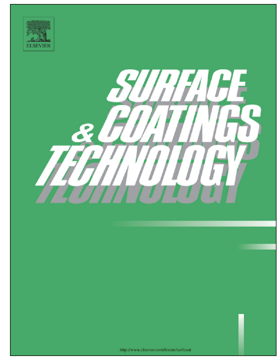
Journal Pre-proof



Automated evaluation of Rockwell adhesion tests for PVD coatings using convolutional neural networks

Bastian Lenz, Henning Hasselbruch, Andreas Mehner

PII: S0257-8972(20)30034-7

DOI: https://doi.org/10.1016/j.surfcoat.2020.125365

Reference: SCT 125365

To appear in: *Surface & Coatings Technology*

Received date: 11 December 2019

Revised date: 10 January 2020

Accepted date: 13 January 2020

IWT

Leibniz-Institut für
Werkstofforientierte
Technologien

Title:

Automated evaluation of Rockwell adhesion tests for PVD coatings using convolutional neural networks

Authors:

1. **M.Sc. Bastian Lenz (Corresponding Author)**
2. **Dipl.-Ing. Henning Hasselbruch**
3. **Dr.-Ing. Andreas Mehner**

Leibniz Institute for Materials Engineering - IWT

Badgasteiner Str. 3

28359 Bremen

Germany

Abstract

An automated method for the classification of the adhesion strength of thin PVD coatings applied on hardened steel substrates is presented in this study using deep neural networks. For the determination of the adhesion strength Rockwell-indentation tests were carried out according to VDI 3198. For this approach, pre-trained convolutional neural networks are adapted to classify microscopic images into the expected adhesion classes HF 1 to HF 6 using transfer learning with a dataset of 1650 already evaluated indentation images. The classification performance of the Matlab implemented network models AlexNet, GoogLeNet and inception-v3 is compared with test and verification images of Rockwell indentations. The inception-v3 network shows good accuracy for polished (roughness Sa < 20 nm), hardened steel substrates with deposited thin coatings of a thickness up to 5 µm. The classifications of the implemented models exhibit an agreement of approximately 85 - 90 % compared to human assessment. The evaluation is robust against disturbance variables such as different exposure times, brightness, image contrasting and magnifications. Different image capture devices can be used with no effect on the classification. The networks show promising results for automated industrial applications, such as in-line adhesion control in coating processes, as they do not require human operator support.

Key Words: Deep learning, convolutional neural networks, Rockwell adhesion testing, thin coatings, hard steel substrates

1 Introduction

The adhesion strength to the substrate material is one of the most important properties of technical coatings, as it significantly influences the applicability for technical applications, particularly for tribologically stressed components. Especially hard coatings, which are produced by physical or chemical vapour deposition (PVD/CVD), show considerable differences of the adhesion strength due to various PVD process parameters, such as process pressure, target power or bias voltage [1]. Additionally, the adhesion can significantly be increased with appropriate surface pre-treatments, for example plasma-nitriding. Interdependencies with other coating properties, for instance hardness, young's modulus or residual stress, lead to various failure models [2]. Therefore, a variety of adhesion testing methods have been developed. Well established techniques for thin, hard coatings are the scratch test and the Rockwell adhesion test.

The Rockwell adhesion test, defined by the VDI 3198 guideline (VDI: Verein Deutscher Ingenieure), is a well-established method for adhesion testing. A conical diamond is loaded on the surface with a defined force, according to DIN EN ISO 6508-1 [3]. The coating damage at the indentation imprint is assigned to six adhesion classes HF 1 – HF 6, which are defined in the VDI guideline 3198, as shown in figure 1 [4, 5]. For adhesion class HF 1, only fine cracks at the edge of the imprint but no delaminations are allowed. Adhesion class HF 2 is characterized by small, unconnected spallings. For adhesion class HF 3 the delaminations are larger in size and are allowed to be connected in the circumferential direction of the imprint. However, some areas at the imprint edge should be free of spalling. Adhesion class HF 4 shows delaminations present all around the indentation imprint, with comparable size to HF 3. The adhesion classes HF 5 and HF 6 (indicating technically insufficient adhesion) show far-reaching delaminations around the imprint, which form a connected area in HF 6.

