Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 232 (2024) 616–625

Procedia Computer Science

www.elsevier.com/locate/procedia

5th International Conference on Industry 4.0 and Smart Manufacturing

# Impact of Blemish Artefacts on Object Detection Models in Autonomous Driving: A Study on Camera Module Manufacturing Defects

Nitin Augustine[a,b], Maximilian Schwab[b], Steffen Klarmann[a], Christian Pfefferer[a], Alexander Schiendorfer[b,*]

[a]*Valeo Schalter und Sensoren GmbH, Valeostraße 1, 86650 Wemding, Germany*
[b]*Technische Hochschule Ingolstadt, Institute AImotion Bavaria, Esplanade 10, 85049 Ingolstadt, Germany*

**Abstract**

With the rise of autonomous driving technology, concerns over the safety of AI-based systems have increased, particularly regarding their effect on human life and infrastructure. The complexity and unpredictable nature of AI-enabled software functions have disrupted traditional safety evaluation methods.In such a scenario, the quality control methods used at the manufacturing environments of sensors used for deploying AI-based perception functions need to be coupled with the downstream functions.This study focuses on investigating the impact of a specific production-related defect called blemish, which is caused by debris on image sensors of camera modules. The research aims to answer two fundamental questions: 1) How can blemish artifacts be mathematically modeled, and what is the process for developing a suitable dataset that incorporates this defect? 2) How can the influence of blemish artifacts on performance of object detection models be traced back to define quality requirements of the camera modules? This work takes a novel perspective by shifting the focus from improving model resilience to improving hardware quality camera modules by studying the effects of manufacturing defects on downstream functions.This research seeks to establish a framework to derive quality targets for front camera module manufacturing which can be extended to other sensors for similar applications.



*Keywords:* Perception function; Object detection; Camera module; Blemish ;

## 1. Introduction: Determining quality criteria for electronic components in autonomous driving

The beginning of the last decade witnessed an exponential use of software and electronics-enabled functionalities within the automotive industry. The increasing concern for safety and the need for convenient transportation systems

* Alexander Schiendorfer. Tel.: +49 841 9348-3531
*E-mail address:* alexander.schiendorfer@thi.de

1877-0509 
10.1016/j.procs.2024.01.061

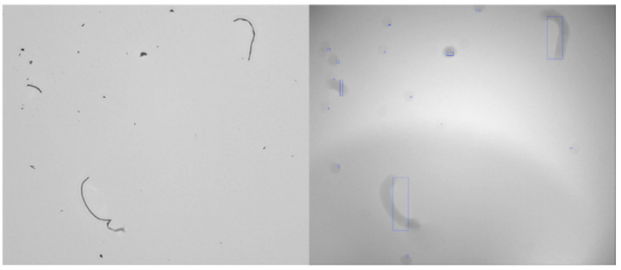


Fig. 1: Effect of particles in cameras: On the left, there is the debris on an image sensor, on the right, there is the blemish on the image output

lead to the development of Advanced Driver Assistance Systems (ADAS) [8]. However, with advances in sensor manufacturing technologies and the rise of deep learning (DL) [3], it is now possible to enable the vehicle to perform more complex tasks with greater levels of autonomy than ever before. Today features like automated braking systems, adaptive cruise control, park assistance, and lane keeping are not rare and available in commercially sold cars. As we are on our path to full vehicle automation, or driverless cars, ensuring the reliability of physical components used for perception is of utmost importance, especially those that relies on AI-based detection and segmentation of objects.

The reliability of components and the resulting performance of AI models are typically disconnected due to different quality standards. It is challenging to assess the reliability of camera-based functions in AI models, particularly deep learning models, prior to deployment because of their black-box nature. Therefore, quality criteria need to be aligned with the model's performance to ensure accurate assessment.

In this paper, we study the effects of one particular production defect called a blemish (see Figure 1) on the object detection function enabled by the camera-based perception performed by a DL-model using synthetic data. Such corruptions are the result of dirt or dust particles present in the manufacturing environment that gets them trapped on the image sensor, covering the pixels it lies on. This may result in partial to complete occlusion of objects present in the field of vision (FOV). A central question is: *"How severe does a defect have to be (e.g., in particle size and intensity) to significantly deteriorate a vision model such that the part has to be scrapped?"*

One challenge is to collect data with such blemishes, as they are generally avoided in the manufacturing stage due to clean room measures. This raises the question of defining the quality criteria for these defects in cameras. A very cautious quality target may lead to excessive scrapping of cameras, resulting in wasted time and money, a lenient target would result in substandard product quality. Another important consideration is that even if such systems existed, it would be necessary to investigate what would happen to the performance of the detection models if such defects were deemed "fit" for use. While many approaches currently aim to integrate ML and DL models into a Decision Support System (DSS) for improved performance and reduced intra-operator variability in quality control, our approach focuses instead on deriving conventional and interpretable quality gates for defects that still yield acceptable results in a downstream deep-learning-based object detection model.

