https://doi.org/10.2352/EI.2023.35.16.AVM-125
© 2023, Society for Imaging Science and Technology

MTF as a performance indicator for AI algorithms?

Patrick Müller[1], Alexander Braun[1]

[1] **University of Applied Sciences Düsseldorf, Düsseldorf, Germany**

Abstract

The modulation-transfer function (MTF) is a fundamental optical metric to measure the optical quality of an imaging system. In the automotive industry it is used to qualify camera systems for driver assistance systems or for future autonomous driving (ADAS/AD). Each modern ADAS/AD system includes evaluation algorithms for environment perception and decision making that are based on artificial intelligence or machine learning (AI/ML) methods and neural networks. The performance of these AI algorithms is measured by established metrics like Average Precision (AP) or precision-recall-curves. In this article we research the robustness of the link between the optical quality metric and the AI performance metric. A series of numerical experiments were performed with an object detection algorithm (cars, pedestrians) evaluated on image databases with varying optical quality. We demonstrate with these that for strong optical aberrations a distinct performance loss is apparent, but that for subtle optical quality differences – as might arise during production tolerances – this link does not exhibit a satisfactory correlation. This calls into question how reliable the current industry practice is where a produced camera is tested end-of-line (EOL) with the MTF, and fixed MTF thresholds are used to qualify the performance of the device-under-test.

Introduction

The MTF is an established metric in the field of optics to quantify the ‘sharpness’of an optical system [1, 2], based on linear system theory. There are many different approaches to measuring the MTF using different types of target test charts (Siemens star, slanted edge, dead leaves or spilled coins, illuminated lines, illuminated points) [3]. To be more precise, many of those methods measure the *spatial frequency response* (SFR), which is basically the same as the MTF but derived from non-harmonic target features (e.g. slanted edge instead of sinusoidally modulated Siemens star). Often MTF and SFR are used synonymous, we use MTF throughout this article.

It is (automotive) industry standard to test a new camera both during production (lens-sensor alignment) as well as end-of-line (EOL) against fixed MTF thresholds [4]. If a MTF value of the device-under-test (DUT) is above the threshold the camera is deemed ready to be delivered to the customer, not causing any failures in the field, if the MTF value is below the threshold the camera will be reworked or – as in the case of the automotive industry – destroyed, throwing away maximum added value. It is worth stressing the relevance of this threshold, as *this is where the money is*! Fig. 1 visualizes this correlation: the developer of the system defines a function limit that the camera needs to achieve under production tolerances (dashed blue line), using MTF as the quality metric. This could be e.g. a certain MTF50 value, or a minimum value at a given spatial frequency. A camera (DUT) is produced and compared to this threshold: if the MTF value is above the threshold, the camera is sold to the customer, with the implicit expectation from automotive quality processes that this system will not cause any accidents or harm in the field. If the MTF value of the DUT is below the threshold it is scrapped, wasting maximum value-add. Reworking the DUT is usually not economically feasible in the automotive industry. Therefore the crossing of function limits and production tolerances (red circle) symbolizes the economic pressure of the whole project: the production plant manager needs the production tolerances to be as wide as possible, avoiding over-engineering and costs, whereas the developer is conscious of the possibly devastating effect a systematic error in the field would cause. Therefore, solidly determining the relation between the test metric (here: MTF) and the actual performance of the DUT in the field is key to economic success.

Nonetheless, exactly this link is not established in a scientifically thorough manner during the development of the camera system. The key challenge for this process is the actual function limits of algorithms based on Artificial Intelligence (AI), machine learning (ML) and neural networks – for brevity subsumed as *AI algorithms* in the following. Due to the superior performance of AI algorithms in the field of Computer Vision (CV) no modern camera-based ADAS/AD system comes without AI algorithms

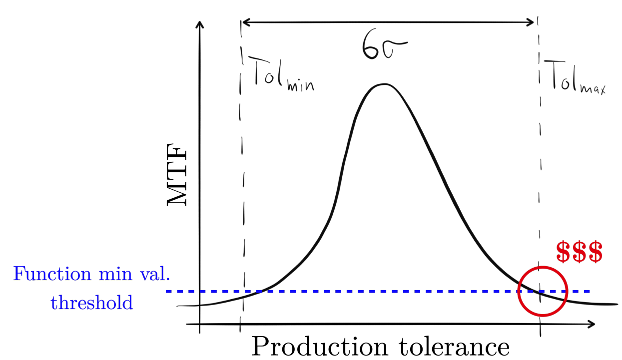


Figure 1: MTF function limits vs. production tolerances. If the limits are too wide, there is a danger of unsafe cameras causing accidents in the field. If the limits are too tight the camera is over-engineered, and good cameras are binned, destroying maximum value-add. In that sense the crossing of the production tolerance limit and the function limit (red circle) symbolizes the economic success of the whole camera project.

