COMPUTER METHODS IN MATERIALS SCIENCE

2024, vol. 24, no. 3, 41–50
https://doi.org/10.7494/cmms.2024.3.0836

Detecting dents in car bodies using machine learning and structured light projection

Izabela Potasz*, Sławomir Potasz, Michał Laska

VUMO sp. z o.o.

Abstract

This article discusses feasible methods for detecting dents in car bodies caused by transportation damage, commuting collisions, and hail. The authors review existing approaches exploiting their limitations, including smartphone-based ML detection algorithms and drive-through tunnels. The paper details the setup for capturing dents using computer vision with industry-grade cameras and structured light projection, emphasizing optimized data acquisition and computer vision setup. A particular emphasis is placed on acquiring high-quality input data thanks to the proper calibration and alignment of cameras, structured light, and the synchronization between them. Challenges related to obtaining high-quality footage in real-life conditions, such as car speed, body color, and lighting conditions, are thoroughly discussed. The method covers algorithms for detecting car paint, optimizing camera parameters, and identifying dents. Data annotation methods are described in detail, ensuring robust training datasets. Validation of the method is based on comparing the results of an inspection by professional car appraisers with algorithm detection outcomes. The results demonstrate the effectiveness of the proposed methods. Additionally, the article explores future research opportunities, such as scratch detection, damage severity estimation, and integrating these systems into automated production lines. The potential for enhancing vehicle inspection processes through advanced computer vision and structured light techniques is also considered.

Keywords: car body inspection, painted surfaces, structured light projection, dents, liquid crystals, light valve, automotive inspection

1. Introduction

Vision systems are extensively integrated into nearly all facets of the automotive industry. These systems play a crucial role in already assembled vehicles, such as reading traffic signs, monitoring speed limits, assisting with lane keeping, and tracking the driver's focus level. They are fundamental to autonomous self-driving technologies. Additionally, vision systems control matrix headlights, enable gesture-based interactions with infotainment systems, assist with vehicle parking, and provide surveillance when the vehicle is parked. In vehicle maintenance, computer vision systems inspect tire conditions, assess wheel alignment, and check for tire rubber expiration. However, a significant area within the automotive industry that computer vision systems still need to address adequately remains: shipping already manufactured cars from factories to car dealerships. In this domain, detecting dents and other surface damage on the car's body is crucial for quality control, claims management, and logistics management. Despite the advances in vision technology, these tasks often rely on manual inspections, which can be time-consuming and prone to human error.

A system for dent detection already exists in the assembly line setup where ambient light is controlled,

* Corresponding author: iza.potasz@vumo.ai
ORCID ID's: 0009-0002-8207-0821 (I. Potasz), 0009-0006-0432-272X (S. Potasz), 0009-0008-9359-3410 (M. Laska)

and the car travels through a light tunnel (Arnal et. al., 2017). This system, however, cannot work when exposed to sunlight; that is often the case in car auctions, logistic plants, and car dealerships. When considering fender-bender accidents or significant car accidents, it is essential to understand how many body parts will be replaced. In this case, a substantial part of the car is damaged and can be written off as a totaled vehicle. In such use cases, regular smartphone-based footage and big datasets can provide decent results (Patil et al., 2017). However, these systems are not precise enough to carefully estimate the number of dents in the car's body, especially after hail damage. Structured light profilometry can also be applied to scan the car body surface precisely and detect all dents (Van der Jeught & Dirckx, 2019). There are, however, a few issues with such a setup: First, it is a slow process that requires numerous images be captured for each camera position. Second, it is a costly process that requires expensive projectors that provide low-intensity light that can be easily surpassed by bright ambient light. It also gives unnecessary high precision of 3D geometry for all dents, so it is optional to do a cost estimate for a car that went through a hailstorm. Off-the-shelf 3D scanners are an obvious solution to this problem, as they can generate an object's 3D model, identify surface imperfections such as dents, and measure their depth. However, they are not suitable for cars. Car body scanning using a 3D range image scanner was also presented (Chen, 2008) but has limitations: a small field of view of up to 0.6 m × 0.6 m means plenty of expensive scanners are required. Moreover, the system generates an actual point cloud, which is too computing-heavy and cumbersome for deployment at a scale where computing costs are an issue. Finally, such systems project light that is not bright enough to overcome sunlight. Automobiles are intricate objects that are difficult to scan due to their complex silhouettes, varying shapes and sizes, and polished, highly reflective metallic surfaces. These characteristics make it impossible to monitor them successfully using motion-tracking 3D handheld scanners, as tracking can be easily lost. To utilize the 3D scanning method, cars must be covered in a matt layer like powder chalk. However, this approach is not feasible on a large scale due to the potential damages caused to the car's paint, low speed, and the high cost of such a process. It also generates significant pollution within the workstation.