Thus, the overall contributions of this work are to provide a statistical proof of quality criteria based on the critical AI based-functions for perception are used such that the scrap rates can be improved and the production efficiency is increased. We use publicly available FOV images that are synthetically tainted with blemishes of varying size (see Section 2), fine-tune a pretrained object detection model based on deep learning, and evaluate how the simulated defects affect the detection quality in Section 3.

### *1.1. Related Work*

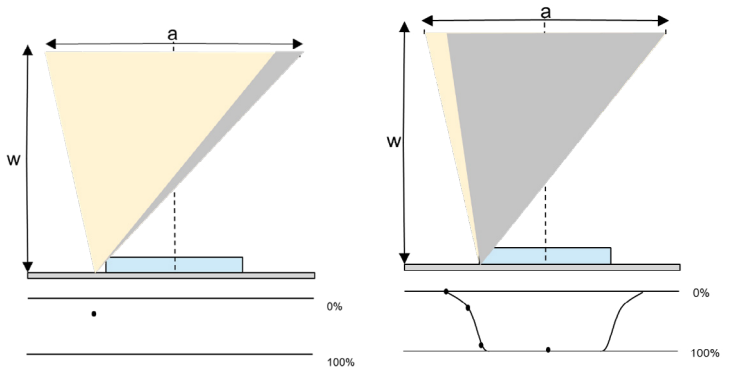
There is ongoing research in the field of investigation of the robustness of perception functions and also methods for improvement. [9] provides an overarching study by defining what corruptions and robustness are in deep neural networks. Further, the study gives an overview of all existing methods to improve corruption robustness including data augmentation and many others [5]. Another relevant work is [2] in which image classifiers were benchmarked in the presence of certain standard corruptions like Gaussian Noise, pixelation etc. However, this work is not specific to the use case of autonomous driving. In this domain, a robust detection benchmark was introduced by [6], investigating the effects of certain corruptions based on Hendryck's framework [2] as well certain datasets containing fog or time

of day corruptions. One of the few papers that deals with investigating the effects of camera-related failures on object detection models is [7] in which the authors identify various failure modes associated with production-related failures of a camera system, including blemish along with other corruptions. The effects of these corruptions are also studied on various object detectors. However, the defects such as blemishes and other dirt that can appear on the camera lens are grouped together and their effect on the models is studied together. In order to define blemish characteristics, we need to trace the physical model of the shadow created by the interaction of light with dust particles on an image sensor. Such studies are also limited in the existing literature. One notable exception is the paper by [10] on the artifacts produced by dust particles on lenses. In this work, the artifacts are modeled using geometrical optics to find the relation between lens parameters and dust size. However, the intensity distribution is only qualitatively modeled and shown to be a Gaussian distribution or a top hat distribution depending on the size of the particle and the diameter subtended by the light rays on the lens. However, all of these works do not elaborate on a specific mathematical model that can be used for injecting blemish corruption into images. This work aims to bridge these gaps by first establishing a model for the blemish artifacts and also defining severity in the context of autonomous driving applications. Moreover, our approach is the first to systematically investigate the effect blemishes of varying size and intensity have on a downstream deep-learning-based object detection model to derive quality gates for manufacturing.

## 2. Methodology: Creating realistic synthetic blemishes and fine-tuning an object detection model

### *2.1. Simulating Blemishes on Existing Image Datasets*

Dust artifacts can be represented as an array of attenuation factors for each pixel in the camera [10]. These artifacts can then be incorporated into test images by pixel-wise multiplication with attenuation kernel values. The attenuation factor can be understood as a proportion of the collection cone blocked by the dust.




(a) Light occlusion is minimal outside the particle (b) Light occlusion increases within the particle

Fig. 2: Light occlusion levels at different positions

From Figure 2, it is clear that the attenuation factor depends on the size of the particle, the image distance and apparent aperture (entrance pupil), the diameter of the entrance pupil as well as the thickness of the particle As the thickness of the particle decreases, the falloff in attenuation reduces and the shadow formed would be approximately the silhouette of the particle itself. As the size of the particle reduces to a point, we can see the constancy in attenuation drop of 100% reduces to a point spot resulting in an approximate Gaussian function.

