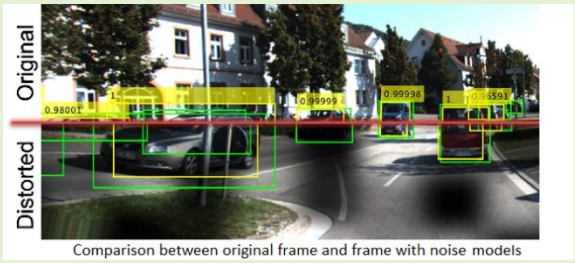
This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3211406

Analysis of Automotive Camera Sensor Noise Factors and Impact on Object Detection

Boda Li, Pak Hung Chan, Gabriele Baris, Matthew D. Higgins, and Valentina Donzella

***Abstract*—Assisted and automated driving functions are increasingly deployed to support improved safety, efficiency, and enhance driver experience. However, there are still key technical challenges that need to be overcome, such as the degradation of perception sensor data due to noise factors. The quality of data being generated by sensors can directly impact the planning and control of the vehicle, which can affect the vehicle safety. This work builds on a recently proposed framework, analysing noise factors on automotive LiDAR sensors, and deploys it to camera sensors, focusing on the specific disturbed sensor outputs via a detailed analysis and classification of automotive camera specific noise sources (30 noise factors are identified and classified in this work). Moreover, the noise factor analysis has identified two omnipresent and independent noise factors (i.e. obstruction and windshield distortion). These noise factors have been modelled to generate noisy camera data; their impact on the perception step, based on deep neural networks, has been evaluated when the noise factors are applied independently and simultaneously. It is demonstrated that the performance degradation from the combination of noise factors is not simply the accumulated performance degradation from each single factor, which raises the importance of including the simultaneous analysis of multiple noise factors. Thus, the framework can support and enhance the use of simulation for development and testing of automated vehicles through careful consideration of the noise factors affecting camera data.**



Comparison between original frame and frame with noise models

***Index Terms*— Assisted and Automated Driving, Environmental perception sensor, Image quality for machine learning, Sensor degradation, Noise factors, Deep neural networks.**

I. Introduction

FROM the initial DARPA Grand Challenge competition, to the testing of Waymo car on Californian roads, automotive manufacturers have been steadily increasing their offering of automated functions such as Adaptive Cruise Control, Lane Assist, Park Assist, etc. The primary drivers for these innovative features include enhanced safety, efficiency, and consumer convenience. However, the development pace has been slower than expected at times, as there are still a number of technical challenges associated with replacing the human driver with automated systems that need to be fully addressed. These technical challenges are limiting the demonstration that automated vehicles are actually safer than a human driver, which in turn impacts the acceptance of them by the wider public. It is key to formally prove the robustness of these systems under challenging and unexpected situations such as failures or degradation of crucial components in the vehicle (e.g. a failure of a perception sensor, such as a camera) under extreme environmental conditions and corner cases [1–3].

Focusing on robustness, assisted and automated functions can be represented as a pipeline of elementary functions, i.e. *sensing – perception – decision – control*. In this pipeline, *sensing* represents the interface between the external world and the vehicle processing systems (either the real physical system or a simulated one in the virtual environment). In assisted and automated driving functions the *sensing* step is implemented by a suitable combination of a number of perception sensors, specifically of cameras, LiDARs, RADARs, and ultrasonic sensors. The data produced by these sensors will be used to perceive the needed elements and build the situational awareness of the ego-vehicle. As a consequence, any alterations to the generated data can mislead the vehicle's understanding of the surrounding environment and of road stakeholders, causing wrong and potentially dangerous planning and acting decisions. Problems related to generated data can include: data out of synchronisation; misalignments in the sensor coordinate systems; degradation of the sensor outputs or of sensor performance; missing or erroneous data; artifacts in the data; etc. It is therefore of remarkable importance to understand how the data generated by perception sensors can be affected by different sources of noise, namely noise factors, and also to try and quantify the impact of the different noise factors on the sensor outputs.

Recent studies have focused on the analysis of a specific noise factor on a sensor type, e.g. rain on LiDAR [4], [5], interference on RADAR [6], [7], fog on camera [8]. However, an in depth investigation of noise factors is required. The automotive environment is highly complex, swiftly changing, and multiple noise sources can impact the sensor outputs simultaneously. Therefore, there is the need to consider the

combination of multiple noise factors, and how to merge and prioritise them. A recently presented framework to support the noise factor analysis of automotive sensors has demonstrated the impact on a LiDAR sensor due to specific environmental noise factors (i.e. rain and obstructions), acting separately or in a combined way [9]. The results clearly demonstrate that considering noise factors individually provides an incomplete prediction and/or quantification of the sensor output degradation, and that the possibility to combine multiple noise factors in simulation and during testing is key to understand the impact on perception. Leveraging the proposed framework [9], this paper performs an innovative and thorough analysis on automotive cameras and combine this analysis with a demonstration of the impact of combined noise factors on camera-based perception.

Cameras are fundamental sensors in automotive as many assisted and automated driving functions (such as lane centring, pedestrian detection, traffic sign recognition) have been traditionally based on camera data [10]. Furthermore, camera data processing can rely on well-established and mature computer vision and machine learning techniques [11]. The framework enables the identification and classification of the different camera noise factors, highlighting the main outputs affected by them and providing guidance to the modelling of degradation of automotive camera images [9]. Moreover, in this paper, the framework is applied to analyse and quantify the impact of two unavoidable and independent noise factors, either separately and/or as a compound noise, namely: 1) windshield tangential distortion, and 2) obstruction due to stains or dirt. The developed single and compounded noise factor models can be used in any testing environment to inject and quantify the effect of camera data degradation. The impact of the noise factors is quantified by evaluating the original and degraded images with two deep neural network (NN)-based object detectors. The evaluation shows that there is a maximum decrease up to ~13% in the precision and recall when the above-mentioned noise factors are applied individually. When combining windshield tangential distortion and obstruction noise factors, the maximum decrease in performance is higher, up to ~18%.

For the first time, this work classifies the sources of noise in automotive cameras based on 5 categories, analysing, and discussing the implications on 4 sensor outputs. The presented results, related two modelled noise factors, demonstrate the significant importance of identifying the noise factors acting on environmental perception sensors and of building parametric models to analyse the effects of single and compound noise factors on automotive sensors to have an accurate evaluation of the implications on perception.

II. Analysis of Noise Factors

A list of camera noise factors is generated through considering the five noise factor types of a typical P-Diagram [9], as shown in Table I. The thirty noise factors, identified from the p-diagram categories, are further investigated to understand how they affect the camera outputs, sec. II.A. Each noise factor can then be modelled and applied to the correct part of the sensor model or collected sensor data to generate degraded data/datasets to be used for further investigation. The combination of different noise factors requires further analysis and can bring to complex relationships between noise models when the factors are not independent, sec. III-IV.

