DOI: 10.2195/lj_Proc_zhu_en_202112_01
URN: urn:nbn:de:0009-14-54509

Automated, AI-based Inspection of Drive Wheels on Overhead Hoist Transport Vehicles

Automatische, KI-basierte Inspektion von Antriebsrädern an Overhead Hoist Transportfahrzeugen

Hailong Zhu, Sebastian Rank, Thorsten Schmidt

Professur für Technische Logistik
Technische Universität Dresden
Dresden, Germany

Overhead hoist Transport systems are used to transport wafers in 300 mm semiconductor factories. These rail-based systems usually consist of hundreds of vehicles to ensure fast and safe transport of wafers between tools. Faults of individual vehicles can cause damage to the transferred goods and production downtimes. To minimize the risk of failure, extensive preventive maintenance of the vehicle's heavily stressed components is required. This includes the chassis and drive wheels. This article describes an automatic inspection approach that can drastically accelerate the inspection process of drive wheels. From the data obtained, we trained a deep convolutional autoencoder network to predict the growth of fractures on the surface of the wheels, which in the end allows us to carry out condition-based predictive maintenance of the vehicles. This approach promises cost savings compared to routine- or time-based strategies for preventive maintenance, as we can carry out maintenance tasks only when they are justified.

[Keywords: faults detection, wear out model, OHT, AMHS, condition monitoring, autoencoder]

Overhead Hoist Transportsysteme werden zum Transport von Wafern in 300-mm-Halbleiterfabriken eingesetzt. Diese schienenbasierten Systeme bestehen in der Regel aus Hunderten von Fahrzeugen, um einen schnellen und sicheren Transport der Wafer zwischen den Bearbeitungsmaschinen zu gewährleisten. Fehler einzelner Fahrzeuge können zu Schäden am Transportgut und/oder gar Produktionsausfällen führen. Um das Ausfallrisiko zu minimieren, ist bis dato eine umfassende vorbeugende Wartung stark beanspruchter Fahrzeugkomponenten erforderlich. In diesem Zusammenhang wurde ein automatischer Inspektionsansatz für die Antriebsräder der Fahrzeuge entwickelt, womit der Inspektionsprozess deutlich beschleunigt werden kann. Aus den Inspektions-/Messdaten konnte ein deep convolutional-Autoencoder-Netzwerk trainiert werden, welches das Wachstum von Frakturen auf der Oberfläche der Räder vorhersagt. Der Inspektionsansatz erlaubt eine zustandsbasierte prädiktive Wartung der Fahrzeuge. Damit sind Kosteneinsparungen gegenüber routinemäßigen oder zeitbasierten Strategien zur präventiven Wartung zu realisieren, da Wartungsaufgaben nur bei tatsächlicher Notwendigkeit vorgenommen werden.

[Schlüsselwörter: Fehlererkennung, Verschleißmodell, OHT, AHMS, Zustandsüberwachung, autoencoder]

1 INTRODUCTION

Overhead hoist Transport (OHT) systems are used to transport wafers in 300 mm semiconductor factories (fabs). These rail-based systems usually consist of hundreds of vehicles to ensure fast and safe transport between the single manufacturing tools [AHS06; HSS16].

The components of the OHT vehicle are heavily stressed because of the 24/7 operation of semiconductor factories. Over time, they wear out and can cause damage to the vehicle itself, its rail, and transported goods and hence contribute to loss of production. In order to minimize the probability of system failure, specialists manually carry out extensive preventive maintenance of the vehicles on fixed time period basis (usually annually, [SZZ18]). This process includes removing the vehicle from the rail, manual inspection of vehicle components and replacement of the potentially worn-out parts. Easy to understand, this approach is quite costly in terms of e.g. personnel efforts, spare parts procurement, and vehicle downtimes. Another challenge is that the quality of inspection and thus the appropriate assessment of the vehicle’s components heavily depend on the expertise and competences of the respective maintenance staff. For example, there is no clear standard describing the grade of wear for drive wheels, which eventually can lead to different decisions of the single technicians on the issue of replacement. In order to automate and standardize the inspection tasks and thus to ensure a high maintenance quality level, we propose the application of an automated inspection station.