A prominent challenge to the industry, therefore, is to develop a comprehensive computer vision system with an advanced data acquisition setup and sophisticated detection algorithms capable of working with complex, highly reflective objects like a car's body in a regular automotive plant setup. This system should detect dents and other popular car body blemishes such as scratches, peeling paint, and variations in paint thickness. The system must effectively inspect cars' reflective surfaces and varied shapes to ensure precise identification and analysis of all these imperfections, ultimately enhancing the quality control process in automotive manufacturing and maintenance. A system for car body inspection that detects dents in a controlled indoor environment has already been developed (Park et al., 2020). This system was able to detect and localize dents on vehicle bodies using a region-based convolutional neural network. Unfortunately, that setup was prepared to work in nearly perfect ambient light conditions of just 500 lx, far from the typical automotive plant/logistic hub setup. The light projection module had only 2,000 lx light intensity. This is why it is so important to develop a system that can operate in real-life scenarios without a need for extensive scanning location preparation. This study proposes purpose-built computer vision systems designed to work with high-intensity ambient light while maintaining high accuracy of convolutional neural network (CNN) detections. Robustness to ambient light and surrounding reflections is essential for achieving reliable and efficient quality control in the automotive industry, ensuring that all imperfections are accurately detected and addressed.

2. Possible data acquisition setups

A smartphone-based system for conducting a gap and flush inspection of a car's body has been successfully developed (Pham et al., 2021). However, such a system is optimized for measuring panel gaps, not dents. Furthermore, the laser was not visible in outdoor ambient light. While there are many research studies working on improving the thermal efficiency of the smartphone, including ways to encode computing data without applying electrical currents, which could result in better heat management (Wang Y. et al., 2019), phones, as of today, still frequently overheat. The example phone message about its overheating is 'iPhone needs to cool down.' Such a system would also be too slow to inspect all dents across the car's body, especially when considering that the iPhone that was used is prone to camera overheating when used for extended periods of time outdoors.

A structured light projection setup using R-CNN networks for car dent detection has also been prototyped in the past (Park et al., 2020). Such a system provides faster feature extraction thanks to the use of

black-and-white stripes that are projected into the car's body. It also helps to standardize the input dataset, regardless of the car's body color. However, it also introduces certain risks. The system used a low-intensity light source, meaning that a slow shutter speed was used that could introduce motion blur if the vehicle was moving through the computer vision data acquisition setup. Such a system is very vulnerable to changing ambient light conditions as well as strong reflections that usually occur in vehicles with dark paint colors. Furthermore, its applicability to real-world scenarios such as automotive plants might be limited due to ambient light intensity levels being tough to control due to constantly opened garage shutter doors and high-speed requirements as human beings were required to scan the car's body centimeter after centimeter. It also makes detecting the top of the car very impractical/impossible due to the height of a car compared to the height of an average person.

As a result, our research was intended to propose a system setup that would overcome ambient light, eliminate strong reflections in the dark car paint, and enable fast shutter speed to acquire the data without motion blur. The proposed data acquisition technique with a strobe light structured light projection setup offers a robust solution by using high-intensity, short-duration light pulses to capture images. Such a setup approach enables the system to be much more robust in changing ambient light conditions, especially when ambient light is over 40,000 lm (8× what was described in the papers). The described setup can work outdoors while exposed to sunlight. The key in this system is precise timing and synchronization between the camera, strobe light source, and mechanical shutter.

3. Developing a selected data acquisition setup

As the very first step, selecting a proper camera is essential to ensure synchronization between the flash lamps and the camera sensor at high shutter speed is possible. Precise synchronization is also critical to ensure the camera sensor captures light exclusively from the structured light projection, minimizing interference with the ambient light. Ideally, the camera sensor should be exposed only when the structured light projection is at its peak power. Any deviation, even as little as 1 ms of integration before or after the optimal projection window, results in ambient light reaching the camera sensor. This significantly reduces the contrast between the reflected structured patterns and the ambient light, introducing noise into the acquired data and compromising the accuracy of the measurements. A blatant but cost-prohibitive setup would be to use a global shutter camera, which simultaneously exposes all lines in the camera sensor. These sensor types are known for their high performance when working with strobe lights. However, to reduce costs and make the computer vision more scalable – as many cameras are required to capture an entire vehicle in one drive-through – the decision was made to use a more affordable rolling shutter camera operating in a global reset shutter sensor mode. This type of sensor mode has some advantages over a global shutter sensor but at a much lower cost, making it a practical choice for applications with multiple cameras.

In the global shutter sensor, all the pixels of the array are exposed simultaneously. After all pixels are correctly exposed, sensor readout is performed, starting from the top and going line by line to the bottom. In the rolling shutter sensor, the sensor starts exposing pixels line by line, again starting from the top and going to the bottom. Sensor readout is done for each line after the exposure is performed.