A. Noise Factor Evaluation

A list of noise factors, Table I, was created through experience, an extensive review of the topic and discussions with experts from industry and academia. The list of noise factors does not include sources and situations that can cause the generation of incorrect data, but instead are better classified as corner cases. For example, surface reflection resulting in "ghosting" of objects is a corner case and therefore is not an entry in the table. However, it is important to complement Table I with these types of corner cases and investigate ways of modelling them. In Table I, the camera outputs affected by each noise factor are highlighted. This work considers a general camera sensor model including as outputs: frame rate (**FR**, in Hz or fps), value of the intensity stored in each pixel ($\mathbf{I_{RGB}}$), position of pixel values ($\mathbf{P_{(x,y)}}$), and dropped frames (**DF**). In particular, dropped frames differ from the frame rate as this output considers when one or more frames are not transmitted, whereas the generation of the specified frames per second stays the same. Additionally, $\mathbf{I_{RGB}}$ can affect any single channel pixel intensity, whereas $\mathbf{P_{(x,y)}}$ affects all the intensity channels in one spatial position of the pixel matrix (i.e. it can be one channel if Bayer images are transmitted, or three channels when transmitting Red-Green-Blue, RGB, ones).

B. Related Work

This section is focused on the works related to two of the noise factors identified in the above analysis (i.e. 21 and 25, Table I). These two noise factors have been selected as they very frequently occur on camera (e.g. even if the sensor is calibrated, some residual distortion is noticeable in the real dataset collected by a camera mounted behind the windshield [12], and obstructions while driving are unavoidable). The proposed noise models (sec. III) and case study on the effect of single and compound noise factors (sec. IV-V) are based on these noise factors.

Camera sensors are made up of different components, for example, the lens is used to focus the radiation for a designed field of view. These lenses will refract the incoming light, to direct it into the correct pixels of the sensor. Many automotive cameras have wide field of view lens, resulting in noticeable optical distortions, which can be decomposed into tangential and radial distortion [13] [14]. Radial distortion, such as pincushion or barrel distortion, changes based on the pixel radial distance from the sensor central point. Conversely, tangential distortion is based on an angle difference between the sensor plane and the lens, resulting in the image being tilted, stretched on one end, and compressed on the other end [13], [15], [16]. Tangential distortion can be also caused when cameras are mounted behind the windshield (e.g. a common position for functions like lane keep assist and sign recognition), that has its own curvature, slightly deflecting

TABLE I

AUTOMOTIVE CAMERA NOISE FACTORS IDENTIFIED THROUGH THE PROPOSED FRAMEWORK [9]. FOR EACH NOISE FACTOR, THE AFFECTED SENSOR OUTPUTS ARE IDENTIFIED: FRAME RATE (**FR**), INTENSITY PER PIXEL ($\mathbf{I_{RGB}}$), POSITION OF PIXEL ($\mathbf{P_{(x,y)}}$), DROPPED FRAMES (**DF**)

Factor Type	ID/ Noise Factor	FR	I_{RGB}	$P_{(x,y)}$	DF	Description
Piece to Piece	01. Alignment		✓	✓		Misalignment between sensor and lens during assembly.[22]
	02. Fabrication Variability	✓	✓			CMOS fabrication variability (photodiode, circuitry, Analogue to Digital Converters).[23] [24]
	03. Lens shape, purity		✓	✓		Lens fabrication variability, resulting in non-ideal absorption and refractions.[25][26][27]
	04. Dark Current variability		✓			There are mechanisms of compensating the current generated by the photodiodes in presence of no light, but usually there is a variability of this current from pixel to pixel. [28]
	05. Image Signal Processing (ISP)	✓	✓	✓		ISP alters the data gathered by the image sensor; functions implemented can include: denoising, demosaicing, colour correction, white balancing, sharpening edges, etc. [29]
Change over Time	06. Ageing of electronics	✓	✓			Degradation of the performance of the electronic components, resulting in effects such as increased/decreased resistance, leaking currents, etc. [30]
	07. Degradation of Lens		✓	✓		Lens wear out and ageing resulting in attenuation and refractions. [20] [31] [32]
	08. Vibration of Mounting		✓	✓		Long term effect of vehicle vibrations resulting in loosening of mounting. [33]
	09. Pollutant Ingress		✓	✓	✓	Ingress of particulates such as dust, water, condensation. [34]
	10. Pixel Degradation		✓			Exposure to electromagnetic waves resulting in degradation of silicon doping and reduction of pixel performance. [35]
	11. Board Ageing	✓	✓		✓	Printed circuit board degradation over time such as whisker/dendritic, connector pin contact degradation. [36]
Usage	12. Misplacement of the sensor		✓	✓		Change in the positioning of the sensor due to terrain and vehicle, causing a variation of the sensor coordinate system (axes and/or angle) with respect to original calibration.[37]
	13. Vehicle impact		✓	✓		Impacts on the camera unit or vehicle which results in misalignment of the sensor/lens. [26]
	14. Chemicals/ Contaminants		✓	✓		Cleaning materials and chemicals may react with the lens surface and cause irreversible damage. [38]
	15. Obstructions		✓	✓		During driving, materials/particles (e.g. water, stains, etc.) can obscure/refract the incoming light.[39] [40]
	16. Lens Scratch		✓	✓		Scratches can reflect and attenuate the incoming light differently to intended.[20]
	17. Vehicle dynamic settings		✓	✓		Adjusting height of vehicle through weight, tyres, pitch, loaded weight etc. thereby changing the sensor coordinate system with respect to original calibration. [41]
Environment	18. Sensor saturation or depletion		✓			Scenes with extreme brightness or very low luminosity (i.e. sunrise, sunset, exiting tunnel) can cause saturation or depletion of areas of pixels and therefore an inaccurate rendering of the scene. [42]
	19. Extreme Temperature	✓	✓		✓	Sensor operating in conditions outside of manufacturer recommended temperature. [43]
	20. Adverse Weather		✓			Conditions such as rain, snow, fog, sleet, frost, mist, etc. [44]-[48]
	21. Optic Obstructions		✓	✓		Obstructions, such as mud, stains, frost, water spray, flies, etc., which are partially or fully on the lens or windshield which can block or refract the light [19]-[21][40]
	22. Low Illumination		✓			Low light resulting in required high pixel gain, creating a larger difference in output intensity between adjacent pixels due to increased noise. [49] [50]
	23. Sun		✓			The sun can cause local saturation, lens flare, IR detection into the colour channels. [51]-[53]
System Interactions	24. Malicious Attacks	✓	✓	✓	✓	Artificial alteration of the image (externally or internally) i.e. cyberattacks, external light source attacks, etc.[54]
	25. Windshield			✓		Curvature of windshield which changes the angle at which light enters the sensor.[17] [55]
	26. Power Supply	✓	✓		✓	Unstable or varying supplied power causing variations in generated signals. [56]
	27. EMI	✓	✓		✓	Electromagnetic Interference (EMI) from start-up/shut-down of electronics, motors, etc. inducing current within the sensor wires and connections.[57]
	28. Saturation of Buffer		✓	✓	✓	Sensor internal buffer saturation, causing problems on transmitted data flow, e.g. inability to process incoming data, incorrect stored data, etc.[58] [59]
	29. LED flicker		✓			Pulsing LEDs in the environment resulting in fluctuations in the generated images.[60]-[62]
	30. Localised Light source		✓			Headlight, flashlight, high beams, laser beams, etc. [63] [64]

incoming light [14], [17], [18]. The developed model for windshield tangential distortion is described in sec. IV.A.