These considerations allow us to make a model for the blemish function which can be represented by a symmetric super-Gaussian function as shown in Figure 3. The attenuation function, denoted as $f(x, y)$, can be represented in 2D

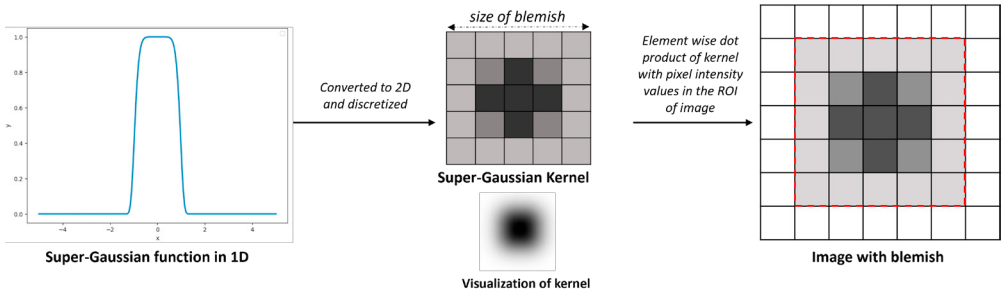



Fig. 3: Blemish injection by pixel manipulation: A super-Gaussian function (see Equation (1)) is applied in two dimensions on an image to make its pixels appear darker than they originally are.

and is symmetric about the $X$ and $Y$ axes. It is given by:

$$f(x, y) = Q \cdot \exp\left(-\left(\frac{(x-x_0)^2}{2} + \frac{(y-y_0)^2}{2}\right)^P\right) \tag{1}$$

The value of $P$ determines the rate of intensity fall-off and $Q$ represents the peak drop in pixel intensity value. This should be calibrated based on the particle thickness and object-entrance pupil distance.

- At $P = 0$, the attenuation function reduces to a normal Gaussian function.
- As $P \to \infty$, the attenuation function tends towards a rectangular function.

The selected datasets, KITTI [1] and BDD100k [11], do not contain any blemishes which is why we transform them into synthetic datasets with blemishes. The symmetric super-Gaussian kernel represents the attenuation matrix which is then applied to the image using pixel-wise multiplication as shown in Figure 3. This multiplication is done by first extracting the region of interest, which in our case, is the area where the blemish is to be injected (represented by the red box), and then applying the pixel-wise operation to individual color channels of the image. The location of the blemish, the diameter, and the peak intensity (the intensity of the darkest pixel) can be varied as per the user or test requirement and results in images as depicted in Section 2.1. The process can be iterated multiple times to create multiple blemishes on the image which also depends on user/test requirement.

(a) Clean Image from dataset [1]

(b) Blemish injected image

Fig. 4: Example of synthetic blemishes, inserted onto a clean image at known positions of objects.

### 2.2. Model fine-tuning

We choose the YOLOv5m model [4] for the purpose of our work as it provides the right trade-off with respect to the size of our dataset and training time. The training is done on the pre-trained model by readjusting weights of the backbone, neck, and head using the BDD100k dataset – with "clean" images only, to prevent the model from learning to ignore blemishes. No additional augmentation was performed during the training of the model as the dataset is sufficiently large and diverse. Further augmentation may result in augmenting the robustness of the model which would lead to possible error-masking when in fact we aim to understand the worst-case scenario. The model is trained

on a BDD100k training dataset which consists of 70k images and validation set of 20k images. At this point, it is important to emphasize that commercial-grade models such as those offered by vendors like Mobileye will probably be even stronger at compensating for minor camera defects. Nevertheless, an open-source fine-tuned YOLOv5 serves as an upper bound estimation of decrease in precision.

### *2.3. Blemish placement scenarios*

Our aim is to evaluate the risk associated with the blemish properties defined by the probability of a blemish appearing on any object, the probability of an object appearing in any specific size group, and the severity of the blemish. We conducted experiments on individual frames (not video sequences) to trace the effect of blemish size and intensity on the performance of the models. The experiments involved varying the diameter of the blemish from 5px to 50px and the intensity from 0.1 to 0.9. For the perception function used in autonomous driving, it is crucial to identify objects near the vehicle under consideration more accurately than those farther away. To assess this, we consider object sizes in the image, as they are inversely correlated with the distance from the vehicle. By dividing the object sizes into four quartiles (0 to 304 pixels, 304 to 811 pixels, 811 to 2947 pixels, and 2947 and more pixels), we gain greater insight into the model inference degradation when a blemish of a particular diameter and intensity is placed on objects that are close to or far from the vehicle. This approach enables us to evaluate the effects of blemishes in relation to the distance of objects from the vehicle and their impact on the overall risk. We conceived two relevant scenarios, one evaluating the worst case that could happen if blemishes directly align with ground truth objects, and another which places blemishes at random to more adequately model the manufacturing distribution.