In a rolling shutter camera's global reset shutter sensor mode, all lines of pixels start the exposure simultaneously, but they finish when the readout for a previous line is completed. This means that the top of the image is darker than the bottom, which was exposed for longer. For example, at 29 Hz acquisition speed, the first line of pixels is exposed for 1 ms, but the last one for 35 ms. This is a severe issue when a light burst from a strobe light lasts for 1 ms: for 97% of the exposure time of the last lines of pixels, the sensor captures ambient light, introducing noise into the acquired data. To counteract this limitation, the shutter speed was fixed, and the image exposure (brightness of an image) was controlled via lens shutter and strobe light. A special light valve was designed and added to the lens, which was necessary to make the rolling shutter sensor with a global reset shutter behave like a global shutter sensor. Usually, such light cut-off is achieved via a mechanical shutter quickly rotating in front of the camera lens, providing a clear, unobstructed view only during the fraction of sensor exposure. Such systems can be successfully deployed at high frequencies (Gembicky et al., 2005). This mechanical shutter must be carefully synchronized with the trigger to only open when all sensor lines are integrated. If not appropriately synchronized, the top-line rows can become overexposed, reducing the system's immunity to ambient light. Figure 1 shows a regular mechanical shutter realized via a high-speed spinning disk with a hollow part.

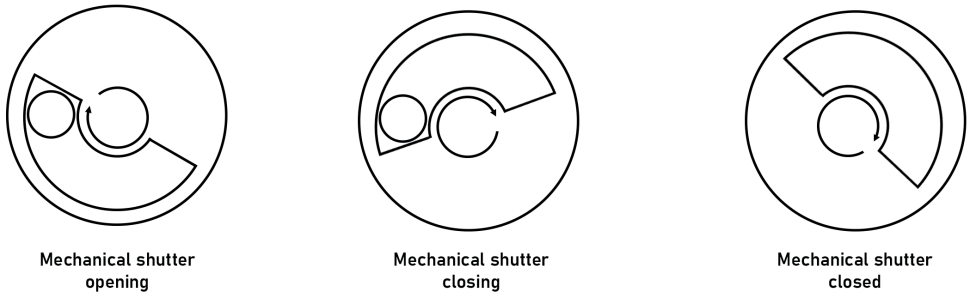

Fig. 1. Regular mechanical shutter realized via high-speed spinning disk with the hollow part

To further reduce manufacturing and maintenance costs and simplify the system, a special light-valve assembly was designed that utilized liquid crystals. It uses liquid crystals to vary the quantity of light passed from a source to a target. By applying voltage, the polarization of anisotropic crystals can be rotated to block the light. A light valve of a transmissive type light valve was tested in this application. A precise transparency control is possible by applying DC voltage of up to 5 V. Liquid crystal Light Valve can reduce transparency by up to 95%, effectively controlling the amount of light reaching the sensor. The light valve darkens at 1-volt DC, and at 4-volt DC, it achieves its maximum light-blocking level of about 95%. This technology requires very little current and has no mechanical moving parts, making it highly efficient, affordable, and reliable. Liquid crystals are very temperature-sensitive. There have already been successful attempts at adjusting liquid crystal haziness by heating the polymer-dispersed liquid crystal (Shin et al., 2022). In this case, heating the Liquid Crystal Light Valve (LCLV) is performed to increase the reaction time of the light valve. The light valve is too slow at room temperature to prevent ambient light from hitting the camera sensor. A comprehensive series of tests was conducted to measure the influence of liquid crystal temperature on the reaction time of the crystals. These tests were crucial to ensure that the liquid crystal light valve operates consistently and accurately under varying temperature conditions. The following measurements, shown in Figure 2, were performed to precisely measure how light valve reaction time changes with temperature. The results collected during the measurements are organized in Table 1, which shows the influence of temperature on the reaction time of a light valve.

The reaction time of liquid crystals drops at higher temperatures. This introduces one challenge: while in the mechanical shutter, controlling the rotational speed of the hollow disc was crucial to ensure proper timing; here, ensuring a constant temperature of liquid crystals was necessary to maintain the high reaction speed of the light valve. At the same time, the maximum operational temperature of liquid crystals is 60°C, meaning that the heating mechanism also requires a control system. To achieve this, a dedicated heater, shown in Figure 3, was designed to maintain the liquid crystals at the desired temperature.

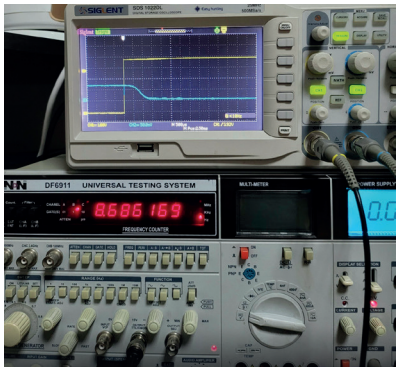

Fig. 2. Measuring reaction time of a light valve

Table 1. Influence of temperature on the reaction time of a light valve

Liquid crystal temperature [°C]	Time to react [ms]	Time to fully closed [ms]
−14	2.0	3.0
20	2.0	3.0
50	0.5	0.5