Moreover, automotive cameras can have different installation positions on the vehicle depending on the external area they need to cover with their field of view. However, vehicle exterior is subject to road conditions which can cause a build-up of dirt or other deposits on the lens or windshield, partially covering the view of the camera. This obstruction could range from tainting a few areas of the image to fully obscuring a part of the image [19]. Einecke et al. have explored the possibility of detecting the presence of these artefacts [20]. They have demonstrated the detrimental effects of small water droplets and stains, which can be frequently splashed onto the lens or windshield obstructing camera's view, for example small water droplets create localised blurring of the image. Some methods have been proposed to

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3211406

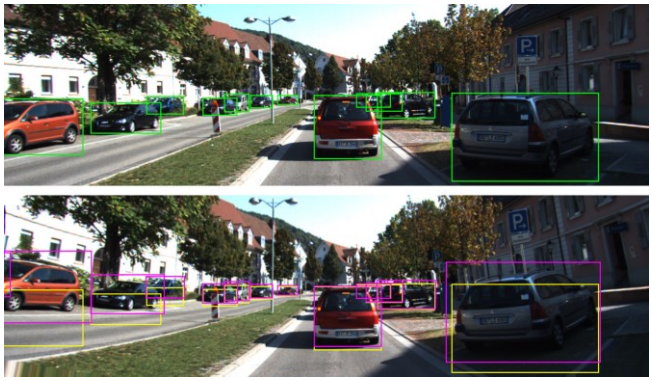


Fig. 1. From top to bottom: (top) original image with ground truth bounding boxes; (bottom) geometrically distorted image with original ground truth bounding boxes in yellow and corrected bounding boxes as new ground truth in purple.

Fig. 2. Detections in the distorted image with IoU: (top) minimum obstruction kernel size l = 12*12 pixels (bottom) minimum obstruction kernel size l = 72*72 pixels.

reduce this kind of noise, e.g. Kokubo et al. investigated a method to remove raindrops from images [21]. In the images used by Kokubo et al. the effects of larger raindrops can be observed, and they are more pronounced than small raindrops. A larger area of the image is blurred, but also each drop creates a larger attenuation around the edges resulting in a visible outline of the raindrop in the image. However, de-noised data will be always an approximation of the original data. On the other hand, the larger patches of stains can obstruct parts of the image fully, rendering them darker, and the edges of thinner patches can tint areas of the image with brown. The obstruction model used in this paper is based on the dataset presented in [19] and described in sec. III.B.

III. Noise Models

This section describes the noise models developed for the selected noise factors (sec. II. B) as a part of this work. The noise models are then used to build noisy sensor datasets, as discussed in the next section, and evaluated to show the effects of individual and combined noise factors.

A. Windshield Tangential Distortion Model

By placing a forward-facing camera behind a windshield (as commonly implemented in most modern vehicles), the external light needs to travel through the windshield before arriving to the sensor. Windshields normally come with certain amount of surface curvature and inclination for aerodynamic and safety reasons. The curvature will be usually small, and a tangential distortion model can be used to describe this distortion [17]. In fact, radial distortion can be attributed to stronger distortion due to lens characteristics and misplacement with respect to the sensor centre, but, as this distortion is normally corrected during the calibration process (by sensor manufacturers), it can be omitted from the model here. Windshield distortion is unlikely to be calibrated by automotive manufacturers for each vehicle. Even a slight misplacement of the camera due to vibrations/changes over time can recreate a misalignment, so windshield distortion is an important noise factor to be investigated.

In the Brown-Conrady polynomial (BCP) model, commonly applied for calibration, the frame distortion caused by lenses and screens can be divided into radial and tangential distortion and thus corrected. The BCP model has three parameters k_1, k_2, k_3 to consider radial distortion and two parameters p_1, p_2 for tangential [71]. Considering only tangential distortion, p_2 had a minimal impact on the outcome, therefore a simplified version of the BCP model considering only p_1 is opted to use, as expressed in (1).

$$(x_{WD}, y_{WD}) = \left(\left(2p_1 x_p * y_p\right), \left(p_1\left(r^2 + 2{y_p}^2\right)\right)\right) \quad (1)$$

Where the original pixel position $(x_p, y_p)$is transformed into (x_{WD}, y_{WD}) by applying the model, and r is the distance from the centre of the sensor (i.e. $r^2 = x^2 + y^2$).

The application of the distortion model affects the position of the objects within the image, as shown in Fig 1. To ensure that the ground truth bounding boxes still correctly encompass the vehicles in the frames, they are also distorted along with the model. Each corner of the original bounding box was distorted by (1) to give the positions of the distorted corners. The maximum and minimum of the x and y coordinates are used to provide the new rectangular bounding boxes. The implications of the change in bounding boxes are highlighted in sec. V.A.

B. Obstruction Model

As mentioned in Table I, there are several different causes/materials that can obstruct the lens of camera. 'Obstructed lens' can cause serious problems, e.g. mud or any other residuals can solidify on the lens (or windshield), covering areas of the pixel matrix. In turn, this covered area can degrade the ability to detect objects in the frames. Uricar's study includes 40 minutes of clean and mud-obstructed images, collected in parallel by identical surround view cameras [19]. The obstruction of one camera is realised by spraying different intensity of mud on lens. From inspecting the frames from the dataset in [19], the main variations associated with the obstructions are the position of the covered area, the size and intensity of the obstructions.

Inspired by Uricar's work, the model proposed in this work creates a mask that changes the intensity of some of the pixels considered to be covered by obstructions. The mask is generated with varying obstruction dimension, and randomly moving this obstructed area in different variants, Fig. 2. For simplicity, the obstructions are considered as squared areas with side l (in pixels). The maximum intensity attenuation of the obstruction is randomly generated and has a Gaussian profile on the area to create a smooth variation of the intensity. For

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3211406

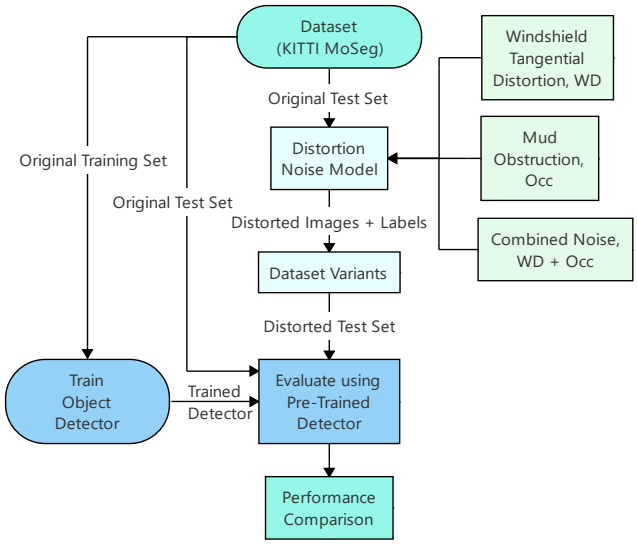


Fig. 3. Schematic representation of the methodology to add noise factors to camera data. We have used a NN based detector to evaluate the effects of image degradation due to single and combined noise factors on the perception step.

each pixel, the noisy intensity is generated as

$$\bar{I}_{Occ} = m_{x,y} * \bar{I}_p \quad (2)$$

where $m_{x,y}$ is the attenuation of each pixel ($0 \leq m_{x,y} \leq 1$).