like Convolutional or Deep Neural Networks (CNN, DNN). But these algorithms are black-boxes in that the output cannot be predicted in a deterministic manner. Rather, it is estimated with a best-practice statistical approach, and with millions of kilometers of test drives on proving grounds and public roads, and in simulation. This situation is described in detail in [4]. Unfortunately, as the function limits of AI algorithms cannot be determined in a mathematical, provable way, the industry runs the danger of a *McNamara fallacy* [5]: reaching a certain level of MTF becomes the actual goal of the development, whereas the application performance of the DUT in the field – a good traffic sign recognition, a stable lane detection – takes a second place.

Nonetheless, linking optical quality to AI performance is still a niche academic research area, in that usually optics designer are no AI/ML experts and vice versa. Still, there are a number of activities. First, the fact that blurring significantly degrades object detection performance has been established with the ImageNet-C image data set, where the *C* stands for corruption [6]. Here Gaussian-like blur was used to degrade sharpness, amongst many other degradation, and the performance declines distinctly. We have extended this work with our own research [7, 8].

Saad et al. [9] applied a measured vignetting intensity profile to the KITTI image dataset and observed a spatially variant reduction in car detection performance. Here, the spatial performance distribution was distinctly correlated to the intensity of the vignetting, where the performance degraded in darker corners of the image. Pezzementi et al. applied a longer list of image degradations (intensity and color noise, distortion, blur) to a data set from the agriculture industry, where a camera was mounted on heavy machinery to detect people working in the fields [10]. Several different detection algorithms where used, and the robustness of each algorithm to the different degradations was quantified using AP and precision-recall curves.

All these works demonstrate the influence of image quality on the detection performance of AI algorithms. We believe with this contribution we address a missing element in all those works as we *quantify* the amount of optical degradation using the MTF on *simulated* test charts. Thus, we can now correlate the amount of degradation in the performance to the quantified optical performance of the imaging system. We show that for large differences in MTF values an AI algorithm performs worse as expected. But for subtle differences as might be found for small production tolerances the AI performance is robust against these changes, calling into question the current practice of equating MTF values end-of-line with the AI performance, demonstrating a possibly dangerous McNamara fallacy.

This paper is structured as follows. After a brief theoretical introduction to optical and AI performance metrics we present our method how to link the optical quality to the performance of AI algorithms. Here, we use object detection for the classes car and pedestrian to demonstrate our results based on experiments on the Berkeley Deep Drive data set [11].

Measuring optical quality and AI performance

In order to link the optical quality to the performance of the AI algorithm both properties have to be quantified by metrics. This section describes both metrics we used. First the optical quality of the lens model we employed is described, as well as a brief review of the lens model itself. Second, the selection of AI algorithms are described together with the choice of metrics.

Optical model and degradation algorithms

The optical model has been used and described in detail before [12, 7, 8]. It comprises of a lens model and a degradation algorithm to apply the lens effects to an image, similar in effect to the *Image Simulation* tool in Zemax [13]. The lens model is based on Zernike polynomials describing the optical wavefront aberration in the pupil plane for different field positions. The main idea is to simulate wavefront aberrations in the pupil plane, and to derive different point-spread functions (PSFs) over field from these aberrations by Fourier transform [1]. The PSFs are then used to degrade the image in image space. Additionally, the Zernike coefficients can be used as model parameters to vary the optical quality, e.g. for a defocus study.

In this work, though, the lens model itself is not parameterized. Instead, the degradation algorithm is varied. Using convolution as a degradation algorithm – an established routine in both academy and industry – is correct only when the PSF is constant, as linear system theory requires both linearity and translation-invariance for a convolution. As the PSF is spatially variant over the field of view (FoV) of the lens the formally correct degradation algorithm is the superposition (SP) approach [1], where each pixel is assigned a unique degradation kernel. In this case this kernel accords to the PSF at that position. We have developed our own degradation algorithm implementations both for convolution in patches and for the superposition approach.