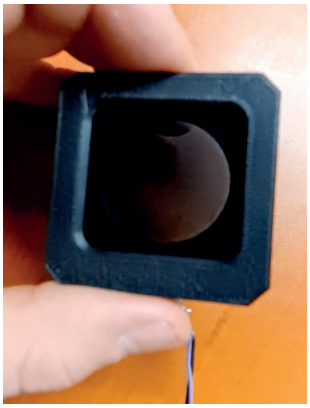

Fig. 3. Light valve surrounded by the heating module

The heater consists of a rectangular frame filled with resistors and equipped with a temperature control sensor. This setup ensures that the light valve consistently operates at 50°C, providing the fastest possible reaction time and significantly improving the system's efficiency in filtering out ambient light. By maintaining this controlled temperature, we can ensure the reliable and consistent performance of the liquid crystal light valve, which is crucial for precise and accurate data acquisition. The design of the heater was carefully considered to avoid any potential overheating, which could damage the light valve. Light valve enabled control over sensor exposure cut-off time by effectively blocking ambient light from hitting the sensor. This resulted in a constant image brightness over the entire image captured by a rolling shutter sensor in a global reset shutter mode. Synchronization in a photography environment can be a very complex task. Some synchronization systems can maintain a sync precision of under 0.5 ms over five minutes of constant video stream (Šmíd & Matas, 2019). In this particular application, only one trigger event is synced to capture an image, making the challenge easier from a synchronization perspective. Afterward, careful synchronization between the light valve, camera shutter, and strobe light was performed to ensure that all systems worked in tandem, delivering optimal results. The light valve was constantly heated and maintained an optimal temperature of 50°C. The light valve was triggered at 2 ms to reach maximum darkness at the 2.5-millisecond mark. The LED strobe light in an over-current light burst lights up 0.5 ms and can only remain lit at full power for about 2 ms before it overheats. The camera shutter speed was set to 4 ms to ensure that the top lines remained exposed for the entire duration of the light burst. For the remaining 2 ms, the camera sensor is cut off from ambient light by a light valve to ensure that the bottom lines of the camera sensor are not collecting ambient light. The schematic of the triggering sequence is detailed in Table 2. The triggering sequence is illustrated in 0.5-millisecond increments.

This precise timing allows the system to effectively minimize ambient light interference. The synchronization ensures that the camera captures the most precise possible images during the brief period when the strobe light illuminates the car's surface. Considering that full daylight can exceed 10,000 lx and direct sunlight reflections can surpass 100,000 lx, the camera's shutter speed, gain, and lens aperture were carefully configured to create a light intensity filter. This filter ensures that only light brighter than 25,000 lx is captured by the sensor. As the camera used has a CMOS sensor type, which is prone to high noise (RadhaKrishna et al., 2021), the camera was set up to generate the most minor possible digital noise, so the digital gain was set at 0 dB. The aperture of the camera lens was set at over f/8.0 to ensure a depth of field of at least 50 cm, which will compensate for car driving errors and mitigate the risk of a car going out of focus. This enabled the computer vision system to not register ambient light with an intensity of less than 40,000 lx. Customized strobe light source could light up to 2 ms with a brightness of 40,000 lx, significantly surpassing the average ambient light intensity levels. Figure 4 presents four different light sources illuminating stickers corresponding to their light intensities of 10,000 lx; 16,000 lx; 25,000 lx; and 40,000 lx. Camera parameters and lens parameters were set up so that only 40,000 lx light source was overexposed.

Fig. 4. Vision system configured in a way that it blocks light under 40,000 lx

Table 2. Triggering sequence illustrated in 0.5-millisecond increments

	0 ms	0.5 ms	1 ms	1.5 ms	2 ms	2.5 ms	3 ms	3.5 ms	4 ms	4.5 ms
Light valve triggered										
Light valve fully on										
Strobe light triggered										
Strobe light fully on										
Camera sensor										

4. Generating the dataset

The hood in Figure 5 is illuminated with a 25,000 lx flood light source. It is a spotlight source mimicking sunlight reflection. Figure 5 shows an image captured at an automated camera setting, representing how it looks with the unaided human eye.

Fig. 5. Test sample photographed with a regular camera in auto mode

Flash Light setup was constructed using high-intensity LED stripes that were triggered using overvoltage of 200% over their regular rated voltage for a concise timeframe of just 2 ms. This enabled the LEDs to produce light as bright as 40,000 lx without burning out. It was another cost-measure to make the system as scalable as possible. LED lights were placed 0.05 m apart in a 1 m × 2.5 m surface covered with LEDs. The sample shown in Figure 6 represents the same hood with 25,000 lx flood light disturbance, but this time with camera and lens parameters adequately adjusted and in sync with the light trigger.