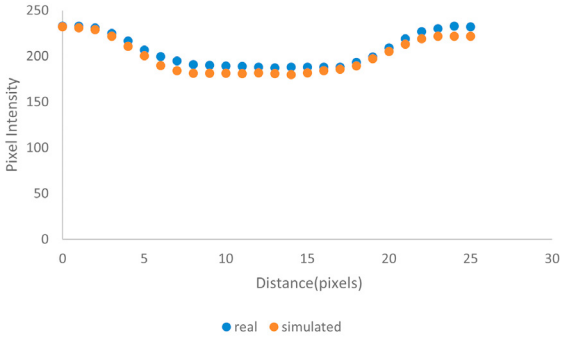
*Stress test – Blemishes placed directly on objects.* In this test, we want to evaluate the object detection model by constraining the location of the blemish. This is done by placing the blemish in the ground truth (GT) bounding box of the objects in the image. This scenario represents the "worst case scenario" as the blemish appears directly on an object to be detected, leading to its occlusion. The diameter of the blemish can be either kept fixed or varied based on a predetermined ratio with respect to the object size. When the diameter is set to a constant value, is possible that the blemish on an object may also result in partial or complete occlusion of nearby objects and augments the effect of existing blemishes, making it difficult to draw a direct conclusion on the blemish characteristics to the model degradation.

*Coverage test – Blemishes placed uniformly at random.* While the stress test is beneficial for understanding the worst-case scenario and allows for establishing a relation between the object size and blemish size on the degradation of the model, in reality, the probability that every object in a frame is covered by a blemish is highly unlikely. In order to simulate a more realistic scenario, we generate blemishes at random locations. Through the coverage test, we aim to establish the correlation between the effective area covered by blemish to the degradation in model performance. This is done by a brute force algorithm that generates blemishes of varying diameter but a fixed intensity at locations such that no two blemishes would overlap until the predefined coverage (expressed as the percentage of image size) is reached. For example, a 0.01% coverage defined by the user or the test for an image of resolution 1280px x 720 px would mean the algorithm would generate blemishes at random locations until 0.01 x 1280 x 720 = 9216 pixels are covered by the blemishes. The result only gives a probabilistic estimate regarding the influence since a comprehensive study should ideally include an extremely large sample.

## 3. Experimental Results

### *3.1. Validity of Synthetic Blemishes*

We discuss the impact of injecting the blemish augmentation on images on the object detection capabilities of our fine-tuned Yolov5. Before evaluating this impact, it is important to check if the mathematical model derived for a blemish within this work is a good representation of blemishes found in reality. In order to do this, we check how well the intensity distribution plot for the theoretical blemish(based on super gaussian equation) and observed blemish correlates, as can be qualitatively compared in terms of distance and pixel intensity in Figure 5a. The goodness of fit can then further be quantified using a $\chi^2$-test, as depicted in Figure 5b.




(a) Pixel intensity profile across one exemplary blemish which gets darker in the center – shown for the synthetic and real one.

| Intensity Range | Synthetic | Real |
|---|---|---|
| 151 - 155 | 35 | 30 |
| 156 - 160 | 25 | 20 |
| 161 - 165 | 16 | 12 |
| 166 - 170 | 13 | 10 |

(b) $\chi^2$ results

Fig. 5: Grounding of simulated blemishes in real data. No significant difference in intensity distribution was detected (significance level 0.05)

### 3.2. Baseline and Metrics

We first look at the results obtained using the clean images in the BDD100K validation dataset shown in Table 1. These results highlight the model's capability to handle clean images and establish a baseline for comparison with the blemish-augmented dataset.