The SP approach is optically the only correct one, but it requires a different PSF for every single pixel. There is a desire to switch to less memory and runtime intensive algorithms to process more images with the simulation. Often for such tasks the so-called isoplanar patches algorithm is used [14], based on the overlap-add criterion [15] and used e.g. in [16]:

$$g(x,y) = \sum_{k=1}^{K} \psi_k(x,y) \cdot (f * h_k)(x,y), \tag{1}$$

where $\psi(x,y)$ denotes the interpolation mask, f the input image (object), h_k the k-th kernel related to the mask and g the final image.

The algorithm reduces the computation of the superposition to the application of a simple convolution per block, which can be accelerated. The algorithm can be used if the PSF changes slowly and is therefore approximately constant in regions of size $K \times L$[16]. To avoid blocking artifacts the patches overlap and are finally interpolated (usually bilinearly) [16, 14]. The parameter block size thus decides how fine the gradations between the PSFs are. More details on the degradation model can be found in [7].

We now compare the SP model with a convolution in patches, where the patch-size is varied. All in all, four different degradation variations are used: SP and patch sizes 80 px, 160 px and 320 px. The differences between all four variations are subtle, and we use these as a proxy to substitute actual production tolerances.

Optical quality

The used lens is a Cooke triplet with distinct chromatic aberrations and astigmatism. These aberrations are stronger than typically encountered in current automotive camera systems, but are useful for the scientific purpose of developing a novel process to link optical quality to AI performance, as it clearly discerns the influences of these aberrations. We apply the lens model directly on top of a data set of already recorded, real images using the four different degradation algorithms discussed above. Thus, the images actually include two sets of lenses: the lens with which the images were recorded, and our Cooke triplet. The recording camera was an iPhone 5, and has a high optical quality. The main influence on the optics will be the Cooke triplet. Finally, we do not consider object space depth, implicitly assuming that all relevant objects are trans-hyperfocal, which for this data set and the selected detection algorithms is an acceptable assumption. Also, the auto-focus capabilites of the recording iPhone are ignored, which influences the defocus values as well. Again, for a real automotive project all these properties have to be rendered correctly, but for the purpose of this article these approximations are good enough to demonstrate the validity of the process.

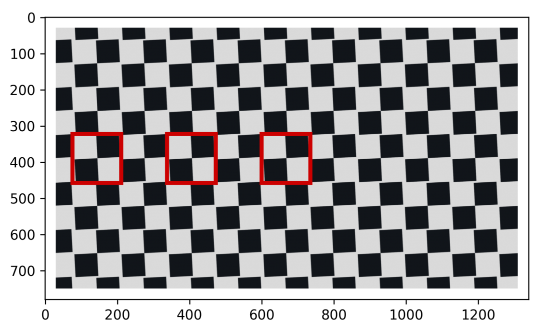


Figure 2: The slanted edge target used to evaluate the MTF performance of the optical model over field. The MTF is evaluated in three field positions indicated by the red squares: center, 0.5 (middle) and 0.7 (edge). The image is rasterized in the required resolution from a vector graphics, and the optical model is used to blur the resulting bitmap using different degradation algorithms. The used target has a contrast ratio of 10:1.

In this article we measure the quality of our optical model with the MTF only. There are many other optical quality metrics that could be used, like distortion and geometrical calibration, signal-to-noise ratio (SNR) or dynamic range (DR)[17]. All these will also have an influence on the application (AI) performance. The process described in this article can be applied to all those other metrics in a straight-forward manner as well, but we focus on the MTF to segregate and isolate the effects.

We determine the MTF with the publicly available and established *sfrmat4*-algorithm by Peter Burns [18], which implements the ISO12233 algorithm using a slanted edge target [19]. Our slanted edge target is shown in Fig. 2 and has resolution 1280×720 pixels. For the purpose of this work three regions of interest (ROI) were selected and marked in red squares in the figure: in the center, at a mid-field and at the edge of the horizontal field of view. The target has a contrast of 10:1 and the evaluation ROI is 45 by 40 pixels to fit the corresponding edge orientation. The MTF is evaluated and plotted for both horizontal and vertical orientation, taking the mean value over two values each for upper and lower (left and right), i.e. four evaluations for each ROI.