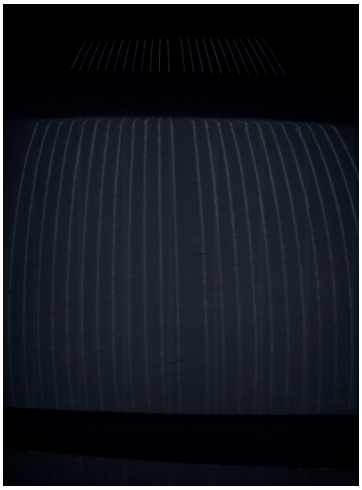

Fig. 6. Test sample photographed with vision system

As one can see, dents are clearly visible in such an image despite the strong 25,000 lx ambient light reflection bouncing off the car's body. Once the data acquisition setup is completed, the training data set generation process can be established. Light is subject to the inverse-square law. According to this law, when the distance between a light emitter and the target is doubled, the observed light intensity decreases by a factor of four. If the light source at 0.3 m distance delivers 40,000 lx to the car's body, at 0.6 m distance, only 20,000 lx will be delivered. That is why it is critical to maintain a consistent light source to car body distance. It is also crucial to have the light as close to the car's body as possible to ensure that enough light intensity is delivered. A simple wheeled trolley was locked in a rail attached to the floor to provide consistent light-to-car body distance. Structured light projection effectively increased visual dent extraction from the car's body, as shown in Figure 7.

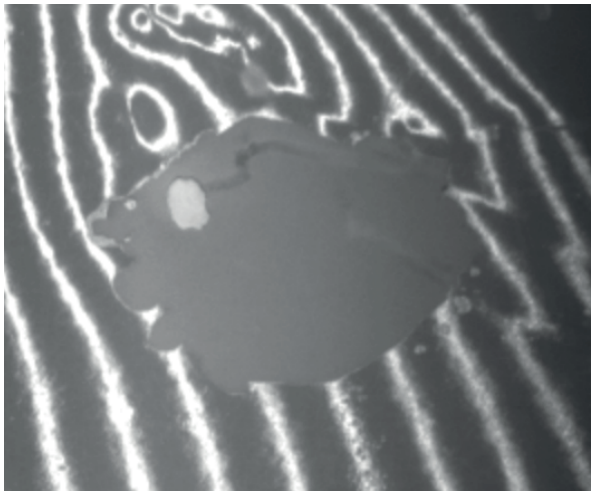

Fig. 7. Sample dent with paint chip recorded with computer vision setup

Fifteen cars with various dent sizes and locations were scanned using the system. It enabled the capturing of over 1,193 different images. With the data acquisition process complete, the next step was to train the algorithms and for this a proper training dataset was essential.

5. Labelling the data

A challenge in labeling the data is ensuring the dataset is 100% accurate and follows ground truth about the car's condition. According to industry standards, the unaided human eye detects, on average, only 55% of car body damage.

The detection rate for dents varies based on factors such as the dent's location, size, and ambient lighting during inspection. Experimental validation was conducted on multiple hail-damaged vehicles to investigate this, examining dents across 37 distinct body elements identified for this study. These vehicles had a total of 434 dents distributed across body panels such as the left A-pillar, grille/front panel, hood, roof, etc.

Findings indicate that larger dents (over 25 mm in diameter) located at body panels near the eye-line are significantly more straightforward to detect, with a 100% detection rate by the unaided human eye. In contrast, more minor dents (under 15 mm in diameter) at more challenging locations, such as the intersection of the hood and grille, showed only around a 25% detection rate. The most difficult dents to detect were those under 10 mm in diameter located at the center of the roof, as inspecting these dents required a ladder, which altered body reflections and made detection more challenging. These small roof dents were identified by the unaided eye at only about a 15% detection rate. Overall, our analysis shows that the average human detection rate for car dents is approximately 55%, aligning with the industry average.

This detection rate drops significantly after two hours of inspecting cars due to eye fatigue. This is why each person doing data labeling could not work for more than two hours a day on labeling the dataset. As one car inspector has a 55% detection rate, the inspector has a 45% chance of missing dent. The concept of independent probabilities is used to find the number of inspectors needed to ensure that all damages are detected. The goal is to have a combined detection rate as close to 100%.

At least six inspectors are required to achieve a combined detection rate of 99%. A team of six specialists was formed for data labeling. Each member had at least two years of experience in visual data labeling in the automotive field. They were chosen after completing a visual test challenge to check their attention to detail and accuracy. From 10 participants, the top six were selected. As each picture in the dataset was labeled by six people independently, the data set is as close to the ground truth as feasible. During the labeling stage, it was crucial to label only actual dents and not the stylish curves of the car's silhouette. Of 1,193 images, 70% were used for training and 30% for testing. During the data labeling stage, specialists needed to review 835 images and mark all visible dents. Dents could be seen because they changed the shape of the straight light stripes reflected on the car's body. When a dent was found, they had to draw a polygon over the dent area. Only dents that appeared in the strobe light area were marked. Dents that did not show up in the structured light were skipped since they would be covered by stripes in another image. To prevent eye fatigue, each specialist could work on data labeling for up to two hours a day. For data labeling, an open-source platform shown in Figure 8 was used.

The training machine had the following parameters: Intel Core i9-13900HK, equipped with 32 GB, LPDDR5 memory, and NVIDIA GeForce RTX 4080 with 16 GB of video memory.