C. Compound Noise Model

The compound model can be easily expressed as in (3), similarly to what was proposed for LiDAR in [9]. In the equation (3), f_{WD} and f_{Occ} represent the noise factor equations (1) and (2). These equations are used to transform the original pixel coordinates and intensity, (x_p, y_p) and $\bar{I}_p$ (here $\bar{\boldsymbol{I}}$ is an array with the 3 RGB values), into the noisy one, (x_n, y_n), $\bar{I}_n$, for every pixel in the sensor. Given the selected noise factors, implications on frame rate and dropped frames are neglected. The proposed equation applies for each pixel in the pixel matrix.

$$[(x_n, y_n), \bar{I}_n] = f_{WD}\left[(x_p, y_p), f_{Occ}(\bar{I}_p)\right] \quad (3)$$

When noise factors are not independent, their combined effect will need to be carefully analysed and formulated. In this case, the two noise models just described are combined as shown in (3), and the parameter p_1 and the size l, respectively described in sec. III.A and sec. III.B, are varied in steps. The tangential distortion model is applied after the obstruction model assuming that the splashes are attached to the outer surface of the windshield and therefore their effect on the pixel will be distorted too due to the curvature of the windshield depending on the pixel position.

IV. Case Study: Methodology

The previously proposed framework has been deployed in this work to analyse the effects of single and compound noise factors on automotive cameras, to identify and classify all their noise sources, sec. II [9]. The effect on the sensor outputs has been highlighted in Table I. The next step consists of applying the developed noise models (sec. III) to camera data (either real or simulated data). To evaluate the impact on perception introduced by the noise factors, a deep neural network for object detection has been used in this work. Average Precision (AP) and max recall are calculated for the dataset with no noise factors and compared to the AP and max recall values for datasets with noise factors at different strengths. The use of NN based perception to evaluate the impact of distortion/degradation of sensor data has been discussed in [64]. The flow of designed experiments is schematically represented in Fig. 3; the generation of the dataset variants (with single and compound noise factors with different intensities) is discussed in Section IV-A, the selected NN in sec. IV.B. As explained in sec, II, two unavoidable and independent noise factors are chosen as a demonstration: distortion due to windshield (i.e. tangential deflection), and obstructed (i.e. stained) lenses.

A. Dataset

Most of the state-of-the-art NN object detectors use supervised training that is based on labelled data, therefore an existing labelled dataset, namely the KITTI MoSeg dataset, is used for our experiments. The KITTI MoSeg dataset is a subset of the original KITTI dataset [66][67]. The MoSeg test set consists of 349 labelled camera images captured by one of the cameras on a vehicle driving in urban streets in the Netherland; the images are collected in good weather condition and with at least one vehicle captured and labelled per frame. The dataset has been pre-processed and rectified. KITTI is a benchmarking automotive dataset used for object classification and segmentation [68]; the classes 'car' and 'van' are merged into a single class, as the focus of this work is on the distortion of the images and their main road stakeholders. The proposed methodology has been tested on other datasets, e.g. the Caltech Cars dataset [69,70], exhibiting results in line with the performance reported in sec. V.

It is a substantial challenge to have curated and labelled datasets for automotive, which not only have enough scenarios to cover sufficient 'real-world' variability and corner cases, but also consider distortion and degradation caused by the noise factors described in the previous section. Thus, in this work, to evaluate and compare the NN performance, an 'ideal' dataset (thereafter named 'the original dataset'), and generated distorted dataset variants are used. Namely, the following dataset variants are generated: 6 variants with tangential distortion of varying magnitude; 6 variants with obstructed lens with different obstruction intensity and position; 36 variants with the compound effect of the two previously mentioned noise factors. Applied noise factors are described in sec. II-III. The values of p_1 and l (Sec. III. A-B) have been selected by visually inspecting the generated frames and comparing them with the datasets previously mentioned in sec. II; namely for tangential distortion it has been selected $p_1 = 0$, -3×10^{-5}, -6×10^{-5}, -9×10^{-5}, -12×10^{-5}, -15×10^{-5}, -18×10^{-5}, and for obstruction $l = 0, 12, 24, 36, 48, 60, 72$ in number of pixels. The p_1 values used for tangential distortion have been selected to mimic real windshield distortion observed in real camera data and datasets, and validated via visual inspection. As previously discussed, in the variants that contain the geometric distortion, it is necessary to distort the ground truth bounding boxes together with their related frame, and depending on the position of the bounding boxes and of the

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and
content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3211406

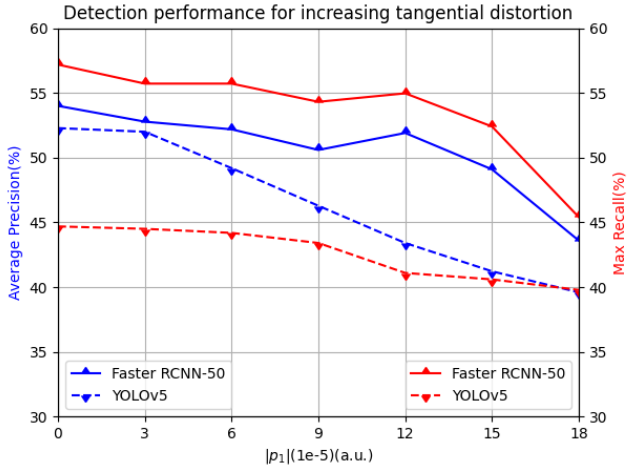


Fig. 4. AP and max recall of the Faster R-CNN and YOLO when evaluating images from the original dataset (p_1=0) and from dataset variants with increasing tangential distortion (i.e. increasing absolute value of p_1).

amount of distortion. Distorted ground truth bounding boxes are potentially bigger in area than original ones so that all the pixels of distorted vehicles can be covered. An example of the original and distorted bounding boxes is shown in Fig. 1.