The classification into the adhesion classes is usually done by light microscopy with a subsequent visual inspection carried out by humans. This task is desired to be automated for various reasons. Firstly, the classification is time-consuming for large numbers of specimens, even though it is easy to perform on individual indentations. Typically, a sample must be taken from the production process and analysed by technical staff in the laboratory. For industrial applications, automatic classification could enable in-line coating adhesion testing, with little or no personnel effort. This would be a reasonable improvement for existing coating processes, both in terms of cost and time. Secondly, the visual classification is based on more or less subjective evaluation of the Rockwell imprint by comparison with the adhesion class images in figure 1. A wrong assessment of the adhesion strength has a major impact on the overall rating of a coating system. Consequently, an automatic classification method eliminates subjective assessments of the adhesion class.

There are different conceivable approaches for automated classification of microscopic pictures of Rockwell imprints. The delaminated area around the indentation imprint could be quantified by image analysis software, similar to existing approaches for the scratch test [6]. Using predefined thresholds for the delamination area would allow a categorisation into the different adhesion classes. This method can easily be implemented, but has some limitations. Firstly, the luminous exposure of the microscopic pictures and the colour of the coatings has to be very similar, to ensure comparable visual contrasts. Additionally, some coating systems might not show a clear contrast to the underlying substrate, preventing analysis of the delaminated area. Therefore, a more convenient and robust approach is needed for a safe automatic evaluation.

In this work the potential of deep artificial neural networks was investigated for classification of Rockwell indentations on a variety of hard PVD coatings on steel substrates. The output of these networks is the classification score for the six adhesion classes, whereby the class with the highest score becomes the prediction.

Deep artificial neural networks mimic the structure of biological neuronal networks in animals or humans. Their use for technical applications has increased significantly in recent years. A variety of different neural network architectures has been developed, each with different objectives and areas of operation [7]. For classification of images primarily convolutional neural networks (CNN) are used [8].Their architecture is characterized by an input- and output layer, with several hidden layers in between, as illustrated figure 2.

Convolutional neural networks take advantage of given hierarchical patterns in the input data and use them to assemble complex information from smaller and simpler patterns. CNNs typically use sequences of alternating convolutional and pooling layers. Convolutional layers reduce the dimension of 2D matrix input neurons (small image parts) by scalar multiplication with a weight vector (convolving) to 1D rows of output neurons (vector). The resulting vector of output signals represents patterns of the 2D input matrix. Learning of patterns results from iterative adjustments of the weight vectors by backpropagation. In convolutional layers all neurons share the same weight vectors. The position of the output neurons (vectors of neurons) represent the spatial position of the pattern in the image. Thus, the output of a convolutional layer represents a map of sub-features (patterns) of the image. Pooling layers reduce the amount of data by subsampling without reducing the dimension. The pooling can be implemented by different procedures, like max-pooling or mean-pooling. With max-pooling, the maximum value of each batch of features is calculated, while mean-pooling gives an average value for each batch. The next convolutional layer combines neighbouring sub-features to more complex features. The output of CNNs is typically given to one or more fully connected output layers, depending on the net architecture. For classification tasks, the number of neurons in the last fully connected layer is equal to the number of output classes. The exact structure and function of the different layers in convolutional neural networks is described in detail in [9]. In this work, pre-trained CNN’s were trained by transfer learning methods to categorise the input images. Transfer learning is a common technique adapting pre-trained neural networks to new tasks. It demands significantly less time and effort than building up and training new networks from scratch [10]. There are many examples, which use transfer learning of pre-trained networks for new classification applications, mostly in the medical field, like tumor classification [11], diabetic macular edema [12] or thoraco-abdominal lymph node (LN) detection [13]. For automated corrosion detection CNNs have been trained to detect corrosion on structures Convolutional Neural Networks techniques [14]. The results are very promising and already outperform other state of the art vision-based detection techniques. Another interesting application is the automated crack detection on large structures using CNNs [15, 16]. The studies suggest that CNNs could also be viable for the classification of Rockwell indentations, if a sufficiently large training dataset is available. The ImageNet database is often used for training and evaluating CNNs [17]. It currently contains about 14 million hierarchically structured and labelled images. By training with these images, the networks are already able to recognize many image properties, so that the adaptation to new areas of application, here Rockwell indentations, requires less computing effort.