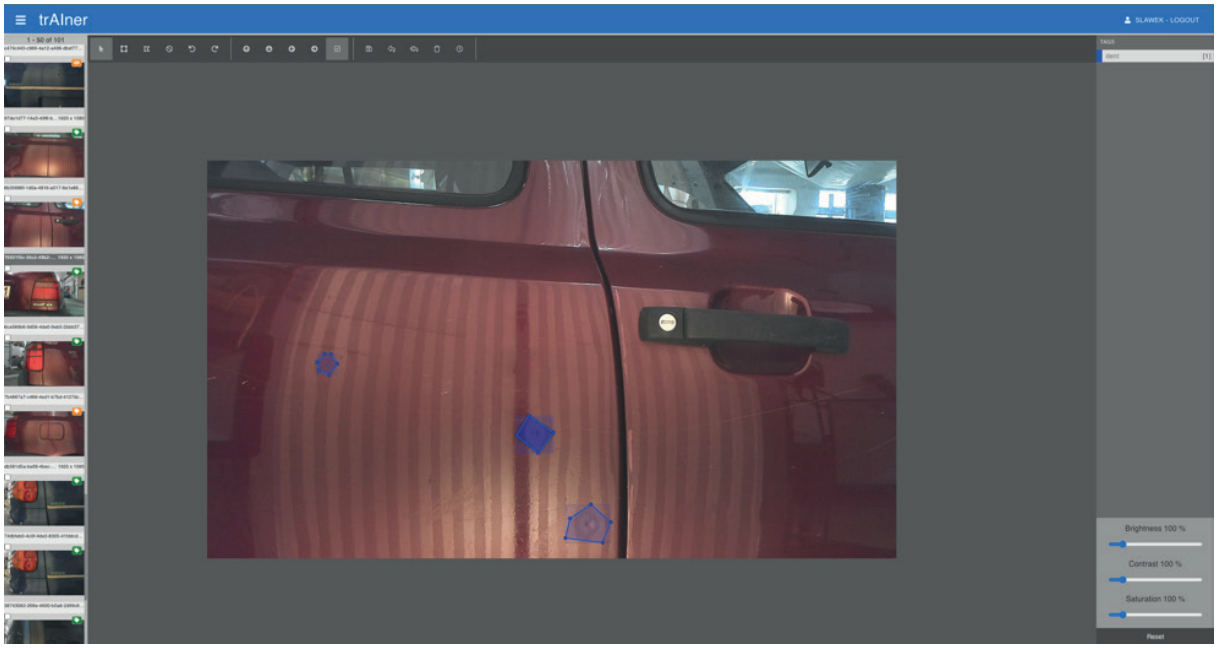

Fig. 8. Data labeling platform

6. Results and discussion

There are many neural networks for object detection in images, for example: Faster R-CNN (Ren et al., 2017), YOLO, FCOS (Tian et al., 2019), CenterNet, SSD (Wei et al., 2016). The YoloR algorithm (Wang Ch.-Y. et al., 2021) was chosen as the Artificial Neural Network for this study. For detecting dents in car body panels using AI, YoloR (Redmon et al., 2016) was selected as the primary algorithm due to its ease of use and strong performance on visual detection tasks. YoloR demonstrates competitive accuracy and speed, particularly evident in the COCO dataset, a widely recognized benchmark in object detection. This benchmark's diverse and challenging scenarios validate YoloR's capability to accurately detect small and irregular shapes, such as dents, which are crucial for assessing vehicle surface quality. Thus, YoloR's adaptability and effectiveness make it well-suited for dent detection in car body panels. The algorithm was needed for detection purposes, not instance segmentation or classification.

YoloR is a modern network developed in 2021, part of many networks in the You Only Learn Once (YOLO) collection of algorithms. It is based on Scaled Yolov4 (Wang Ch.-Y. et al., 2020) architecture but introduces a few improvements. One of the upgrades is unified representation. One of these advancements is introducing a unified representation, designed not solely for a single task, such as detection or segmentation, but is generalized to apply simultaneously to multiple tasks that a network can perform. It enables the network to perform detection, segmentation, or instance segmentation tasks. Due to improved generalization, the network also achieves better performance on individual tasks, such as detection. Data augmentation was applied because cars with hail damage are not easily obtainable. Data augmentation is an effective way of improving the scarce dataset that was captured (Mumuni et al., 2022). The dataset was then further augmented using two methods. The dataset was first augmented three times by adding random rotations (between −25° and +25°, horizontal and vertical flip transformations) and adjusting the black point of images. Then, the dataset was augmented using mosaic augmentation and built into the YoloR algorithm. Its principle of operation is to take four images; these four images are then cropped, resized, and merged into one mosaic-like image. This further increased the data size by twice the size. A sample mosaic-like image is shown in Figure 9.

The YOLO-R algorithm in the P6 version was trained for 100 epochs utilizing an NVIDIA GeForce RTX 4080 GPU. The remaining training parameters were retained at their default YOLO-R settings, including the learning rate, momentum, and weight decay. Transfer learning was employed during training, with initial weight values derived from the pre-trained YOLO-R model on the COCO dataset (https://cocodataset.org/, Microsoft COCO: Common Objects in Context). The training dataset comprised collected and augmented images, resized to a resolution of 1,024 × 1,024 pixels.