AI algorithms and performance metrics

In this article, an object detector is used to localize and recognize persons and cars on an automotive data set. A state-of-the-art neural network for object detection is used, which has been trained on the MS COCO dataset [20] and is examined here on the Berkeley Deep Drive dataset[11] without fine-tuning. The analysis is limited to the validation set of the BDD100k detection dataset. The 10000 images all have the same size and were extracted from videos taken by an iPhone5 camera in US cities. On the dataset used, a total of 13425 people of different sizes are marked. The people are also approximately equally distributed in the horizontal direction. This facilitates later analysis. The labels were converted to COCO annotation format for use in the python COCO API for evaluation [20].

The Cascade Mask R-CNN algorithm was used for object detection only, since instance segmentation labels are not available for the BDD10k data set. The Deep Neural Network is present in the Detectron2 model zoo [21] with pre-trained weights and uses a Resnet-152 backbone [22] pre-trained on ImageNet [23]. The algorithms are not re-trained on the degraded data set for this study, a process that is included in ongoing work.

We use *Intersection-over-Union* (IoU), *mean average precision* (mAP) and *Precision-Recall-Curves* to quantify the performance of the detection network[24].

Linking optical quality and AI performance

To investigate the influence of the interpolation density on the object detection, a three-step procedure is chosen: First, a PSF data set is generated with the optical model, producing 920k PSFs per color channel. These PSFs are input to the various degradation algorithms and applied to the unmodified image data set, resulting in multiple image data sets that now contain the effects of different PSF sampling density. Finally, the inference for object recognition is performed for all image data sets modified in this way. The different results per step are compared to each other [7].

As mentioned before, three different model approximations of the superposition approach are discussed. The 1280×720 imager is divided into isoplanar square blocks of size 80^2,160^2 and 320^2 as summarized in Tab. 1. The number of used color PSFs range from 48 to 432, about a factor of 10, while the superposition uses a total of 2764800 PSFs. The total number of used PSFs increases if the PSF grid needs to be filled with edge PSFs for a given block size [7].

Metric	SP	ISO_{320}	ISO_{160}	ISO_{80}
Block size	–	320^2	160^2	80^2
# PSFs (luminance)	921600	16	48	144
# PSFs (color)	2764800	48	144	432

Table 1: Chosen algorithms for Cooke-Triplet optical model on imager of size 1280×720 and three color channels [7]

Summarizing, the BDD100k validation data set is degraded using the single optical model with the four different algorithms, resulting in five data sets: baseline, SP, ISO_{320}, ISO_{160} and ISO_{80}. In addition to degrading the image data set, the slanted edge target introduced before is degraded with these four algorithms as well, allowing for a quantification of the optical quality of the degradation method. This additional quantification of the optical quality is one central novelty of our approach, first presented in [7].

MTF as performance indicator?

Now it is possible to link the optical quality to the AI performance. We can compare four different sets of degraded image quality with four sets of AI performance. The detection algorithm performs on these five data sets, yielding the AI algorithm performance as AP and Precision-Recall-Curves. The baseline is not included in the MTF visualizations, as it just yields the sinc-function as the Fourier transform of the rectangular pixel of the given size.

Optical quality results

We show three graphs depicting the MTF curves for all four degradation algorithms, for the center of the image, for mid-field and for the edge of the field. The optical quality degrades with increasing field as expected from the lens setup, and the astigmatism becomes clearly visible (difference between dashed and solid line). Recall that the MTF is principally evaluated in horizontal and vertical direction, which on the horizontal axis of the field of view accords to tangential and sagittal, respectively [13, 25].

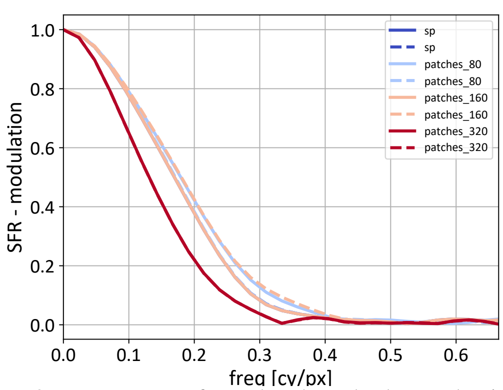


Figure 3: MTF curves from the slanted edge evaluation at the center of the field. Dashed curves depict vertical and solid lines horizontal evaluation.