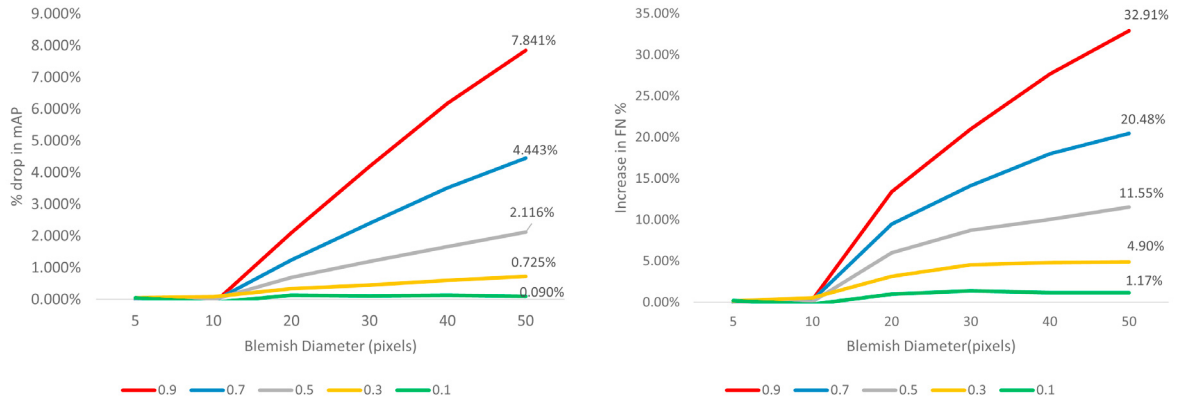
The mAP is a comprehensive metric that measures average precision values at various recall values across all object classes. This gives a holistic evaluation of the overall detection capability of the model at a given threshold of Intersection over Union (a measure of how well the predicted and ground truth bounding boxes overlap) – here we chose 50%, then called mAP50. In the context of autonomous driving, another critical metric is False Negatives which indicates the situations when the model failed to recognize the presence of an object. For example, when looking at car objects only, our fine-tuned YOLOv5 reaches a mean average precision (mAP) of 74 % and fails to detect 74,450 objects out of 236,790, i.e., 31%. This is safety-critical because instances like missing a vehicle or pedestrian could have fatal consequences. By understanding the baseline metrics given in Table 1, we can put the deterioration of mAP and FN in the presence of blemishes into perspective.

Table 1: Baseline Object Detection Metrics

| Object Class | Number of instances | Precision | Recall | mAP@50 | False Negatives |
|---|---|---|---|---|---|
| all | 428306 | 73% | 45% | 50.1% | NA |
| traffic light | 62593 | 67% | 56% | 57% | 27747 |
| traffic sign | 79672 | 74% | 56% | 62% | 35349 |
| car | 236790 | 77% | 69% | 74% | 74450 |
| person | 30648 | 72% | 49% | 56% | 15659 |
| bus | 4004 | 68% | 53% | 59% | 1885 |
| truck | 9824 | 68% | 54% | 60% | 4546 |
| rider | 1504 | 71% | 39% | 45% | 915 |
| bike | 2287 | 69% | 42% | 48% | 1324 |
| motor | 934 | 65% | 37% | 40% | 591 |
| train | 50 | 1 | 0.00 | 0.00 | NA |

### 3.3. Stress test

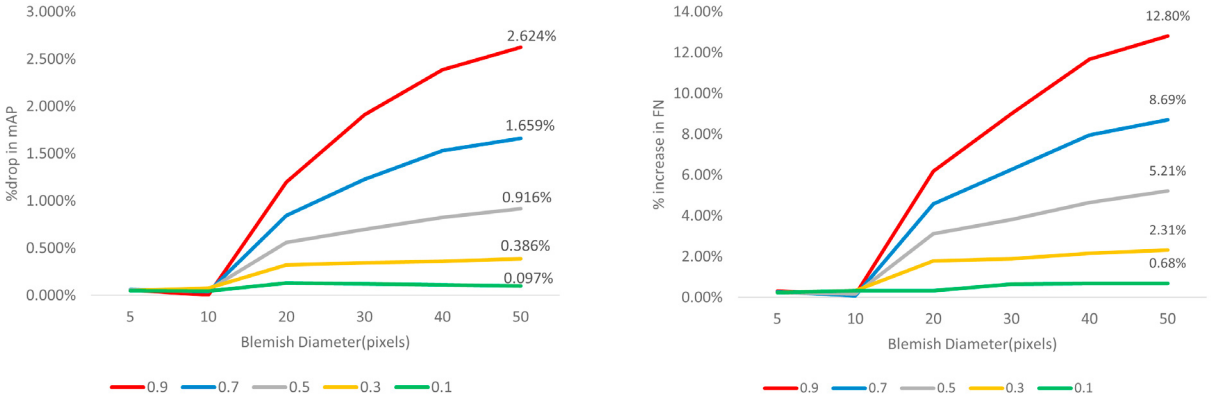
We first evaluate the effect of blemish diameter and intensity on the object detection model's inference. During this test, we also want to gain insight into how the object size (directly proportional to the proximity of the object from the vehicle) on which the blemish appears affects the occlusion levels and corresponding model performance degradation.