The *AP* (Average Precision) measure (1) is used to compare the results:

$$\mathrm{AP} = \int_0^1 \rho(r)\,dr \tag{1}$$

where: $\rho(r)$ is the precision in the function of recall, this metric is a standard measure of detection effectiveness on images used in datasets such as COCO (Lin et al., 2014). Figure 10 shows precision during training for 100 epochs.

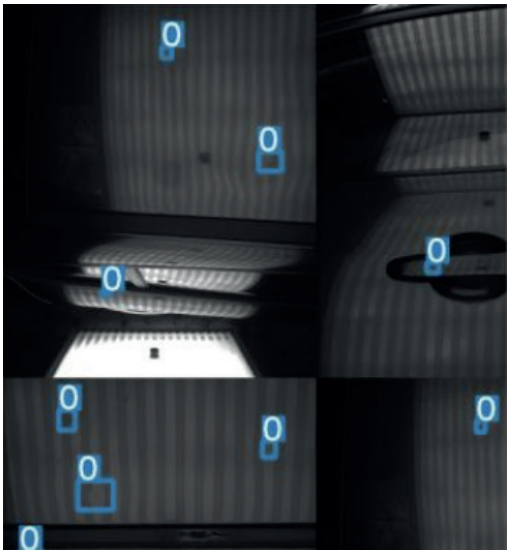

Fig. 9. Sample mosaic augmentation

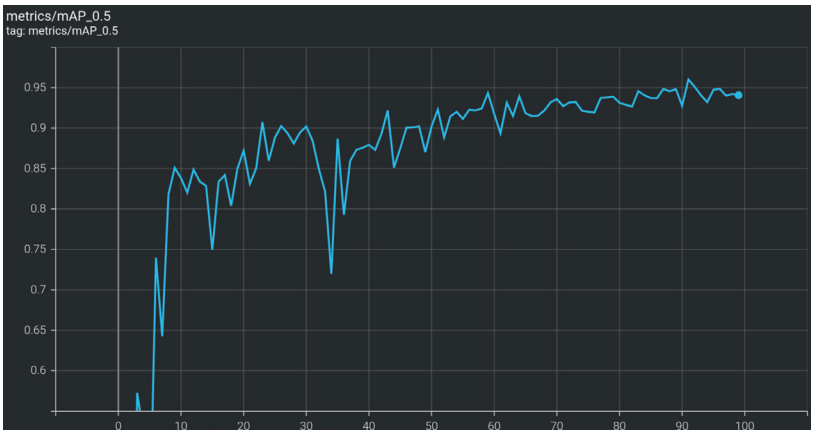

Fig. 10. Precision during training for 100 epochs

Upon completing the training process, the network achieved a detection precision rate of *AP* 50 = 95.2%. Figure 11 shows an example detection indicating the presence of a dent with a probability of 0.8.

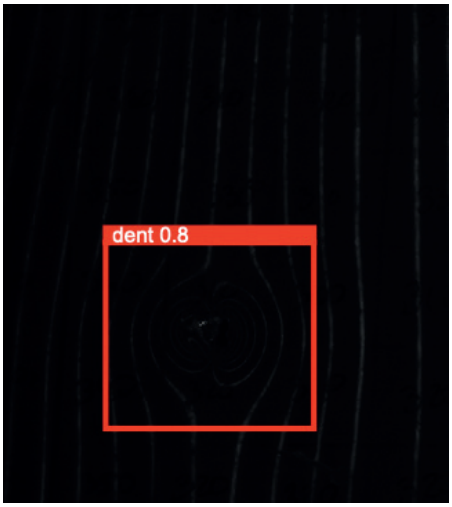

Fig. 11. Sample detection

7. Conclusion

This study delivered a comprehensive system capable of registering and detecting car body dents even in harsh ambient light conditions matching real-life operations at a car logistics plant. Ambient light interfering with data capturing presents a significant challenge for the structured light projection method in real-life scenarios of the automotive industry. Sunlight generates very bright reflections, especially on metallic car surfaces and dark car paint colors, making it crucial to overcome these reflections and project clear patterns onto the car's body. Given that light follows the inverse-square law, a light source that is much weaker than sunlight but positioning it very close to the target object can still overcome sunlight reflections. This is particularly effective when combined with a proper camera and lens setup. The proposed computer vision system addresses this challenge by significantly blocking ambient light from getting registered via the camera sensor thanks to a mechanical shutter composed of a light valve, overvoltage-driven strobe light, and exact synchronization with the camera sensor. Moreover, the setup was cost-optimized to work with a much cheaper rolling shutter camera operating in global reset shutter mode. By optimizing the proximity and intensity of the structured light source, it is possible to preserve high visibility projected stripe patterns on the car's body. This approach effectively filters out the interference caused by sunlight reflections, allowing for the accurate detection of defects such as dents, scratches, peeling paint, and irregularities in paint thickness. As a result, the system achieves a 95.2% detection accuracy despite direct sunlight exposure. This result was achieved on a testing dataset that was not used for algorithm training. Although the results of this study deliver very good results, there is still significant room for improvement, especially in developing detections for other body damage types.