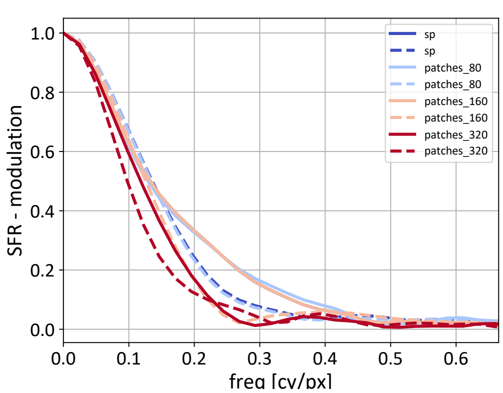


Figure 4: MTF curves from the slanted edge evaluation at the middle of the field. Dashed curves depict vertical and solid lines horizontal evaluation.

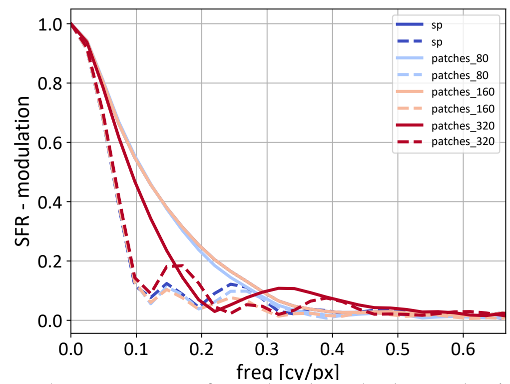
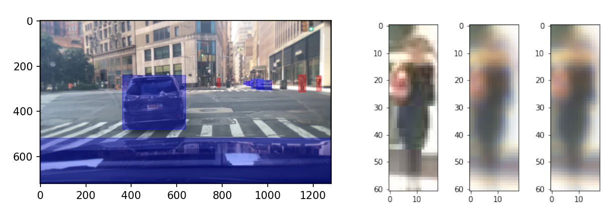


Figure 5: MTF curves from the slanted edge evaluation at the edge of the field. Dashed curves depict vertical and solid lines horizontal evaluation.

Ideally, all degradation curves would be identical, as the optical lens model is the same, and only the numerical application is different. As a trend, though, the coarse simulation for ISO_{320} is always lower in quality in comparison to the other simulations. Interestingly, both ISO_{80} and even ISO_{160} seem to be an adequate approximation of the superposition algorithm, at least in terms of the optical quality, as the MTF curves for those three algorithms are in very close proximity. In the center this is especially clear, as all curves are on top of each other except the one for ISO_{320}. Mid-field the ISO_{320} again differs distinctly, but interestingly the vertical (dashed) curve for ISO_{160} also differs, a fact that is not repeated at the edge of the field.

In a production environment single numerical values would be derived from these curves. These would indicate that the ISO_{320} 'lens' has a lower performance, with a MTF50 value of 0.131 cy/px in comparison to 0.174 cy/px for the others, or an MTF value of 22% @ 0.2 cy/px in comparison to 40% for the others. For any automotive quality processes these would clearly be very different lenses, and the values are indeed so different that it is easy to image the ISO_{320} 'lens' being a 'fail', and the others all a 'pass'. The interesting questions now is: how do these 'lenses' perform on an actual task, like person detection?

AI performance results



(a) Example detection. (b) (c) (d)

Figure 6: Example detection of a person with a small size a) The whole image, b) – d) crop of a single person at the right edge, with b) original, c) SP and d) ISO_{320}. The original instance image b) has a confidence score = 93.4%, c) SP: 90.2%, d) ISO_{320}: 91.3% [7].

An example detection of a person shown in Fig. 6 visualizes the task and the effect different degradation algorithms have on the image quality. The whole image is shown as well as three crops of a single person, for the original (baseline) version, as well as for the SP and the ISO_{320} algorithm. The effect of the degradation is distinct, the person looks much blurrier. Nonetheless, there is hardly a visual difference between the two different degradation algorithms, even though the MTF at the edge of the field indicates differently. In accordance with the visual impression, the confidence scores of the person detection are almost the same for SP and ISO_{320} at 90.2% and 91.3%, respectively, a slight drop from 93.4% for the original.