B. NN Based Object Detector

For the detection task, the most common NN object detectors can be broadly classified into one or two stage detectors. As a part of this work, a recent pre-trained YOLO. V5 [72] was deployed as a widely used one stage detector, and a faster region proposal deep convolutional NN (Faster R-CNN) architecture was selected as a two stage detector, offering almost real-time performance [11]. The selected pre-trained Faster R-CNN is based on a ResNet50 backbone, which works with RGB inputs and is a relatively small network (size of about 96 MB, 50 layers, 25.6 M parameters); this NN offers a good trade-off in terms of speed and accuracy, which are crucial in automotive applications. In all cases, the NNs have been re-trained and optimised using the original dataset; then they have been used to evaluate the original and variant datasets (assuming that pre-trained NNs will be deployed for perception in assisted and automated driving functions). The evaluation of results is based on AP and max recall, calculated for the 'vehicle' class, as described in [64], with a threshold for the Intersection over Unit (IoU) of 0.5.

V. Case Study: Results and Discussion

A. Detection Performance with Windshield Tangential Distortion

Fig. 4 shows AP and max recall when the selected NNs are used to evaluate datasets with only one noise factor, i.e. increasing tangential distortion (namely the absolute value of the parameter p_1 is increasing). This figure demonstrates that tangential distortion impairs the object detection performance of the network: AP is decreasing by almost 11% and 14%, and max recall by about 12% and 5%, for Faster R-CNN and YOLO respectively, across the considered values of distortion ($0 \leq |p_1| \leq 0.00018$) when compared to the results for the original data. However, by visual inspection of the results, there are interesting frames where distortion can fictitiously increase AP

Fig. 5. From top to bottom: detections in the original image with IoU values; detections in the distorted image with IoU value ($|p_1| = 0.00012$).

and max recall due to missing ground truth bounding boxes where there are some vehicles in the frames, as shown in Fig. 5 – ground truth and NN detected bounding boxes are in green and yellow respectively. The upper frame shows the detections on the original dataset and the lower on the distorted dataset with $|p_1| = 0.00012$. Fig. 5 shows that the distortion applied to the left most vehicle avoids the detection of this vehicle, which was not labelled in the ground truth. This vehicle is erroneously detected as a false positive in the original dataset and not detected at all in the variant dataset; however, due to the incorrect ground truth, the precision calculated on this single image will be higher in the case of distorted data. This case also shows the importance of a meticulous labelling of the ground truth in the datasets used for training and evaluation of NNs, as errors in the original labels can cause a fictitious decrease in the NN performance, when actually the NN would be able to identify correctly more objects than the ones labelled in the dataset. The importance of ground truth bounding boxes is discussed for MoSeg in [73] and it is outside the aim of this paper. Moreover, some fluctuations in the performance can be introduced by the process of distorting the bounding boxes (as discussed in section III.A), as their final size can be bigger with respect to the distorted targets, Fig. 1 (purple boxes in the lower image). The further away are the ground truth bounding boxes from the centre of the image, the more is the size increase of the distorted boxes due to tangential distortion. For distorted objects closer to the frame edges, it becomes less precise to represent the labelled targets in rectangular bounding boxes and the box sizes become remarkably bigger than the targets; therefore, it can be easier for the NN to have a higher IoU and the number of true positive detections can oscillate slightly, maintaining an overall decreasing trend with increasing distortion.

B. Detection Performance with Obstruction

Fig. 2 shows a visual example of the distorted images generated to build the dataset variants with partial obstruction due to stains. In the lower image, as the size of the areas covered by obstructions increases, more objects (and therefore vehicles) are covered by stains and this coverage has an impact on AP and max recall, as shown in Fig.6. However, in a few images, when the covered areas are sufficiently big and positioned nearby the object edges, the obstructed areas improve detection performance by highlighting the vehicle through darkening out the

This article has been accepted for publication in IEEE Sensors Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2022.3211406

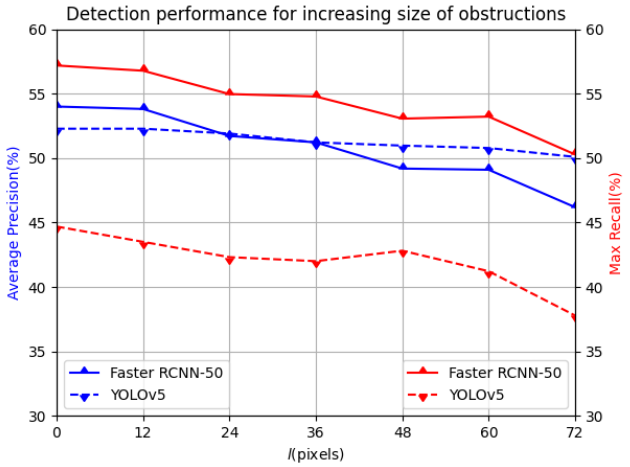


Fig. 6. AP and max recall of the Faster R-CNN and YOLO when evaluating images from the original dataset (l=0) and from dataset variants with increasing size of the obstruction kernel on the image (i.e. increasing l).

surrounding area. Nonetheless, calculating the metrics over the entire datasets, these rare cases are negligible.

Fig. 6 shows AP and max recall when the NNs are used to evaluate datasets with this single noise factor, i.e. obstruction caused by stains (namely the size l of the stain areas on the image has being increased). Similarly, to what discussed in sec. V.A, the detection performance is negatively affected when the noise factor strength is augmenting, with AP decreasing by 8% and 3%, and max recall decreasing by about 7% and 7%, for Faster R-CNN and YOLO respectively, across the considered values of distortion when compared to the results for the original dataset. It is worth noting that if the covered areas are smaller than 12x12 pixels the NN performance is negligibly affected (<1%).

C. Detection with Combined Distortion and Occlusion

The values of AP when evaluating datasets with the compound noise factor are shown in Fig.7. In this figure there is a trend similar to the results presented in Fig. 4 and Fig. 6. In fact, with the compound noise there is a performance decrease every time the magnitude of one of the two noise factors is increased. The slope of the plots for increased absolute value of p_1 (i.e. for increased tangential distortion) is comparable to the plots in Fig. 3, however when the obstruction noise (coloured plots in figures) increases, the maximum achievable AP decreases. The overall AP can drop by about 17.6% and 15.8% for the selected Faster RCNN and YOLO.v5 respectively. This result clearly shows that the compound noise affects the NN performance to a greater extent with respect to the single noise factors, and a separate noise factor analysis does not give a clear picture of the impact of multiple noise factors acting simultaneously on the sensor.

VI. Conclusion

This paper presents, for the first time, a thorough noise factor analysis carried out for automotive camera sensors. A breakdown list of thirty different noise sources and how they can affect sensor outputs in terms of frame rate, pixel position shift, intensity, and dropped frames is proposed and discussed. Moreover, two independent noise sources have been modelled and analysed, namely distortion due to windshield and obstruction. Their separate and combined effects on the images of well-established automotive datasets have been quantified by using state of the art perception algorithms based on deep neural networks. The results show that for all the considered noise factors an increase in the magnitude of noise causes a noticeable decrease in the performance, and the highest performance degradation is achieved for the compound noise. This work can pave the way to a comprehensive analysis and modelling of environmental perception sensors (and their noise factors) used for assisted and automated driving.