2 Material and Methods

The Rockwell indentation tests were carried out with a commercial Rockwell hardness testing device (Testor Model HT1). The normal load is 1471 N, with a holding time of 4 s, using a conical Rockwell-diamond with a tip radius of 0.2 mm and an aperture of 120°. A large variety of different coating systems was tested for the training dataset. The majority of the coatings were deposited using reactive magnetron sputtering [18-21]. Marginal conditions are a polished, or at least nearly smooth (Sa = 2 nm to 2 µm), substrate surface and coatings with a thickness of 0.5 to 5 µm. The deep learning tests were carried out using the deep learning and parallel computing toolboxes of Matlab (The MathWorks Inc.), as well as pre-trained network models provided in the Matlab Add-On Center [22]. AlexNet, GoogLeNet and inception-v3 were selected as pre-trained deep convolutional neural networks. These networks, which differ mainly in their complexity, are commonly used for image classification applications. They were also selected because they can be trained in reasonable time with the available hardware and are therefore suitable for a first evaluation. AlexNet, presented 2012 by Krizhevsky et al. [8], uses input images with a size of 227x227 pixels and rgb colour range. It has a depth of 8 layers and is trained to categorise images into 1000 classes, using the LSVRC-2010 ImageNet training set with, at that point of time, a total of 1.3 million images as training data. A evolution of AlexNet is represented by GoogLeNet, which was presented 2015 by Szegedy et al. [23]. The architecture with a depth of 22 layers is considerably more complex compared to AlexNet. The image input size is 224x224 pixels. As a third and most advanced network, the inception-v3 network was used. It was developed by Szegedy et al. in 2016 [24]. The depth is 48 layers and the input images have to be 299x299 pixels in size. All network models in this work are frequently used for image classification tasks and are suitable for transfer learning, as they are already able to extract and evaluate properties of images. AlexNet and GoogLeNet only use max-pooling layers for data reduction, while inception-v3 additionally uses mean-pooling layers. The last fully connected layer and the output layer of the networks were modified to six output neurons for the six adhesion classes HF 1 - HF 6.

A dataset of 1652 classified microscopy images of Rockwell indentations was created for training and testing the different networks. A light microscope (Metallux 3, Ernst Leitz Wetzlar GmbH) and a laser-confocal microscope (Keyence VK-X1000) with magnifications from 50 to 100 and different exposure options have been used to create a diverse dataset. 12 images per adhesion class were randomly selected from the total dataset and separated for testing the different CNNs. The remaining 1580 indentation images were used as training lesson for the networks. Of each adhesion class at least 200 images were present, so that a biasing of the networks towards an overrepresented class is prevented. Nonetheless, adhesion class HF 3 contained more images than the other classes. A reason for this is that pictures in this category show more variance in terms of size and shape of the delaminations among each other. The training images were augmented with a random translation in X and Y direction, up to 5 pixels. It has to be noted, that the training of the networks is influenced by the human classification of the training data. In order to minimise this source of error, the training data was cross-checked and adjusted by four other scientific staff members. To ensure that different lighting, colour and contrast ratios of the impressions have no effect on the subsequent classification, they are included in the training data for each adhesion class to a sufficient extent. For verification purposes a separate dataset of 36 images with adhesion classes HF 1 to HF 6 was prepared. Figure 3 shows six typical images, one from each adhesion class, of the verification dataset. Since the nets can evaluate independently of the absolute imprint size, images without scale bars were used.

The different illumination, colour and enlargement of the individual indentations can be seen. For hardened steel with a hardness of about 58 – 62 HRC, the imprint diameters are in the range of 520 to 580 µm.

The solver for transfer learning was stochastic gradient descent with momentum and a decreasing learning rate. The maximum number of epochs was set to 50. As hardware conventional desktop components was used: an 8 core CPU, 16GB RAM and an Nvidia GTX1070 GPU.

3 Results

3.1 Training of the networks

Figure 4 shows training accuracy and loss of the three networks during training over 50 epochs. The training accuracy is defined as the percentage of correct predictions for each training mini batch. To compute the training loss, a cross entropy function is used in Matlab. The training of the networks took 10 minutes with AlexNet, 15 minutes with GoogLeNet and about 1 hour with the inception-v3 network model. All models reached a mini-batch accuracy of nearly 100 %. The mini-batch size for AlexNet and GoogLeNet was 128 images and 32 images for inception-v3, due to reaching the memory limit of the GPU. It can be concluded from the training graphs that inception-v3 converges fastest to a correct classification.

3.2 Testing

The trained networks were tested with the above mentioned 72 testing images, which are collectively shown in figure 5.