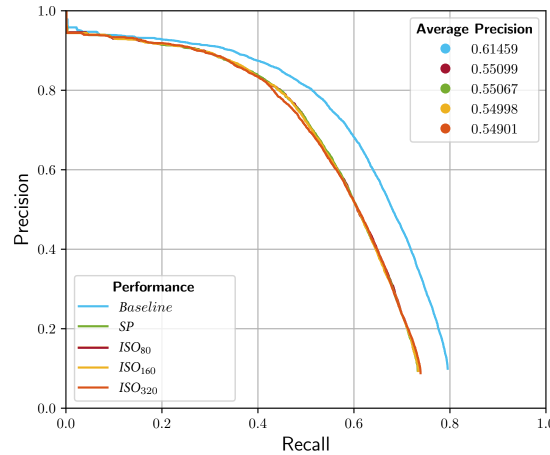
Consequently, we observe in the precision-recall evaluation in Fig. 7 a clear drop in AI performance going from the original to the degraded images for both person (a) and car (b) detection. But the detection algorithm demonstrates the same performance for all four different degradation algorithms, as all four precision-recall-curves are on top of each other. Also, the average precision (AP, inset legend) confirms this result with a drop of approx. 6.5% from original to degraded for the person detection in (a), and again a constant detection performance at about 55.0% for all degradation algorithms. The graph for car detection in (b) yields similar results, though the drop in AP from original to degraded is now only 2%. While this could indicate a robustness of car detection in comparison to person detection, we assume this to be object size dependent, which is also currently under investigation.

Summary and outlook

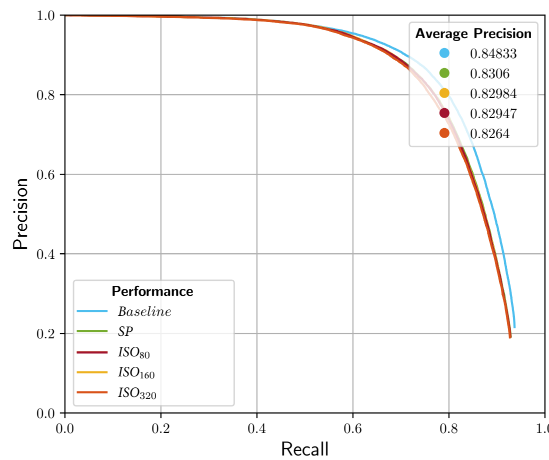
We presented a numerical simulation method to link optical quality of an imaging system to the performance of an AI algorithm working on those images. A baseline data set – here the Berkeley Deep Drive (BDD) – was degraded in four different ways, using a self-developed physical-realistic optical lens model. On this group of five different data sets (original plus four degradations) the same AI algorithm was run, a person and car detection based on Cascade Mask R-CNN from the Detectron2 model zoo. We use the degradation algorithms not only on the BDD data set, but also on an slanted edge target that allows for a quantitative evaluation of the MTF performance. Now the optical quality – in terms of the MTF – can be linked to the AI performance – given here by Average Precision and Precision-Recall-Curves.

The results show a good correlation between the unmodified and the degraded BDD data sets: the degradation leads to a distinct reduction in AI performance for person and car detection. On the other hand, while the optical performance of the four degradations exhibits clear differences, the AI performance is constant for all four. Because the MTF values differ so much, and yet the AI performance is constant, this questions the established practice of the automotive industry of simply selecting a fixed MTF threshold during end-of-line testing. Clearly, every algorithm has its own requirements that need to be determined individually. Our novel process of linking optical quality to AI performance can help with this determination using physical-realistic simulations.

There were several constraints in this study. The AI algorithm was trained on COCO, not on the BDD. Further, no retraining was performed that took the degraded image data sets into account, to gauge how much of the performance loss could be recovered. This leads to the larger question how low certain optical metrics can go while keeping – or recovering – the AI performance with clever training on blurred images. Finally, we showed the result for one select AI algorithm and two classes only. Apparently, it would be of interest to study other detection algorithms



(a)



(b)

Figure 7: Precision vs. Recall and Average Precision (AP) for Cascade Mask R-CNN pedestrian and car detection, evaluated for all degradation algorithms. Baseline refers to the unmodified database.

and classes of algorithms, and to test which group of algorithm is sensitive to which class of optical degradations. This is left for future work.

References

[1] Joseph W. Goodman. *Introduction to Fourier optics*. Fourth edition. New York: W.H. Freeman, Macmillan Learning, 2017. 546 pp. ISBN: 978-1-319-11916-4.

[2] Glenn D. Boreman. *Modulation Transfer Function in Optical and Electro-Optical Systems*. SPIE, July 1, 2001. ISBN: 978-0-8194-8045-3. DOI: 10.1117/3.419857. URL: https://spiedigitallibrary.org/ebooks/TT/Modulation-Transfer-Function-in-Optical-and-Electro-Optical-Systems/eISBN-9780819480453/10.1117/3.419857 (visited on 09/25/2020).