Acknowledgment

Boda Li acknowledges the support by the State Scholarship Fund by China Scholarship Council (CSC) from the Ministry of Education of P.R. China. Dr Donzella acknowledges that this work was supported by the Royal Academy of Engineering under the Industrial Fellowships scheme. The Authors wish to acknowledge the HVM CATAPULT. The Authors thank Gunny Dhadyalla, AESIN, and Dr Anthony Huggett, ON Semiconductor, for the extremely fruitful discussions.

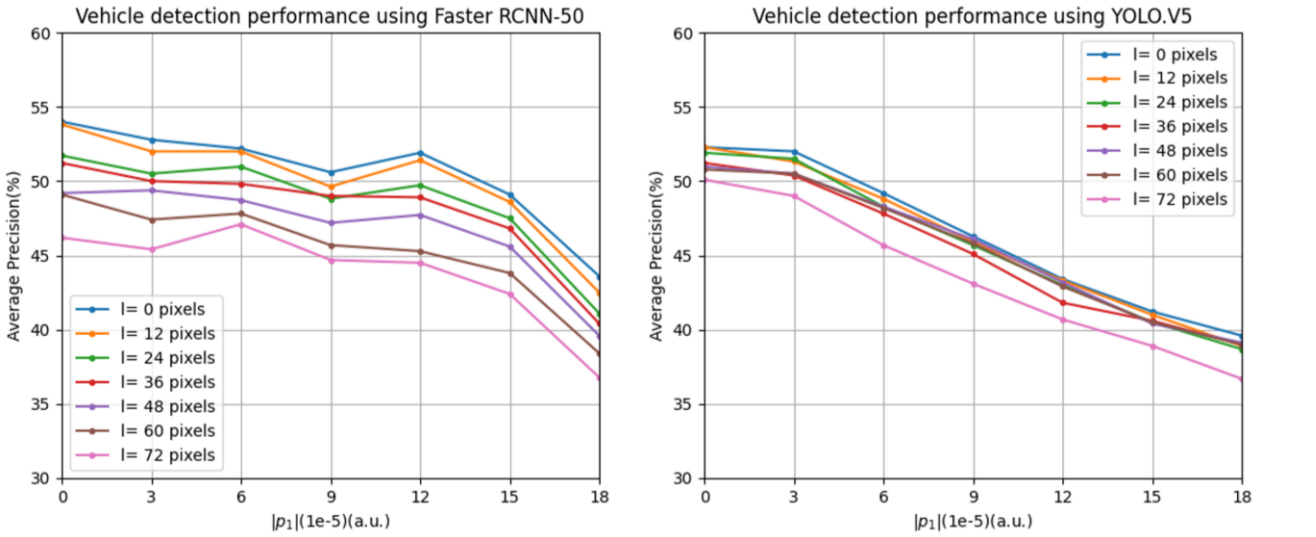


Fig. 7. Faster RCNN (left) and YOLO (right) versus increasing tangential distortion due to the windshield. Different colours represent different amounts of occlusion due to the increase of the obstruction kernel size.

REFERENCES

[1] S. M. P. Nidgi Kalra, "How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability?," 2014.

[2] W. Kröger, "Automated vehicle driving: background and deduction of governance needs," *J. Risk Res.*, vol. 24, no. 1, pp. 14-27, 2021.

[3] J. Robinson, J. Smyth, R. Woodman, and V. Donzella, "Ethical considerations and moral implications of autonomous vehicles and unavoidable collisions," *Theor. Issues Ergon. Sci.,* 2021.

[4] K. Montalban, C. Reymann, D. Atchuthan, P. E. Dupouy, N. Riviere, and S. Lacroix, "A quantitative analysis of point clouds from automotive lidars exposed to artificial rain and fog," *Atmosphere(Basel).,* vol. 12, no. 6, 2021.

[5] J. P. Espineira, J. Robinson, J. Groenewald, P. H. Chan, and V. Donzella, "Realistic LiDAR with Noise Model for Real-Time Testing of Automated Vehicles in a Virtual Environment," *IEEE Sens. J.,* vol. 21, no. 8, pp. 9919-9926, 2021.

[6] S. Alland, W. Stark, M. Ali, and M. Hegde, "Interference in Automotive Radar Systems: Characteristics, mitigation techniques, and current and future research," *IEEE Signal Process Mag.,* vol. 36, no. 5, pp. 45-59, 2019.

[7] F. Uysal and S.Sanka, "Mitigation of automotive radar interference," in *2018 IEEE Radar Conf.,* pp. 405-410, 2018.

[8] K. M. Hudd, M. Thornton, and A. Richards, "Automotive sensing: assessing the impact of fog on LWIR, MWIR, SWIR, visible, and lidar performance," pp. 50, 2019.

[9] P. H. Chan, G. Dhadyalla, and V. Donzella, "A Framework to Analyze Noise Factors of Automotive Perception Sensors," *IEEE Sensors Lett.,* vol. 4, no. 6, 2020.

[10] F. E. Sahin, "Long-range, high-resolution camera optical design for assisted and autonomous driving," *Photonics,* vol. 6, no. 2, 2019.

[11] M. C. García, J. T. Mateo, P. L. Benítez, and J.G. Gutiérrez, "On the performance of one-stage and two-stage object detectors in autonomous vehicles using camera data," *Remote Sens.,* vol.13, no. 1, pp.1-23, 2021.

[12] M. Braun, S. Krebs, F. Flohr and D. Gavrila, "EuroCity Persons: A Novel Benchmark for Person Detection in Traffic Scenes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 8, pp. 1844-1861, 2019. Available: 10.1109/tpami.2019.2897684.

[13] K. Lelowicz, "Camera model for lens with strong distortion in automotive application," in *2019 24th Int. Conf. Methods Model. Autom. Robot. MMAR 2019,* pp. 314-419, 2019.

[14] A. Furnari, G. M. Farinella, A. R. Bruna, and S. Battiato, "Affine covariant features for fisheye distortion local modeling," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 696-710, 2017.

[15] S. Lee and H. Hua, "A robust camera-based method for optical distortion calibration of head-mounted displays," in *Proc. - IEEE Virtual Real.,* pp. 27-30, 2013.

[16] S. A. Schneider and K. Saad, "Camera behavioral model and testbed setups for image-based ADAS functions," *Elektrotechnik und Informationstechnik,* vol. 135, no. 4–5, pp. 328–334, 2018.

[17] F. Verbiest, M. Proesmans, and L. Van Gool, "Modeling the Effects of Windshield Refraction for Camera Calibration," *Lect. Notes Comput. Sci.* vol. 12351, pp. 397–412, 2020.