Figure 6 a - c shows the confusion charts for each trained network. With this representation method, the counts of predictions is plotted over the predicted (x) and true (y) classes. Correct classifications therefore lie on the diagonal of the resulting matrix. The necessary computing time to classify the 72 test images was less than 10 seconds for all networks. The number of correct predictions of AlexNet was 63, 62 for GoogLeNet and 61 for inception-v3. AlexNet and inception v3 showed no deviations in critical adhesion classes HF 5 and HF 6, while GoogLeNet misclassified one indentation as HF 4. The maximum deviation between prediction and true class for GoogLeNet and inception v3 was one level, while AlexNet classified a single HF 3 indentation as HF 1. Based on these results, inception-v3 appears to be the best neural network for the classification of Rockwell indentations, despite having the lowest number of correct predictions.

3.3 Network Verification

The accuracy of the networks was further investigated by evaluating a verification dataset containing 36 images of 6 different coating systems, wherein each represents an adhesion class. Figure 3 exemplary shows one imprint from every coating system, respectively adhesion class. Basic information about the six coating systems are provided in table 1.

These coating systems were deposited on hardened (60±2 HRC), polished (Sa < 10 nm) steel samples. Figure 7 shows the average rating of the six coating systems made by human visual evaluation, AlexNet, GoogLeNet and inception-v3. The mean adhesion class of each coating was calculated by averaging the individual indentation evaluations.

The high adhesive strength HF 1 (coating system 1) was correctly recognized by all networks with small standard derivations. For the coating system 2, the adhesion class was slightly underestimated by GoogLeNet, overestimated by inception-v3, while AlexNet matched the human classification. It must be noted that this coating system represents a borderline case between HF 1 and HF 2 and is therefore difficult to classify, even for trained technical staff, compare figure 3 b). Coating system 3 was correctly classified by GoogLeNet and inception-v3, but slightly underestimated by AlexNet. However, the standard deviations still overlap. Coating system 4 was correctly classified by all networks. The coating systems 5 and 6 with the technically insufficient adhesion classes HF 5 and HF 6 were correctly categorised with one exception. AlexNet classified the coating system 5 as HF 4, which is critical, because HF 4 indicates a technically usable adhesion strength.

4 Conclusions

In this work three different deep convolutional neuronal networks were trained with a dataset of about 1650 Rockwell imprints to perform an automated adhesion evaluation. The focus of the work was to evaluate the possibility of using deep neural networks as an alternative, or addition, to existing automated image analysis approaches. After training, GoogLeNet and especially inception-v3 showed accurate results for the automated classification of a test batch of coating systems with different adhesion classes. AlexNet also showed promising results, but still exhibited some critical classification errors. With a relatively small data set, it was possible to train the networks by transfer learning in order to perform the classification reliably. The good performance of the inception-v3 network could be partially related to the increased input image size of 299x299 pixels, compared to 224x224 (GoogLeNet) and 227x227 (AlexNet) pixels. The complexity of the networks can also be significant for the classification accuracy. This would explain the better performance of GoogLeNet compared to AlexNet. By extending the training data set, it will be possible to evaluate further surface and layer types in the future. The necessary computing time for the ratings is less than ten seconds for 100 images with the inception-v3 net, using conventional desktop PC-hardware. Even at this early stage the evaluation is already qualified for practical use. Another positive aspect is that the VDI guideline 3198 does not have to be changed or adapted. A stand-alone tool with a graphical user interface for indentation evaluation was created, which is ready for practical use. Extending the training dataset is planned to increase the accuracy. Further, more complex network models will be tested in the future. For example, the network model NASNet-Large with an input image size of 331x331 pixels, which is also available within the Matlab CNN toolbox [22]. The combination of the results of this work with adhesion testing machines is very promising for academic and industrial applications. For this purpose it is planned to use the Python compatibility of Matlab to transfer the shown concepts into real application tasks. Since the CNN networks generate probabilities for the individual adhesion classes for individual imprints, a finer classification of the adhesion class could also be calculated. It is planned to double check this approach with conventional image analysis of the delaminated area around the indentations. In addition, the methods shown are to be applied to other examination methods, for example the scratch test. By comparing different adhesion testing methods, the results of this work will also be reviewed towards their applicability for real applications.