(a) Effect on mAP

(b) Change in FN%

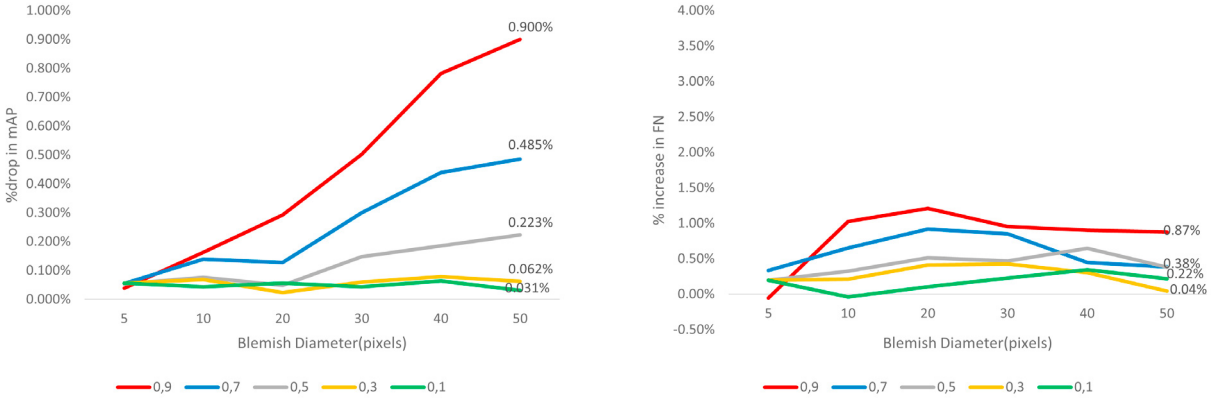
Fig. 6: Drop in mAP and increase in FN for the stress test across all object sizes and varying intensity drops (0.9 referring to a 90% lower intensity, i.e., darker pixels, 0.1 to a 10% lower intensity).




(a) Change in mAP

(b) Change in FN %

Fig. 7: Drop in mAP and increase in FN for the stress test on the lowest quartile group object size (0 to 304 pixels).




(a) Change in mAP

(b) Change in FN%

Fig. 8: Drop in mAP and increase in FN for the stress test on the highest quartile group object size (2947 and more pixels).

Looking at Section 3.2, we can inspect the results across all object classes and sizes. For blemishes up to 10 pixels in size, there is no noticeable change in either mAP nor FN. Using a blemish intensity with just 0.1 (referring to a maximum intensity drop, i.e., darkening, of 10% of the original pixel intensity in the region of interest of placement in the image) only leads to an mAP drop of 0.097 % – even if the blemish diameter is 50 pixels (cf. Figure 6a). Similarly, blemishes with intensity 0.9 would led to an increase of 12.8 % in the false negative rate (FN), i.e., missed objects (cf. Figure 6b). Roughly translated to our baselines in Table 1, instead of a 30% FN rate for cars, we could expect about a 43% FN rate. Figures 7 and 8 show the same plots for the lowest and highest of the four quartiles of object sizes – with the others (not depicted) leading to similar results.

When looking at the largest objects (i.e., the closest ones in the FOV) in Figure 8, even a distribution of high-intensity high-diameter blemishes does not provoke more than a 1% increase in FN or decrease in mAP. For these objects most nearby, the effect of blemishes is relatively low. If we are willing to accept a 1% change in mAP or FN, across all objects (Section 3.2), blemishes with an intensity of 0.1 or 0.3 are tolerable up to almost all diameters (at least in terms of mAP). We defer a discussion of acceptance thresholds to Section 3.6.

### *3.4. Likelihood Study*

We want to evaluate how likely the situations simulated in the above experiments can happen in real scenarios, i.e., being so unfortunate that objects are directly occluded by blemishes. Hence, we conduct a probabilistic study to evaluate this by generating coordinates sampled from a uniform distribution equal to the number of objects in the frame. The blemishes that happen to fall in the bounding boxes within an image are counted. We are interested in the ratio of the number of coordinates that appear on an object to the total number of objects in a given image. The number of overlaps or multiple blemishes occurring in the same bounding box is not considered here. The experiment was repeated a total of 10 times for each condition to ensure confidence.

Table 2 shows, in the event that the number of particles on sensor happen to be equal to number of objects in any given frame, how likely can it cover any object of a particular size group. Though there are an equal number of objects in each group, the probability of a particle appearing on objects of the largest size groups is the highest (almost 1 in 10 objects), which also follows our intuition.

