automatic labeling for solar cell quality inspection based on generative adversarial network,” *Sensors*, vol. 21, no. 13, pp. 1–22, 2021.

[26] Z. Zuo, B. Shuai, and G. Wang, “Convolutional recurrent neural networks: learning spatial dependencies for image representation,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2015, 2015, pp. 18–26.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/auto-2022-0104).

Bionotes

Alexander Zeiser
BMW Group, Ohmstraße 2, 84030 Landshut, Germany
aljosha.koecher@hsu-hh.de

Alexander Zeiser studied Mechanical Engineering at Politecnico di Milano. He focused on Manufacturing Technologies and Production Management. After his graduation, in 2020 he started working at BMW Group plant Landshut in the department of production digitalisation. Besides, he is a PhD student at Leiden Institute for Advanced Computer Science (LIACS), Netherlands.

Bekir Özcan
BMW Group, Ohmstraße 2, 84030 Landshut, Germany

Bekir Özcan graduated in Mechanical Engineering with specialisation in data mining and machine learning at the Technical University of Darmstadt. His main research interests are in the field of deploying deep learning in manufacturing.