DOI: 10.2195/lj_Proc_zhu_en_202112_01
URN: urn:nbn:de:0009-14-54509

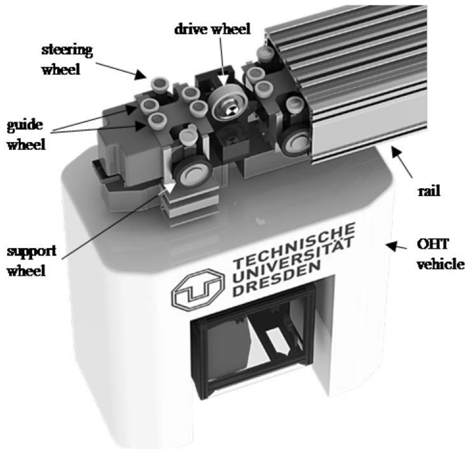
The wear and tear process of the drive wheel happens gradually over time. With data collected from multiple measurements on multiple vehicles, we trained a deep convolutional autoencoder to predict how the fractures on the driving wheel develops in time.

This paper is structured as followed: Section 2 introduces the objective and describes an (automatic) inspection station. In section 3 we present a autoenocder model to predict the development of fractures on drive wheel surfaces. Finally, Section 4 concludes and gives an outlook.

2 APPROACH

2.1 SUBJECTS OF INSPECTION

Depending on the manufacture, OHT vehicles have different wheels. For our experiments, we applied a Murata SCR 350 vehicle equipped with drive wheels, guide wheels, steering wheels, and support wheels, as shown in Figure 1. A test environment was brought up and the corresponding data for the evaluation was collected in a 300 mm fab of a partner.



The long-term goal of our research is to inspect all the mentioned components automatically. For this paper, we will focus on measurements and analysis of the driving wheel although the subsequently presented approach is supposed to be applicable for the other mentioned wheels (and OHT-manufactures), too.

During operation, the radius of the drive wheel decreases and the surface fissures. A worn-out drive wheel can result in nonsufficient traction and inaccurate positioning of the vehicle in its rails. A worn-out wheel also releases more particles than wheel in proper condition, which is quite problematic in clean room areas. To prevent such issues, the condition of the drive wheel should be inspected and evaluated thoroughly.

2.2 THE INSPECTION STATION

In order to capture the condition of the drive wheel, we developed an inspection station. We used the term “station”, because the measurements are done not directly on the “productive” OHT path but on a rail next to it. The actual measuring unit is assembled on a part of straight rail. An opening allows optic sensors to inspect the drive wheel. To do so, we mounted a laser profile sensor on the rail to obtain the 3D model of a drive wheel’s surface. Figure 2 shows how the laser sensor scans the drive wheel. The scanning procedure can be described as followed: While the roller device is locked, the vehicle enters the maintenance station and stops with the drive wheel directly underneath the sensor’s laser axis. Afterwards the roller device is unlocked which allows the drive wheel to rotate and the laser sensor positioned above is able to capture the profile data of the wheel’s surface. By recording a series of profiles, from the laser profile sensor’s data a 3D scan of the wheel’s surface is obtainable.

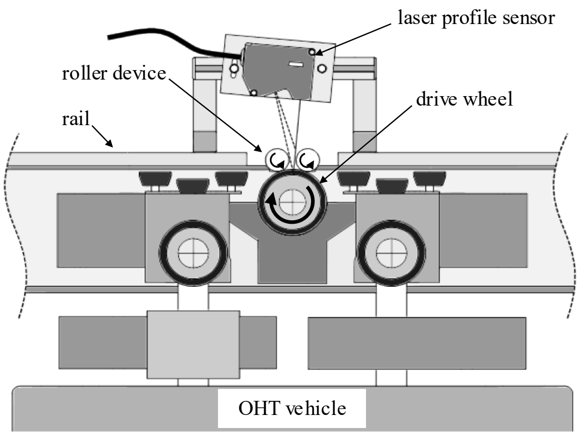
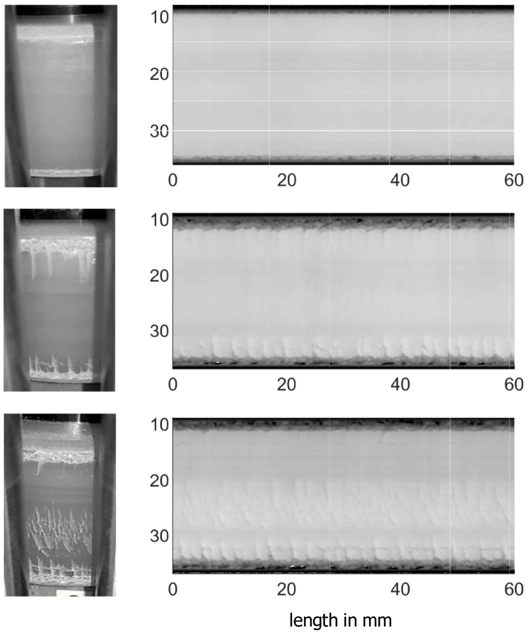