Table 2: Likelihood of blemish appearing on an object. Size ranges refer to quartiles.

| Scenario | Percentage |
|---|---|
| Number of blemishes is equal to number of objects in FOV | 11.4% |
| Number of blemishes is equal to the number of all objects in 0-0.033% size range in an image | 0.1% |
| Number of blemishes is equal to the number of all objects in size range 0.033-0.088% in an image | 0.2% |
| Number of blemishes is equal to the number of all objects in size range 0.088-0.32% in an image | 0.7% |
| Number of blemishes is equal to the number of all objects in size range 0.32-89.9% in an image | 11.3% |

### *3.5. Coverage test*

The variation in mAP and false negatives across the experiments is noted for 3 different coverage levels: 0.01%, 0.1% and 1% with a fixed intensity of 0.9. The maximum intensity is chosen as the 0.9 to find the worst case scenario as observed from the previous studies. Tables 3 and 4 show the minimum and maximum change in metrics, across ten trials, for varying coverage percentages and FN and mAP, respectively. We find that when 1% of the pixels are covered, there is larger variation across the ten experiment trials for false negatives while the mAP change remains consistent. Also, we find that 1% coverage of pixels also leads to a maximum increase in 1 object additionally missed per every 100 objects detected.

### *3.6. Key Findings*

The main objective of our study is to determine the critical thresholds for blemish characteristics such as diameter and intensity, as well as the critical object size at which blemishes have detrimental effects on the perception model.

Table 3: Percentage change in False Negatives

| Percentage of pixels covered by blemishes | Min Change | Max Change |
|---|---|---|
| 1% | 0.4% | 1.2% |
| 0.1% | 0% | 0.34% |
| 0.01% | 0.23% | 0.25% |

Table 4: Percentage change in mAP

| Percentage of pixels covered by blemishes | Min Change | Max Change |
|---|---|---|
| 1% | 0% | 0.31% |
| 0.1% | 0.11% | 0.21% |
| 0.01% | 0.0% | 0.11% |

Table 5: Risk framework to derive threshold for blemish properties: Color coding shown instead of numeric values.

| Experiment | Critical Diameter | Critical Intensity | Likelihood of blemish on object | Severity due to change in metric |
|---|---|---|---|---|
| All objects | 20 | 0.1 | | |
| Object size between 0 and 304 pixels in area | 20 | 0.1 | | |
| Object size between 304 and 811 pixels in area | 20 | 0.3 | | |
| Object size between 811 and 2947 pixels in area | 20 | 0.3 | | |
| Object size larger than 2947 pixels in area | 50 | 0.9 | | |

This would further provide guidance on the acceptable particle size based on assumptions about blemish properties which could be applied directly in manufacturing to predict scrap. Our aim is to evaluate the risk associated with the blemish properties defined by:

$$\begin{aligned}\text{Risk} = {} & \text{Probability of blemish appearing on any object} \\ & \times \text{Probability of object appearing in any specific size group} \\ & \times \text{Severity}\end{aligned}$$

The probability study gave insight into how likely a blemish can appear on an object when the number of blemishes is equal to the number of target objects on which they are supposed to appear in the experiments that expose the critical diameter and intensity. However, the probability of a blemish appearing on the sensor should be derived from the production setting.

The most conservative case is defined by the first experiment where we analyzed the worst-case scenario of a blemish appearing on every object in an image. However, it is important to note that from the probability study, we found that the likelihood of a randomly generated blemish appearing on an object is 0.113. On average, there are 18.93 objects per frame. Considering this and assuming independence for every object, the probability for a blemish to appear on every object simultaneously is $0.113^{19} \approx 10^{-18} \approx 0$. This indicates that this scenario is extremely unlikely. Still, the critical diameter and intensity derived from the worst-case scenario are a diameter of 20 pixels and intensity 0.3 for mAP, a diameter of 20 pixels and intensity 0.1 for FN.

For groups of objects based on size ranges (expressed as a percentage of image size), we note the likelihood in Table 5 for both the metrics along with the severity. The overall severity is defined by two factors:

- Severity due to the size of the object: This comes from domain knowledge that the severity of damage is more when the detection of an object near the vehicle gets affected by the vehicle than the ones farther from it.

- Severity due to the metric change: This comes from the experiments conducted where the effect of model performance due to blemish is measured in terms of mAP and false negatives.