[3] Uwe Artmann. “Image quality assessment using the dead leaves target: experience with the latest approach and further investigations”. In: IS&T/SPIE Electronic Imaging. Ed. by Nitin Sampat, Radka Tezaur, and Dietmar Wüller. San Francisco, California, United States, Feb. 27, 2015, 94040J. DOI: 10.1117/12.2079609. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2079609 (visited on 09/25/2020).

[4] Alexander Braun. “Automotive mass production of camera systems: Linking image quality to AI performance”. In: *tm - Technisches Messen* (June 15, 2022). Publisher: Oldenbourg Wissenschaftsverlag. ISSN: 2196-7113. DOI: 10.1515/teme-2022-0029. URL: https://www.degruyter.com/document/doi/10.1515/teme-2022-0029/html?casa_token=gZVA2FRN_sMAAAAA%3AYNYJxEi-WhG9DEFWSLxGbcvTo8eeDzqikr8gXfBzwV_m_T9ZSEgrfDSkeOf7Us-237Os-g5o5Tqg (visited on 12/12/2022).

[5] Dietmar Jannach and Christine Bauer. “Escaping the McNamara Fallacy: Towards more Impactful Recommender Systems Research”. In: *AI Magazine* 41.4 (Dec. 28, 2020), pp. 79–95. ISSN: 2371-9621, 0738-4602. DOI: 10.1609/aimag.v41i4.5312. URL: https://ojs.aaai.org/index.php/aimagazine/article/view/5312 (visited on 02/03/2023).

[6] Dan Hendrycks and Thomas Dietterich. “Benchmarking Neural Network Robustness to Common Corruptions and Perturbations”. In: *arXiv:1903.12261 [cs, stat]* (Mar. 28, 2019). arXiv: 1903.12261. URL: http://arxiv.org/abs/1903.12261 (visited on 03/08/2022).

[7] Patrick Müller and Alexander Braun. “Simulating optical properties to access novel metrological parameter ranges and the impact of different model approximations”. In: *2022 IEEE International Workshop on Metrology for Automotive (MetroAutomotive)*. 2022 IEEE International Workshop on Metrology for Automotive (MetroAutomotive). July 2022, pp. 133–138. DOI: 10.1109/MetroAutomotive54295.2022.9855079.

[8] Patrick Müller, Alexander Braun, and Margret Keuper. “Impact of realistic properties of the point spread function on classification tasks to reveal a possible distribution shift”. In: Workshop on Distribution Shifts, 36th Conference on Neural Information Processing Systems (NeurIPS 2022). New Orleans, LA, Dec. 3, 2022, p. 13.

[9] Kmeid Saad and Stefan-Alexander Schneider. “Camera Vignetting Model and its Effects on Deep Neural Networks for Object Detection”. In: *2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE)*. 2019 IEEE International Conference on Connected Vehicles and Expo (ICCVE). Graz, Austria: IEEE, Nov. 2019, pp. 1–5. ISBN: 978-1-72810-142-2. DOI: 10.1109/ICCVE45908.2019.8965233. URL: https://ieeexplore.ieee.org/document/8965233/ (visited on 04/09/2021).

[10] Zachary Pezzementi et al. “Putting Image Manipulations in Context: Robustness Testing for Safe Perception”. In: *2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. 2018 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR). Philadelphia, PA: IEEE, Aug. 2018, pp. 1–8. ISBN: 978-1-5386-5572-6. DOI: 10.1109/SSRR.2018.8468619. URL: https://ieeexplore.ieee.org/document/8468619/ (visited on 04/09/2021).

[11] Fisher Yu et al. “BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning”. In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. ISSN: 2575-7075. June 2020, pp. 2633–2642. DOI: 10.1109/CVPR42600.2020.00271.

[12] Patrick Müller, Mattis Brummel, and Alexander Braun. “Spatial Recall Index for Machine Learning Algorithms”. In: *London Image Meeting*. Society for Imaging Science and Technology, 2021.

[13] Zemax. *OpticStudio — Optical, Illumination & Laser System Design Software - Zemax*. URL: https://www.zemax.com/products/opticstudio (visited on 04/15/2021).

[14] James G. Nagy and Dianne P. O’Leary. “Fast iterative image restoration with a spatially varying PSF”. In: Optical Science, Engineering and Instrumentation ’97. Ed. by Franklin T. Luk. San Diego, CA, United States, Oct. 24, 1997, p. 388. DOI: 10.1117/12.279513. URL: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.279513 (visited on 10/27/2020).