Figure 2. Section view of the inspection station and vehi

Figure 3 shows three examples of drive wheel’s surfaces with different wear-out grades. The scans were captured with the inspection station described before.

The top photo and image a) show a relatively new wheel with slight abrasions on both sides of the wheel. The photo and image b) show a wheel in the middle of its predicted life cycle; we can see micro fractures on the edge of the wheel. The last photo and image c) show the surface of a wheel right at the end of its life cycle. We can see fractures in the center of the wheel surface. The performance of the wheel will decline if not replaced.

DOI: 10.2195/lj_Proc_zhu_en_202112_01
URN: urn:nbn:de:0009-14-54509



3 ANALYSIS METHODS AND RESULTS

3.1 ABRASIVE MODELING OF RUBBER WHEELS

Not explicitly mentioned so far, this paper focuses on rubber wheels. In this context, the abrasive wear process of rubber were studied since decades, both experimentally and theoretically [GeP83; StD88; Sch58; IUS05; MuR92]. Experimental studies on the friction and wear of rubber wheels were also carried out [IUS05]. In various studies, the weights and diameter of the rubber probe is used to indicate the degree of wear [GeP83; StD88; Sch58; MuR92]. Even though this leads to a confident assessment, weight and diameter are difficult to measure if the drive wheel is supposed to remain mounted on the vehicle. However, the fractures vertical to the direction of travel on the wheel surface are also valid markers for the condition of the wheel surface [IUS05; MuR92]. So for our study we used these vertical fractures as an indicator of wear-out degree and try to predict their progress/growth.

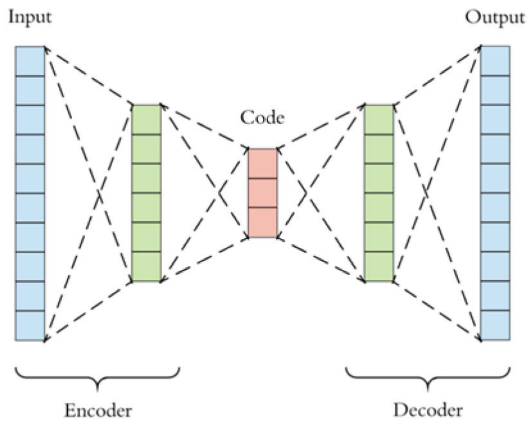
From our experiments we learned that the 3D model derived from the laser profile sensor's data has excellent quality in terms of resolution and accuracy; usually down into the sub micrometer range. Therefore, we are even able to observe subtle changes on the wheel's surface between two subsequent measurements. It allows us to identify single wheels in itself and even their singular fractures over multiple measurements

3.2 CONVOLUTIONAL AUTOENCODER MODEL FOR PREDICTION

In the past decades, machine learning algorithms and techniques have been of great interest in research and developed and widely used to perform data analysis on large datasets [LBD89]. For instance, the datasets commonly used to train image recognition algorithms contains millions of images. In case of supervised learning, annotated data is required, and the performance of the model depends heavily on the quality of the dataset used. However, annotated data is not always available due to the nature of the raw data and/or the costly/time consuming labeling process.