[18] C. Krebs, P. Müller, and A. Braun, "Impact of Windshield Optical Aberrations on Visual Range Camera Based Classification Tasks Performed by CNNs," presented at *London Imaging Meet.,* vol. 2021, no. 1, pp. 83–87, 2021.

[19] M. Uřičář, J. Ulický, G. Sistu, H. Rashed, P. Křížek, D. Hurych, A. Vobecký, and S. Yogamani, "Desoiling Dataset: Restoring Soiled Areas on Automotive Fisheye Cameras," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pp. 4273-4279, 2019.

[20] N. Einecke, H. Gandhi, and J. Deigmöller, "Detection of camera artifacts from camera images," in *2014 17th IEEE Int. Conf. Intell. Transp. Syst.,* pp. 603-610, 2014.

[21] Y. Kokubo, S. Asada, H. Maruyama, M. Koide, K. Yamamoto, and Y. Suetsugu, "Removing raindrops from a single image using synthetic data," in *Proc. Int. Conf. Pattern Recognit.,* pp. 2081-2088, 2020.

[22] A. Fetić, D. Jurić and D. Osmanković, "The procedure of a camera calibration using Camera Calibration Toolbox for MATLAB," in *2012 Proceedings of the 35th International Convention MIPRO, 2012,* pp. 1752-1757, 2012.

[23] K. Kurita, T. Kadono, R. Okuyama, S. Shigemastu, R. Hirose, A. Onaka-Masada, Y. Koga, and H. Okuda, "Proximity gettering technology for advanced CMOS image sensors using carbon cluster ion-implantation technique: A review," *Physica Status Solidi (a),* vol. 214, no.7, pp.170-216, 2017.

[24] V. Goiffon, M. Estribeau, O. Marcelot, P. Cervantes, P. Magnan, M. Gaillardin, C. Virmontois, P. Martin-Gonthier, R. Molina, F. Corbiere, S. Girard, P. Paillet, and C. Marcandella, "Radiation Effects in Pinned Photodiode CMOS Image Sensors: Pixel Performance Degradation Due to Total Ionizing Dose," in *IEEE Transactions on Nuclear Science,* vol. 59, no. 6, pp. 2878-2887, Dec. 2012.

[25] D. H. Cha, H. J. Kim, Y. Hwang, J. C. Jeong, and J. H. Kim, "Fabrication of molded chalcogenide-glass lens for thermal imaging applications," *Appl. Opt.* vol. 51, pp.5649-5656, 2012.

[26] N. Vekariya, and F. Hultenius Syversen, "Cost Calculating of Optoelectronics-An Analysis of Cost Items for a Camera Application used in the Automotive Industry at Volvo Cars Corporation," 2011.

[27] Kowa Optimed, "Polychromatic Diffraction MTF LM3JC10M (3.7mm F2.8)", 2021.

[28] A.N. Abarca Prouza and A.J.P.A.M. Theuwissen, "In-pixel temperature sensors for dark current compensation of a CMOS image sensor", Ph.D., Delft University of Technology, 2021.

[29] G. Portelli, and D. Pallez, "Image signal processor parameter tuning with surrogate-assisted particle swarm optimization," *in International Conference on Artificial Evolution (Evolution Artificielle),* pp. 28-41, 2019.

[30] D.G. Yang, F.F. Wan, Z.Y. Shou, W.D. van Driel, H. Scholten, L. Goumans, and R. Faria, "Effect of high temperature aging on reliability of automotive electronics," *Microelectronics Reliability,* vol. 51, no. 9-11, pp. 1938-1942, 2011.

[31] S. Reichel and F. Lentes, "Blue glass lens elements used as IR cut filter in a camera design and the impact of inner quality onto lens performance," *Proc. SPIE 8550*, Optical Systems Design 2012.

[32] Y. Lai, X. Kuang, P. Zhu, M. Huang, X. Dong, and D. Wang, "Colorless, transparent, robust, and fast scratch-self-healing elastomers via a phase-locked dynamic bonds design," *Advanced Materials,* vol. 30, no. 38, pp.1802556, 2018.

[33] T. T. Nguyen, X. D. Pham, J. H. Song, S. Jin, D. Kim and J. W. Jeon, "Compensating Background for Noise due to Camera Vibration in Uncalibrated-Camera-Based Vehicle Speed Measurement System," in *IEEE Transactions on Vehicular Technology,* vol. 60, no. 1, pp. 30-43, 2011.

[34] P. Khapane, V. Chavan, and U. Ganeshwade, "Water Ingress Analysis and Splash Protection Evaluation for Vehicle Wading using Non-Classical CFD Simulation," *SAE International Journal of Passenger Cars - Mechanical Systems,* vol. 10, no. 1, pp.183, 2022.

[35] P.R. Rao, X. Wang, A.J.P.A.M. Theuwissen, "Degradation of CMOS image sensors in deep-submicron technology due to γ-irradiation," *Solid-State Electronics,* vol. 52, no. 9, pp. 1407-1413, 2008.

[36] P. S. Karagiannopoulos, N. M. Manousakis, C. S. Psomopoulos, "A Novel ILP Formulation for PCB Maintenance Considering Electrical Measurements and Aging Factors: A "Right to Repair" Approach," *Energies,* vol. 15, no. 1, pp.183, 2022.

[37] C. Hughes, M. Glavin, E. Jones, and P. Denny, "Wide-angle camera technology for automotive applications: a review," *IET Intelligent Transport Systems,* vol. 3, no. 1, pp. 19-31, 2009.

[38] B. Marina, "The effects of accelerated aging on optical components: Application to vehicle camera systems," degree of Master, Linköping University, 2019.

[39] M. Uřičář, P. Křížek, G. Sistu and S. Yogamani, "SoilingNet: Soiling Detection on Automotive Surround-View Cameras," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC),* pp. 67-72, 2019.

[40] F. Pizzati, P. Cerri, and R. D. Charette, "Model-based occlusion disentanglement for image-to-image translation," in 2020 *European conference on computer vision,* pp. 447-463, 2020.

[41] H. E. Tseng and D. Hrovat, "State of the art survey: active and semi-active suspension control," *Vehicle System Dynamics,* vol. 53, no. 7, pp. 1034-1062, 2015.

[42] Z. Zhang, C. Forster and D. Scaramuzza, "Active exposure control for robust visual odometry in HDR environments," in *2017 IEEE International Conference on Robotics and Automation (ICRA),* pp. 3894-3901, 2017.

[43] A. Pelamatti, J. Belloir, C. Messien, V. Goiffon, M. Estribeau, P. Magnan, C. Virmontois, O. Saint-Pé, P. Paillet, "Temperature Dependence and Dynamic Behavior of Full Well Capacity in Pinned Photodiode CMOS Image Sensors," in *IEEE Transactions on Electron Devices,* vol. 62, no. 4, pp. 1200-1207, 2015.

[44] S. Hasirlioglu and A. Riener, "A Model-Based Approach to Simulate Rain Effects on Automotive Surround Sensor Data," in 2*018 21st International Conference on Intelligent Transportation Systems (ITSC),* pp. 2609-2615, 2018.