5 Acknowledgements

We gratefully thank the Deutsche Forschungsgemeinschaft DFG for funding the project ME 1954/3-1 "Tribologische Transfermechanismen und großflächige Mikrokontaktsimulation der Festschmierstoffbereitstellung aus PVD-Schichten für trockenlaufende Zahnradstufen“ within the priority program 2074 „Fluidfreie Schmiersysteme mit hoher mechanischer Belastung“. We also gratefully acknowledge the framework program “horizon 2020” of the European Union for funding the project “CROwned Spline Surface Optimization using New Treatments”. Both projects share the aim of developing new coating systems and surface treatments for dry running tribological machine elements. Both projects require objective and reliable determination of the coating adhesion to the substrates.

List of References

1. Robertson, J., Diamond-like amorphous carbon. Materials Science and Engineering: R: Reports, 2002. 37(4): p. 129-281.

2. Hetzner, H., Systematic Development of Amorphous Carbon Coatings under Consideration of the Requirements of Sheet-Bulk Metal Forming. 2014.

3. Mittal, K.L., Adhesion measurement of films and coatings. Vol. 640. 1995: VSP.

4. DIN, E., 6508-1, Metallische Werkstoffe-Härteprüfung nach Rockwell. Teil1: Prüfverfahren (Skalen A, B, C, D, E, F, G, H, K, N, T)(ISO 6508-1: 2005), 2006.

5. Normen, V.D.I., VDI 3198. VDI-Verlag, Dusseldorf, 1991: p. 1-8.

6. Zawischa, M., et al., Scratch resistance of superhard carbon coatings – A new approach to failure and adhesion evaluation. Surface and Coatings Technology, 2016. 308: p. 341-348.

7. LeCun, Y., Y. Bengio, and G. Hinton, Deep learning. nature, 2015. 521(7553): p. 436.

8. Krizhevsky, A., I. Sutskever, and G.E. Hinton. Imagenet classification with deep convolutional neural networks. in Advances in neural information processing systems. 2012.

9. Goodfellow, I., Y. Bengio, and A. Courville, Deep learning. 2016: MIT press.

10. Haykin, S., Neural Networks and Learning Machines, 3/E. 2010: Pearson Education India.

11. Huynh, B.Q., H. Li, and M.L. Giger, Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. Journal of Medical Imaging, 2016. 3(3): p. 034501.

12. Karri, S.P.K., D. Chakraborty, and J. Chatterjee, Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. Biomedical optics express, 2017. 8(2): p. 579-592.

13. Shin, H.-C., et al., Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE transactions on medical imaging, 2016. 35(5): p. 1285-1298.

14. Atha, D.J. and M.R. Jahanshahi, Evaluation of deep learning approaches based on convolutional neural networks for corrosion detection. Structural Health Monitoring, 2018. 17(5): p. 1110-1128.

15. Zhang, L., et al. Road crack detection using deep convolutional neural network. in 2016 IEEE international conference on image processing (ICIP). 2016. IEEE.

16. Cha, Y.J., W. Choi, and O. Büyüköztürk, Deep learning- based crack damage detection using convolutional neural networks. Computer- Aided Civil and Infrastructure Engineering, 2017. 32(5): p. 361-378.

17. Deng, J., et al. Imagenet: A large-scale hierarchical image database. in 2009 IEEE conference on computer vision and pattern recognition. 2009. Ieee.

18. Decho, H., et al., Influence of Chromium, Zirconium and Silicon on the Wear Resistance of Titanium Aluminum Nitride (TiAlN) Multilayer Films. HTM Journal of Heat Treatment and Materials, 2018. 73(1): p. 41-53.

19. Kovac, J., et al., Tensile properties of magnetron sputtered aluminum–scandium and aluminum–zirconium thin films: A comparative study. Surface and Coatings Technology, 2013. 215: p. 369-375.

20. Böhmermann, F., et al., Dry rotary swaging–approaches for lubricant free process design. International Journal of Precision Engineering and Manufacturing-Green Technology, 2015. 2(4): p. 325-331.

21. Decho, H., et al., Optimization of chromium nitride (CrNx) interlayers for hydrogenated amorphous carbon (a-C:H) film systems with respect to the corrosion protection properties by high power impulse magnetron sputtering (HiPIMS). Surface and Coatings Technology, 2016. 293: p. 35-41.

22. The MathWorks, I. Pretrained Deep Neural Networks. 2019; Available from: https://de.mathworks.com/help/deeplearning/ug/pretrained-convolutional-neural-networks.html.