In our case, the growth of fractures is not intuitive to human eyes recognition, thus makes manual annotation impossible. In order to exploit the unlabeled surface scan/3D models described in section II.B, we propose using a convolutional deep autoencoder model to firstly extract the characteristics of the fractures on wheel's surfaces and secondly make a prediction on fractures' development. We pursuit this approach, because this hybrid model has been proven to be effective on feature extractions on unintuitive image datasets like CT scans [CSZ17] and radar-based signals [SÖG18].

An autoencoder [SMS15] is a neural network that learns to describe the input with a low dimension representation and reconstruct the input from that representation. As shown in Figure 4, an autoencoder is constructed by two main parts: an encoder that maps the input into a lower dimensional representation, known as "code", and a decoder that maps the code to a reconstruction of the original input.



However, the autoencoder can also be used to generate or reconstruct images [SÖG18], and for motion prediction [WDG16]. In our case, aiming to predict the progress of fractures on wheels, their 3D scans serve as the input of the autoencoder, and the difference between subsequent measurements as the output of the autoencoder. In detail, the scanned 3D surface model of a wheel is comparable to a

DOI: 10.2195/lj_Proc_zhu_en_202112_01
URN: urn:nbn:de:0009-14-54509

grayscale image, difference being that each pixel is represented by a 32-bit real number instead of an 8-bit integer. Because of this similarity, we decided to use convolutional layers for both, the encoder and the decoder, which is proven to be more effective than applying "simple" ANN (artificial neuro-network) for image processing [ZeF14]. The convolutional layer is the core building block of a convolutional neuro network [CNN]. Each of the convolutional layer consist of a set of learnable filters—each has a small receptive field. During the learning process, each of these filters is convolved across width and height of the input image, computing the dot product between the filter and a part of the input, generating a 2-demensional activation map of that filter. By calculating the overall cost function, we can determine the performance of the respective set of filters in terms of detecting the desired type of feature. By backpropagation, the parameters of the filters are updated accordingly. This process will be repeated until the performance of the CNN reaches expectations/defined thresholds.

Figure 5 shows the structure of our convolutional autoencoder network. There are three convolutional layers in the encoder and decoder part, respectively. The encoder tries to extract the features of the input 3D model, which are important for prediction, whereas the decoder tries to generate a feasible prediction based on these features.

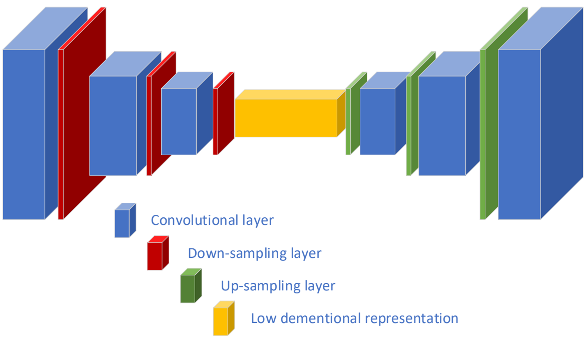


Figure 5. Structure of proposed convolutional autoencoder

3.3 DATA-PREPARATION AND TRAINING OF THE CONVOLUTIONAL AUTOENCODER

To obtain a sufficient amount of data to test/validate our model, we captured the state of driving wheels of 14 OHT vehicles for 10 times over a time span of 9 months each. Data were collected during live operation from one of our partners' 300 mm-wafer fab. According to well-experienced technicians in charge, in real world the lifespan of a driving wheel is usually about 12 to 18 months.

For our test design, in cooperation with the mentioned technician, we selected vehicles with different wear-out conditions to make sure our datasets represent the entire lifespan of driving wheels. Apparat from that, the corresponding OHT vehicles were selected randomly and did not get any special treatment/attention at all. We also performed the data gathering processes with extra precaution to avoid errors, outliers and inconsistencies. Eventually, the dataset used to train the autoencoder consist of 24 000 data pairs (characteristics of a data pair: see next paragraph), extracted from 10 measurements of 12 wheels. The verification set consist of 4000 data pairs from 10 measurements of the remaining 2 wheels.