[45] S. Hasirlioglu and A. Riener, "Introduction to rain and fog attenuation on automotive surround sensors," in *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC),* pp. 1-7, 2017.

[46] Y. Hamzeh, S. A. Rawashdeh. "A Review of Detection and Removal of Raindrops in Automotive Vision Systems," *Journal of Imaging*, vol. 7, no. 3, pp.52, 2021.

[47] Y. Hamzeh, Z. A. El-Shair, A. Chehade and S. A. Rawashdeh, "Dynamic Adherent Raindrop Simulator for Automotive Vision Systems," in *IEEE Access*, vol. 9, pp. 114808-114820, 2021.

[48] T. Rothmeier and W. Huber, "Let it Snow: On the Synthesis of Adverse Weather Image Data," *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 3300-3306, 2021.

[49] B. Moomaw, "Chapter 11 - Camera Technologies for Low Light Imaging: Overview and Relative Advantages", in Methods in Cell Biology, G. Sluder and D. E. Wolf, Ed. *Academic Press*, 2013.

[50] S. Gilroy, J. O 'dwyer and L. Bortoleto, "Characterisation of CMOS Image Sensor Performance in Low Light Automotive Applications," in *Irish Machine Vision and Image Processing conference 2017,* pp. 278, 2017.

[51] M. A. Esfahani and H. Wang, "Robust Glare Detection: Review, Analysis, and Dataset Release," *arXiv preprint arXiv:2110.06006*, 2021

[52] I. Jatzkowski, D. Wilke and M. Maurer, "A Deep-Learning Approach for the Detection of Overexposure in Automotive Camera Images," in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pp. 2030-2035, 2018.

[53] K. Yoneda, N. Ichihara, H. Kawanishi, T. Okuno, L. Cao and N. Suganuma, "Sun-Glare region recognition using Visual explanations for Traffic light detection," in *2021 IEEE Intelligent Vehicles Symposium (IV),* pp. 1464-1469, 2021.

[54] D. H. Sharath Yadav and A. Ansari, "Autonomous Vehicles Camera Blinding Attack Detection Using Sequence Modelling and Predictive Analytics," *SAE Technical Paper,* 2020.

[55] C. Krebs, P. Müller, and A. Braun, "Impact of Windshield Optical Aberrations on Visual Range Camera Based Classification Tasks Performed by CNNs," in *London Imaging Meeting*, vol. 2021, no. 1, pp. 83-87, 2021.

[56] K. Yonemoto and H. Sumi, "A CMOS image sensor with a simple fixed-pattern-noise-reduction technology and a hole accumulation diode," in *IEEE Journal of Solid-State Circuits,* vol. 35, no. 12, pp. 2038-2043, 2000.

[57] N. Mutoh, M. Nakanishi, M. Kanesaki and J. Nakashima, "EMI noise control methods suitable for electric vehicle drive systems," in *IEEE Transactions on Electromagnetic Compatibility,* vol. 47, no. 4, pp. 930-937, 2005.

[58] A. El Gamal and H. Eltoukhy, "CMOS image sensors," in *IEEE Circuits and Devices Magazine*, vol. 21, no. 3, pp. 6-20, May-June 2005.

[59] J. A. Leñero-Bardallo, R. Carmona-Galán and Á. Rodríguez-Vázquez, "A Wide Linear Dynamic Range Image Sensor Based on Asynchronous Self-Reset and Tagging of Saturation Events," in *IEEE Journal of Solid-State Circuits,* vol. 52, no. 6, pp. 1605-1617, June 2017.

[60] M. Innocent, S. Velichko, D. Lloyd, J. Beck, A. Hernandez, B. Vanhoff, C. Silsbv, A. Oberoi, G. Singh, S. Gurindagunta, R. Mahadevappa, M. Suryadevara, M. Rahman, and V. Korobov, "Automotive 8.3 MP CMOS Image Sensor with 150 dB Dynamic Range and Light Flicker Mitigation," in *2021 IEEE International Electron Devices Meeting (IEDM),* vol. 30, no. 2, pp.1-4, 2021.

[61] B. Deegan, "LED flicker: Root cause, impact and measurement for automotive imaging applications," *Electronic Imaging*, vol. 2018, no.17, pp.146, 2018.

[62] N. Behmann and H. Blume, "Psychophysics Study on LED Flicker Artefacts for Automotive Digital Mirror Replacement Systems", *Electronic Imaging,* vol. 2020, no. 11, pp. 234, 2020.

[63] P. Sevekar and S. B. Dhonde, "Nighttime vehicle detection for intelligent headlight control: A review," *2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT)*, pp. 188-190, 2016.

[64] E. Oldenziel, L. Ohnemus and S. Saralajew, "Provident Detection of Vehicles at Night," *2020 IEEE Intelligent Vehicles Symposium (IV)*, pp. 472-479, 2020.

[65] P. H. Chan, G. Souvalioti, A. Huggett, G. Kirsch, and V. Donzella, "The data conundrum: compression of automotive imaging data and deep neural network-based perception," presented at *London Imaging Meet.,* vol. 2021, no. 1, pp. 78–82, 2021.

[66] M. Siam, H. Mahgoub, M. Zahran, S. Yogamani, M. Jagersand and A. El-Sallab, "MODNet: Moving Object Detection Network with Motion and Appearance for Autonomous Driving," 2017.

[67] G. Andreas, L. Philip, S. Christoph and U. Raquel, "Vision meets Robotics: The KITTI Dataset," *International Journal of Robotics Research (IJRR),* 2013.

[68] E. Mohamed et al., "Monocular Instance Motion Segmentation for Autonomous Driving: KITTI InstanceMotSeg Dataset and Multi-Task Baseline," in *2021 IEEE Intelligent Vehicles Symposium (IV),* pp. 114-121, 2021.

[69] M. Weber and P. Perona, 'Caltech Cars 1999'. CaltechDATA, Apr. 06, 2022. doi: 10.22002/D1.20084.

[70] B. Philip, P. Updike, and P. Perona, 'Caltech Cars 2001'. CaltechDATA, Apr. 06, 2022. doi: 10.22002/D1.20085.

[71] J. Wang, F. Shi, J. Zhang, and Y. Liu, "A New Calibration Model and Method of Camera Lens Distortion," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5713–5718, 2006.

[72] G. Jocher et al., "ultralytics/yolov5: v6.2 - YOLOv5 Classification Models, Apple M1, Reproducibility, ClearML and Deci.ai integrations", Zenodo.org, 2022. [Online]. Available: https://zenodo.org/record/7002879#.Yv53si7MJPY. [Accessed: 18-Aug- 2022].

[73] B.Li, G.Baris, P.H.Chan, A.Ramen and V.Donzella, "Testing ground-truth errors in an automotive dataset for a DNN-based object detector", in *Proc. of the International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, 2022 (under review)