23. Szegedy, C., et al. Going deeper with convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

24. Szegedy, C., et al. Rethinking the inception architecture for computer vision. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

Tables

Table 1: Properties of the verification coating systems

Adhesion Class	HF 1	HF 2	HF 3	HF 4	HF 5	HF 6
Coating System	single-layer TiN	multi-layer a-C:H(:W)				
Indentation Hardness HIT [GPa]	25.7	8.0	12.9	18.1	8.0	8.7
Thickness [µm]	1.4	2.1	2.0	2.3	4.3	3.5

List of Figure Captions

Figures

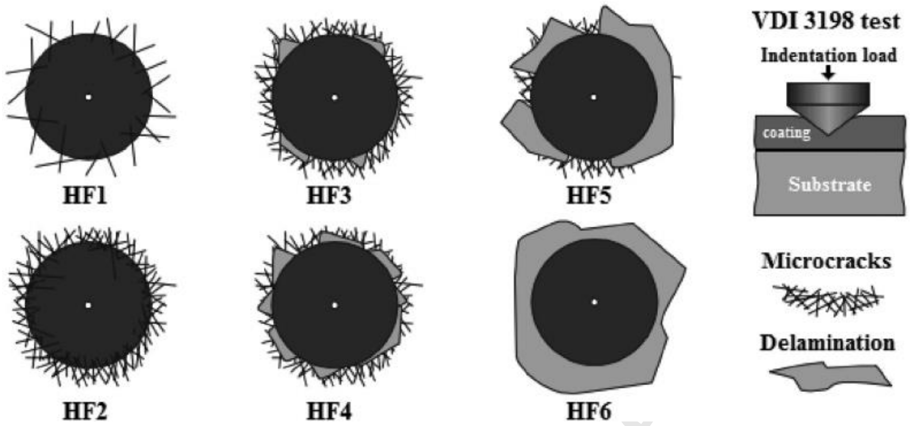


Figure 1

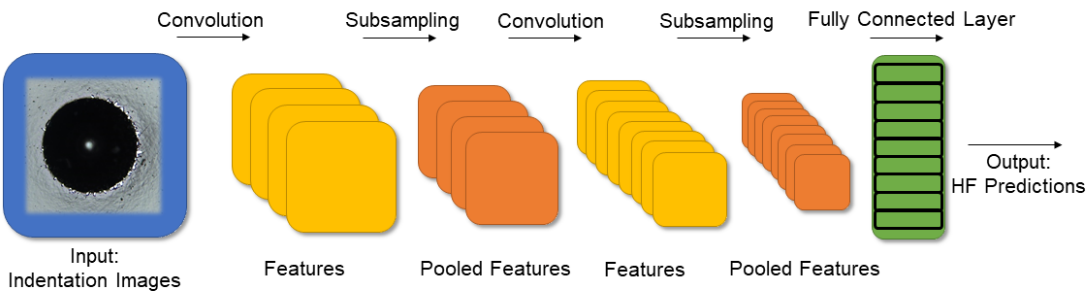


Figure 2