For the data preparation process, we first performed an overall pattern matching on the sample data, so that every fracture can be tracked through multiple measurements over time. Each data pair consists of a 3D model of a potentially fractured area form a measurement as the input of the autoencoder and a 3D model of the difference between the original and the sequential measurement of the same area. We used the difference between two measurements instead of a second measurement as our observations and findings revealed that the development of fractures is not significant between two sequential measurements. In other words, it became obvious that if we use the second measurement as the output of the autoencoder, our autoencoder will be more focused on reconstructing the surface model than predicting its changes (as intended). In this context and exemplarily, Figure 6 shows two sequential measurements in a) and b) of a fracture and their grey scale difference in c). To make the difference more visible, we have magnified it by a factor of 5.

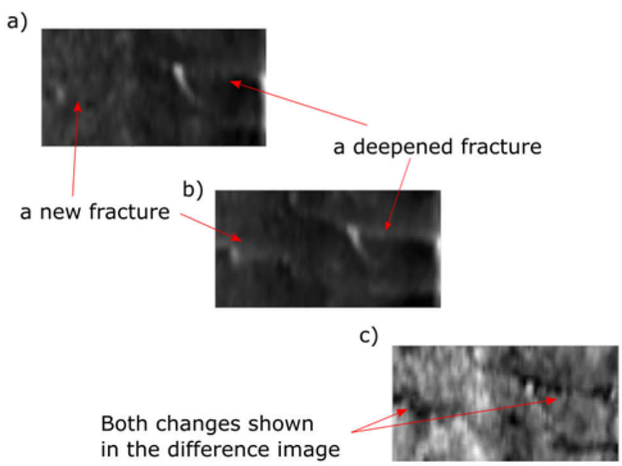


Figure 6. Training data: a) surface of a fracture from measurement 1; b) surface of the same fracture from a subsequent measurement 2 (one month later); c) difference between the two measurements

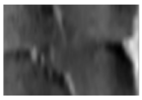
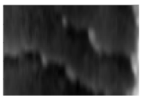
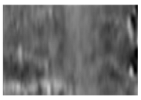
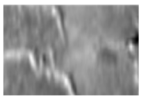
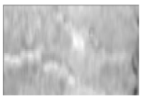
For the obligatory training process of the hybrid model, we used the tensorflow and keras library as they are the most commonly used model-training platforms. We applied batch normalization to accelerate the training [IoS15] and "adadelta" [Zei12] as the optimizer. We trained the autoencoder for 800 epochs, where the cost function flattened after approximately 650 epochs. We selected these hyperparameters based on results of multiple experiments' results [BBK15].

3.4 USE CASE: RESULTS AND EVALUATION

Following result assessment is performed on a qualitative basis because, unfortunately, a universally defined standard on evaluation of surface fractures is missing.

Table 1 shows three representative examples from the verification data set consisting of 4000 data pairs, including a first measurement as input, the difference when measured a second time a month later, and the predicted difference calculated by our convolutional autoencoder each.

Table 1. Examples from the verification dataset: Difference calculated on the basis of two consecutive measurements and difference predicted by autoencoder.

	example 1	example 2	example 3
input			
difference observed from measurements			
difference predicted by autoencoder			

In Table 1, in the input line, the white area is where the fracture is. In the lines showing the difference observed from measurements and their predicted counterparts, white means further/additional wear out.

Figure 7 gives a more detailed impression, how the characteristics (in our case the development of fractures) are successfully predicted by our autoencoder model. We can see that the convolutional autoencoder can predict the area in which the fracture is more likely to develop. Nevertheless, we must also admit that our prediction lacks details. There are meanly two reasons for this: First, so far we strongly assume that the development of fractures is not a deterministic process, whereas the autoencoder can only predict the possibility of further fracture development of an area. The second reason is that during the coding process some information on the details (inherently) go lost. This is why autoencoders are used as noise filters—as described in chapter 3.2 and [SMS15]. One way to improve the performance on details of the prediction might be to train a generative adversarial network (GAN, see [GPM14]). In theory, a GAN network can generate a possible prediction with more details, but may need more training data and computing resources. In this case, we could use the convolutional autoencoder as the so-called commentator (supportive model) in the training process of a GAN network. We will determine the feasibility of the GAN network in a feature research.