[15] Thomas G. Stockham. “High-speed convolution and correlation”. In: *Proceedings of the April 26-28, 1966, Spring joint computer conference on XX - AFIPS ’66 (Spring)*. the April 26-28, 1966, Spring joint computer conference. Boston, Massachusetts: ACM Press, 1966, p. 229. DOI: 10.1145/1464182.1464209. URL: http://portal.acm.org/citation.cfm?doid=1464182.1464209 (visited on 10/27/2020).

[16] Stefan Harmeling, Hirsch Michael, and Bernhard Schölkopf. “Space-Variant Single-Image Blind Deconvolution for Removing Camera Shake”. In: (2010), p. 9.

[17] *IEEE Standards Association P2020 — Automotive Image Quality Working Group*. URL: https://site.ieee.org/sagroups-2020/ (visited on 01/31/2022).

[18] Peter D. Burns and Don Williams. “Camera Resolution and Distortion: Advanced Edge Fitting”. In: *Electronic Imaging* 2018.12 (Jan. 28, 2018), pp. 171–1–171–5. ISSN: 2470-1173. DOI: 10.2352/ISSN.2470-1173.2018.12.IQSP-171. URL: http://www.ingentaconnect.com/content/10.2352/ISSN.2470-1173.2018.12.IQSP-171 (visited on 09/25/2020).

[19] *ISO12233:2017, Photography — Electronic still picture imaging — Resolution and spatial frequency responses*. Standard. Volume: 2017. Geneva, CH: International Organization for Standardization, 2017.

[20] Tsung-Yi Lin et al. "Microsoft COCO: Common Objects in Context". en. In: *Computer Vision – ECCV 2014*. Ed. by David Fleet et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2014, pp. 740–755. ISBN: 978-3-319-10602-1. DOI: 10.1007/978-3-319-10602-1_48.

[21] Yuxin Wu et al. *Detectron2*. 2019. URL: https://github.com/facebookresearch/detectron2.

[22] Kaiming He et al. "Deep Residual Learning for Image Recognition". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, June 2016, pp. 770–778. ISBN: 978-1-4673-8851-1. DOI: 10.1109/CVPR.2016.90. URL: http://ieeexplore.ieee.org/document/7780459/ (visited on 04/07/2022).

[23] Olga Russakovsky et al. "ImageNet Large Scale Visual Recognition Challenge". In: *International Journal of Computer Vision* 115.3 (Dec. 2015), pp. 211–252. ISSN: 0920-5691, 1573-1405. DOI: 10.1007/s11263-015-0816-y. URL: http://link.springer.com/10.1007/s11263-015-0816-y (visited on 04/20/2022).

[24] Mark Everingham et al. "The Pascal Visual Object Classes (VOC) Challenge". In: *International Journal of Computer Vision* 88.2 (June 2010), pp. 303–338. ISSN: 0920-5691, 1573-1405. DOI: 10.1007/s11263-009-0275-4. URL: http://link.springer.com/10.1007/s11263-009-0275-4 (visited on 06/10/2022).

[25] *Understanding tangential/sagittal in OpticStudio and how to rotate rays – Knowledgebase*. URL: https://support.zemax.com/hc/en-us/articles/1500005488921-Understanding-tangential-sagittal-in-OpticStudio-and-how-to-rotate-rays (visited on 02/03/2023).

Author Biography

Patrick Müller received his M.Sc. in 2018. His Master's thesis examined the influence of a Point Spread Function Model to Digital Image Processing algorithms. He is currently pursuing his Doctorate at Hochschule Düsseldorf and University of Siegen with a focus on the application of optical models to digital images, their assessment and correlation with the performance of Computer Vision algorithms.

Alexander Braun received his diploma in physics with a focus on laser fluorescent spectroscopy from the University of Göttingen in 2001. His PhD research in quantum optics was carried out at the University of Hamburg, resulting in a Doctorate from the University of Siegen in 2007. He started working as an optical designer for camera-based ADAS with the company Kostal, and became a Professor of Physics at the University of Applied Sciences in Düsseldorf in 2013, where he now researches optical metrology and optical models for simulation in the context of autonomous driving. He's member of DPG, SPIE and IS&T, participating in norming efforts at IEEE (P2020) and VDI (FA 8.13), and currently serves on the advisory board for the AutoSens and the Electronic Imaging conference.