References

Arnal, L., Solanes, J. E., Molina, J., & Tornero, J. (2017). Detecting dings and dents on specular car body surfaces based on optical flow. *Journal of Manufacturing Systems*, *45*, 306–321. http://doi.org/10.1016/j.jmsy.2017.07.006.

Chen, H. (2008). *Automatic Dent Detection on Car Bodies* [Master thesis, Hong Kong University of Science and Technology]. Rare & Special e-Zone. https://doi.org/10.14711/thesis-b1023250.

Gembicky, M., Oss, D., Fuchs R., & Coppens, P. (2005). A fast mechanical shutter for submicrosecond time-resolved synchrotron experiments. *Journal of Synchrotron Radiation*, *12*, 665–669. https://doi.org/10.1107/S090904950501770X.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common objects in context. In D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014. 13th European Conference, Zurich, Switzerland, September 6–12, 2014, Proceedings* (pt. 5, pp. 740–755). Springer Cham. https://doi.org/10.1007/978-3-319-10602-1_48.

Mumuni, A., & Mumuni, F. (2022). Data augmentation: a comprehensive survey of modern approaches. *Array*, *16*, 100258. https://doi.org/10.1016/j.array.2022.100258.

Park, S. H., Tjolleng, A., Chang, J., Cha, M., Park J., & Jung, K. (2020). Detecting and localizing dents on vehicle bodies using region-based convolutional neural network. *Applied Sciences*, *10*(4), 1250. https://doi.org/10.3390/app10041250.

Patil, K., Kulkarni, M., Sriraman, A., & Karande, S. (2017). Deep learning based car damage classification. *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. https://doi.org/10.1109/ICMLA.2017.0-179.

Pham, L. H., Tran, D. N.-N., Byun, J. Y., Rhie, Ch. H., & Jeon, J. W. (2021). A smartphone-based laser measuring system for gap and flush assessment in car body. *IEEE Transactions on Industrial Electronics*, *68*(7), 6297–6307. https://doi.org/10.1109/TIE.2020.2992971.

RadhaKrishna, M. V. V., Govindh, M. V., & Veni, P. K. (2021). Review on image processing sensor. *Journal of Physics: Conference Series*, *1714*, 012055. https://doi.org/10.1088/1742-6596/1714/1/012055.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). YOLO You Only Look Once: unified, real-time object detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/CVPR.2016.91.

Ren, S., He, H., Girshick, R., & Sun, J. (2017). Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031.

Shin, H., Ryu, J. W., Gwag, J. S., Lee, H.-M., Hong, J.-Y., Kim, S.-W., Lee, Y.-S., & Kim, J. M. (2022). Temperature-controllable liquid crystal window shutter. *Journal of the Korean Physical Society*, *80*, 273–278. https://doi.org/10.1007/s40042-022-00422-0.

Šmíd, M., & Matas, J. (2019). *Rolling shutter camera synchronization with sub-millisecond accuracy*. arXiv:1902.11084, https://doi.org/10.48550/arXiv.1902.11084.

Tian, Z., Shen, C., Chen, H., & He, T. (2019). FCOS: fully convolutional one-stage object detection. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. https://doi.org/10.1109/ICCV.2019.00972.

Van der Jeught, S., & Dirckx, J. J. (2019). Deep neural networks for single shot structured light profilometry. *Optics Express*, *27*(12), 17091–17101. https://doi.org/10.1364/OE.27.017091.

Wang, Ch.-Y., Bochkovskiy, A., & Liao, H.-Y. M. (2020). Scaled-YOLOv4: scaling cross stage partial network. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/CVPR46437.2021.01283.

Wang, Ch.-Y., Yeh, I.-H. & Liao, H.-Y. (2021). *You Only Learn One Representation: Unified Network for Multiple Tasks*. arXiv: 2105.04206. https://doi.org/10.48550/arXiv.2105.04206.

Wang, Y., Zhu, D., Yang, Y., Lee, K., Mishra, R., Go, G., Oh, S.-H., Kim, D.-H., Cai, K., Liu, E., Pollard, S. D., Shi, S., Lee, J., Teo, K. L., Wu, Y., Lee, K.-J., & Yang, H. (2019). Magnetization switching by magnon-mediated spin torque through an antiferromagnetic insulator. *Science*, *366*(6469), 1125–1128. https://doi.org/10.1126/science.aav8076.

Wei, L., Anguelov, D., Erhan, D., Szegedy, Ch., Reed, S., Fu, Ch.-Y., & Berg, A. C. (2016). SSD: single shot multibox detector. In B. Leibe, J. Matas, N. Sebe, M. Welling (Eds.), *Computer Vision – ECCV 2016. 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings* (pt. 1, pp. 21–37). Springer Cham. https://doi.org/10.1007/978-3-319-46448-0_2.