With the forecast generated by the autoencoder, we can predict the growth of fractures on driving wheels—more precisely, the fractures expected one month in advance. So, if the prediction of a wheel's surface meets the criteria for replacement, maintenance of the vehicle should be scheduled in the short term. In this way, we can now plan the maintenance work based on the condition of the driving wheel, which might be far more efficient than preventive actions.

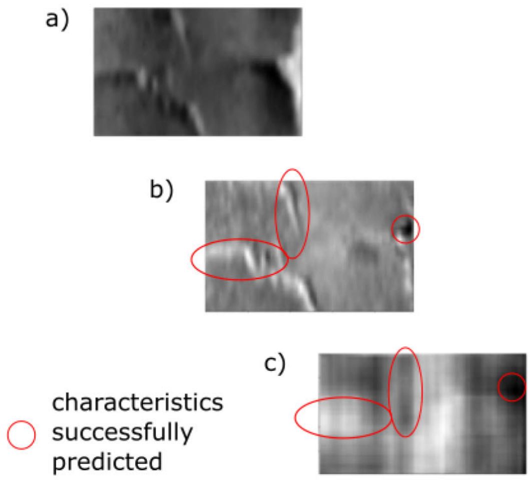


Figure 7. A detailed presentation of an example: a)input; b)difference; c)prediction of difference

4 CONCLUSION

In this article, we introduced and applied an automatic inspection station for OHT vehicles, especially for their drive wheels. We deployed the inspection station in a 300 mm-semiconductor-fab and conducted measurements of multiple vehicles in live operation. From the obtained data and its 3D models of the wheels, we trained a deep convolutional autoencoder network in order to forecast the growth of fractures on the wheels' surface. So far, the results are quite promising, as we, firstly, proved that the conditions of the wheel surface can be captured by the automatic inspection station and its installed laser profile sensors. Secondly, an autoencoder model can be used to predict the development of fractures—nevertheless, we pursue to gather more scans to further quantitatively and statistically validate the result of the prediction model. From practitioner point of view, with the help of the inspection station and the prediction model, it is possible to establish condition-based preventative maintenance of OHT vehicles, which is more efficient than preventive or reactive maintenance.

DOI: 10.2195/lj_Proc_zhu_en_202112_01
URN: urn:nbn:de:0009-14-54509

LITERATURE

[AHS06] Agrawal, Gaurav K.; Heragu, Sunderesh S.: *A survey of automated material handling systems in 300-mm SemiconductorFabs.* In: IEEE transactions on semiconductor manufacturing vol. 19 / 2006, number 1, pp. 112–120.

[BBK13] Bardenet, Rémi; Brendel, Mátyás; Kégl, Balázs [et al.]: *Collaborative hyperparameter tuning.* International conference on machine learning. PMLR: 2013, pp. 199–207.

[CSZ17] Chen, M.; Shi, X.; Zhang, Y. [et al.]: *Deep Features Learning for Medical Image Analysis with Convolutional Autoencoder Neural Network.* In: IEEE Transactions on Big Data vol. / 2017, pp. 1–1.

[GeP83] Gent, AN; Pulford, CTR: *Mechanisms of rubber abrasion.* In: Journal of Applied Polymer Science vol. 28 / 1983, number 3, pp. 943–960.

[GPM14] Goodfellow, Ian J.; Pouget-Abadie, Jean; Mirza, Mehdi [et al.]: *Generative adversarial networks.* In: arXiv preprint arXiv:1406.2661 vol. / 2014.

[HSS16] Hammel, Christian; Schmaler, Robert; Schmidt, Thorsten [et al.]: *Empowering existing automated material handling systems to rising requirements: FA: Factory automation.* 2016 27th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC). IEEE: 2016, pp. 87–93.

[IoS15] Ioffe, Sergey; Szegedy, Christian: *Batch normalization: Accelerating deep network training by reducing internal covariate shift.* International conference on machine learning. PMLR: 2015, pp. 448–456.

[IUS05] Iwai, T.; Uchiyama, Y.; Shimosaka, K. [et al.]: *Study on the formation of periodic ridges on the rubber surface by friction and wear monitoring.* In: Wear vol. 259 / 2005, number 1–6, pp. 669–675.