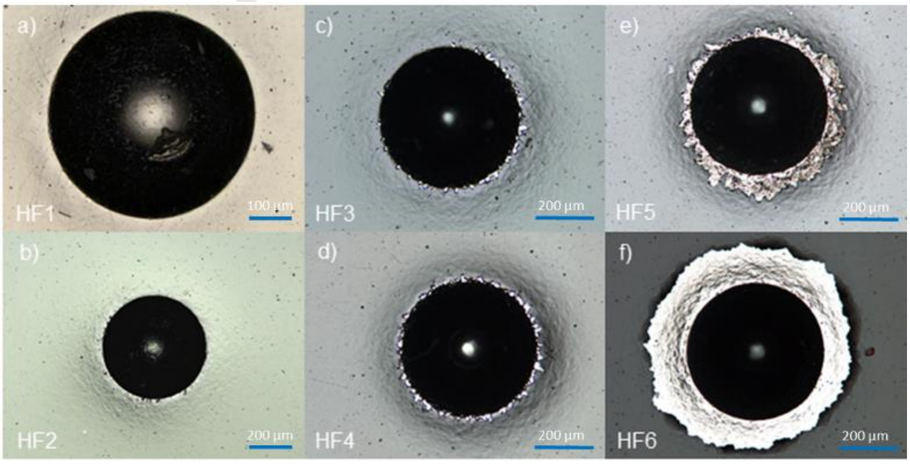


Figure 3

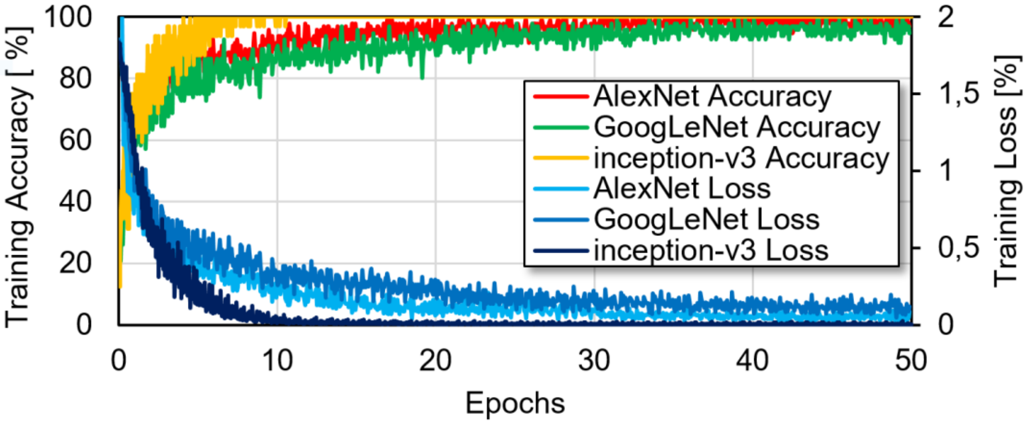


Figure 4

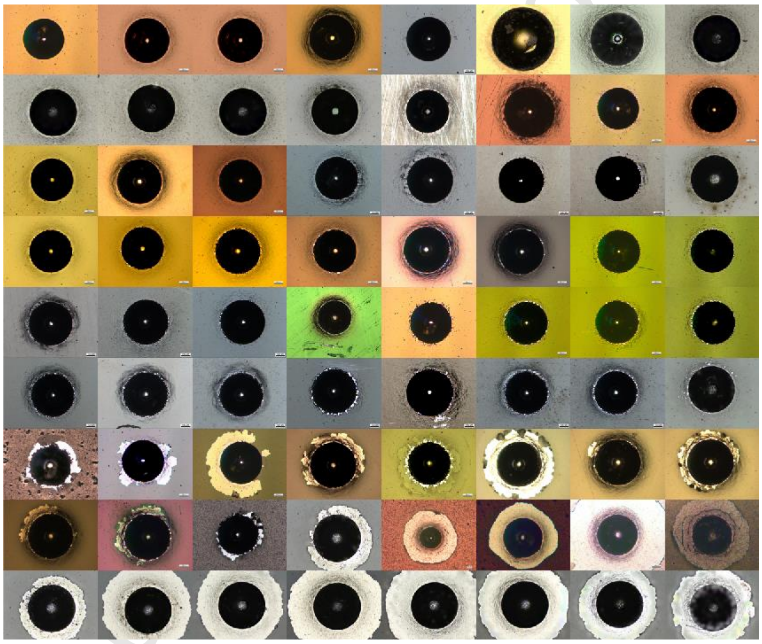


Figure 5

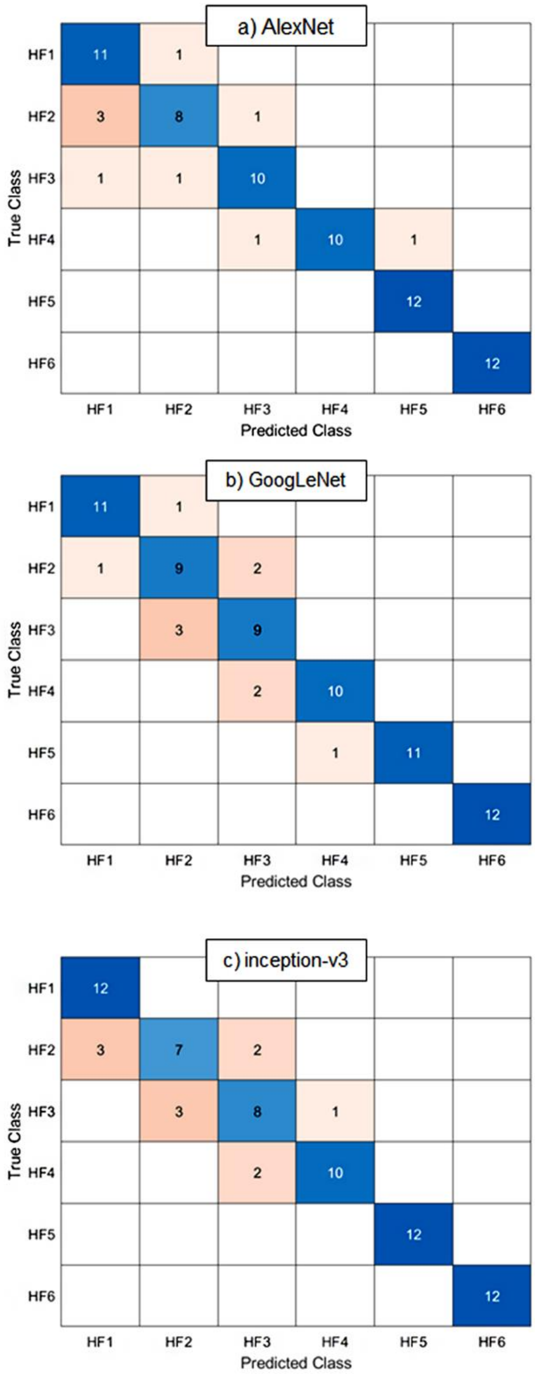

a) AlexNet
b) GoogLeNet
c) inception-v3
True Class
Predicted Class
HF1
HF2
HF3
HF4
HF5
HF6

Figure 6

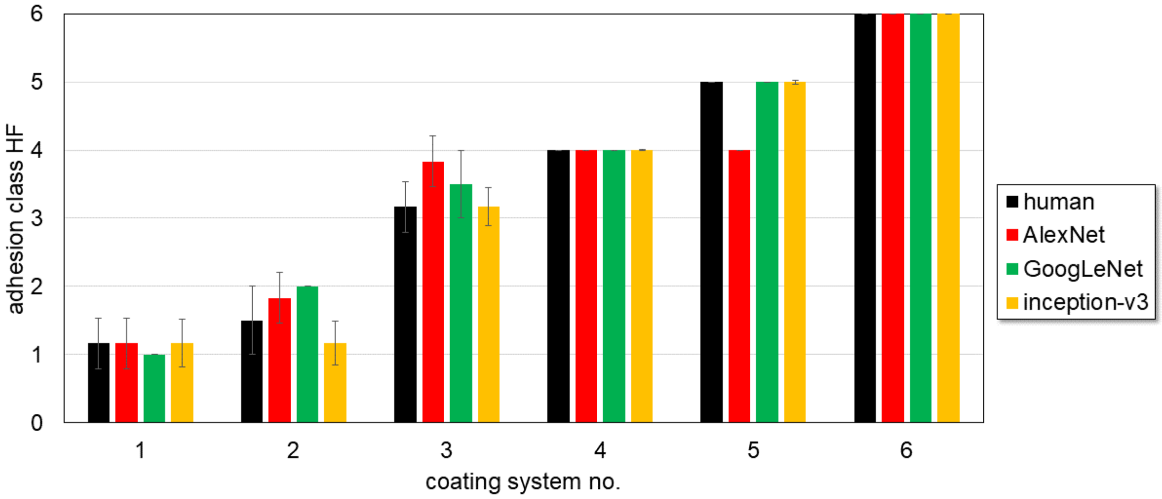

6
5
4
3
2
1
0
adhesion class HF
1
2
3
4
5
6
coating system no.
human
AlexNet
GoogLeNet
inception-v3

Figure 7

Credit Author Statement

1. M. Sc. Bastian Lenz:

Conceptualization, Methodology, Software, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Visualization

2. Dipl.-Ing. Henning Hasselbruch:

Validation, Writing - Review & Editing

3. Dr.-Ing. Andreas Mehner:

Resources, Writing - Review & Editing, Supervision, Project administration, Funding acquisition

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Highlights

- Convolutional neural Networks (CNN) are trained to classify Rockwell indentations into the adhesion groups HF 1 to HF 6 according to VDI 3198.
- The classifications of the implemented models exhibit an agreement of approximately 85 - 90 % compared to human assessment.
- The networks show promising results for automated industrial applications, such as in-line adhesion control in coating processes, as they do not require human operator support.