In Table 5, we only mention the severity due to size of object since the critical diameter and intensity mentioned is derived from the severity due to metric change. Also note that the likelihood is defined only by the probability experiment and does not depict how probable it is in a real case scenario. This is because a blemish with a large diameter can also affect the performance of the model when it appears near an object. This scenario is not considered in the study as the diameter of a blemish is found to be smaller than most object sizes that are of relevance in the FOV. The likelihood in real scenarios is also contributed by other factors like how likely a blemish of particular intensity can occur. While it is difficult to establish general acceptance thresholds, an analysis like the one proposed, enables quality engineers to get an insight into the behavior of blemishes and set appropriate limits on intensity and diameter during manufacturing. This has to be aligned with the current scrap rate and quality goals.

## 4. Conclusion and Future Work

We presented a methodological blueprint for deriving quality gates for manufacturing front camera modules taking into account the effect of synthetic blemishes on the object detection performance in a downstream deep learning model. To ensure the safety and reliability of these modules, which are vital for autonomous driving applications, a testing system aligned with functional performance was proposed. Experiments were conducted to determine critical blemish diameter and intensity, considering worst-case and average-case scenarios throughout different object sizes. Safety limits were established, and the risk associated with blemishes appearing on objects was assessed. By presenting a methodology for analyzing the trade-off between risk and quality, stakeholders can make informed decisions regarding blemish thresholds. This work is based on simulating blemishes using a mathematical approximation and conducting experiments to establish critical quality gates for blemish properties. However, real production environments may exhibit variability in particle sizes, shapes, and thickness on the image sensor. One approach to address this could involve using variational autoencoders or generative adversarial networks to find distributions adjacent to the simulated ones, introducing greater testing variations. Another direction of research is exploring more DL architectures and datasets. Finally, the same testing strategy could be adapted to include other use cases of downstream computer vision tasks such as semantic segmentation or scene forecasting.

**Acknowledgments**

This work was partially funded by Bavarian Ministry of Economic Affairs, Energy and Technology in the BayVFP project PALIM – Performance-Accelerated Learning for Intelligent Manufacturing (DIK-2195-0062 / DIK0331/02).

## References

[1] Geiger, A., Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets robotics: The kitti dataset. The International Journal of Robotics Research 32, 1231–1237.

[2] Hendrycks, D., Dietterich, T., 2019. Benchmarking neural network robustness to common corruptions and perturbations. arXiv preprint arXiv:1903.12261 .

[3] Huang, Y., Chen, Y., 2020. Autonomous driving with deep learning: A survey of state-of-art technologies. arXiv preprint arXiv:2006.06091 .

[4] Jocher, G., Stoken, A., Borovec, J., NanoCode012, ChristopherSTAN, Changyu, L., Laughing, Hogan, A., Lorenzomammana, Tkianai, YxNONG, AlexWang1900, Diaconu, L., Marc, Wanghaoyang0106, Ml5ah, Doug, Hatovix, Poznanski, J., , L.Y., Changyu98, Rai, P., Ferriday, R., Sullivan, T., Xinyu, W., YuriRibeiro, Reñé Claramunt, E., Hopesala, Dave, P., Yzchen, 2020. ultralytics/yolov5: v3.0. Zenodo. doi:10.5281/zenodo.3983579.

[5] Lai, Z., 2022. Cameras in advanced driver-assistance systems and autonomous driving vehicles, in: Advanced Driver Assistance Systems and Autonomous Vehicles: From Fundamentals to Applications. Springer, pp. 213–243.

[6] Michaelis, C., Mitzkus, B., Geirhos, R., Rusak, E., Bringmann, O., Ecker, A.S., Bethge, M., Brendel, W., 2019. Benchmarking robustness in object detection: Autonomous driving when winter is coming. arXiv preprint arXiv:1907.07484 .

[7] Secci, F., Ceccarelli, A., 2020. On failures of rgb cameras and their effects in autonomous driving applications, in: 2020 IEEE 31st International Symposium on Software Reliability Engineering (ISSRE), IEEE. pp. 13–24.

[8] Shaout, A., Colella, D., Awad, S., 2011. Advanced driver assistance systems-past, present and future, in: 2011 Seventh International Computer Engineering Conference (ICENCO'2011), IEEE. pp. 72–82.

[9] Wang, S., Veldhuis, R., Strisciuglio, N., 2023. The robustness of computer vision models against common corruptions: a survey. arXiv preprint arXiv:2305.06024 .

[10] Willson, R.G., Maimone, M.W., Johnson, A.E., Scherr, L.M., 2005. An optical model for image artifacts produced by dust particles on lenses .

[11] Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., Madhavan, V., Darrell, T., 2020. Bdd100k: A diverse driving dataset for heterogeneous multitask learning, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 2636–2645.