[LBD89] LeCun, Yann; Boser, Bernhard; Denker, John S. [et al.]: *Backpropagation applied to handwritten zip code recognition.* In: Neural computation vol. 1 / 1989, number 4, pp. 541–551.

[MuR92] Muhr, AH; Roberts, AD: *Rubber abrasion and wear.* In: Wear vol. 158 / 1992, number 1–2, pp. 213–228.

[Sch58] Schallamach, A.: F*riction and abrasion of rubber.* In: Wear vol. 1 / 1958, number 5, pp. 384–417.

[SMS15] Sedhain, Suvash; Menon, Aditya Krishna; Sanner, Scott [et al.]: *Autorec: Autoencoders meet collaborative filtering.* Proceedings of the 24th international conference on World Wide Web. 2015, pp. 111–112.

[SÖG18] Seyfioğlu, M.S.; Özbayoğlu, A.M.; Gürbüz, S.Z.: *Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities.* In: IEEE Transactions on Aerospace and Electronic Systems vol. 54 / 2018, number 4, pp. 1709–1723.

[SZZ18] Siegel, Armin; Zhakov, Artem; Zhu, Hailong [et al.]: *Trailer based rail monitoring in overhead hoist transport systems.* 2018 International Symposium on Semiconductor Manufacturing (ISSM). IEEE: 2018, pp. 1–4.

[StD88] Stupak, PR; Donovan, JA: *Fractal analysis of rubber wear surfaces and debris.* In: Journal of materials science vol. 23 / 1988, number 6, pp. 2230–2242.

[WDG16] Walker, Jacob; Doersch, Carl; Gupta, Abhinav [et al.]: *An uncertain future: Forecasting from static images using variational autoencoders.* European Conference on Computer Vision. Springer: 2016, pp. 835–851

[Zei12] Zeiler, Matthew D.: *Adadelta: an adaptive learning rate method.* In: arXiv preprint arXiv:1212.5701 vol. / 2012.

[ZeF14] Zeiler, Matthew D.; Fergus, Rob: Visualizing and understanding convolutional networks. European conference on computer vision. Springer: 2014, pp. 818–833.

DOI: 10.2195/lj_Proc_zhu_en_202112_01
URN: urn:nbn:de:0009-14-54509

Dipl.-Ing. Hailong Zhu, Research Assistant at Chair of Material Handling, Dresden University of Technology. Hailong Zhu was born 1988 in Hebei, China. Between 2013 and 2018 he studied electronic engineering at the Dresden University of Technology.

Address: Professur für Technische Logistik, Technische Universität Dresden, Münchner Pl. 3, 01062 Dresden, Germany,
Phone: +49 351 463 32532, Fax: + 49 351 463 35499,
E-Mail: hailong.zhu@tu-dresden.de

Dr.-Ing. Sebastian Rank, postdoctoral researcher at the Chair of Logistics Engineering, Institute of Material handling and Industrial Engineering, Technische Universität Dresden.

Sebastian Rank received a M.S. degree in Economics and Engineering and a Ph.D. degree in Engineering at Technische Universität Dresden. His research interests include autocorrelation in logistics systems and AMHS simulation in semiconductor areas. He is a member of ASIM.

Prof Dr.-Ing. Thorsten Schmidt, Head of the Chair of Logistics Engineering, Institute of Material handling and Industrial Engineering, Technische Universität Dresden.

Thorsten Schmidt is full professor at the TU Dresden and heads the Chair of Material Handling in the Mechanical Engineering faculty since 2008. He holds a diploma degree in mechanical engineering from the TU Dortmund and a Master degree in industrial engineering from the Georgia Institute of Technology. He received his Ph.D. from the TU Dortmund in 2001. His research areas are the design and optimization of facility logistics and production systems including a focus on the machinery and components involved. He currently works on energy efficient control strategies in material flow, fast approximation in early planning stages by means of standard design modules, online data analysis, formal verification of control logic, performance analysis of decentralized and selfcontrolled systems, lightweight structures in material handling and stress analysis on wire ropes and toothed belts, respectively.