

## Powering AI-driven car damage identification based on VeHIDE dataset

Van-Dung Hoang, Nhan T. Huynh, Nguyen Tran, Khanh Le, Thi-Minh-Chau Le, Ali Selamat & Hien D. Nguyen

To cite this article: Van-Dung Hoang, Nhan T. Huynh, Nguyen Tran, Khanh Le, Thi-Minh-Chau Le, Ali Selamat & Hien D. Nguyen (25 Jun 2024): Powering AI-driven car damage identification based on VeHIDE dataset, Journal of Information and Telecommunication, DOI: [10.1080/24751839.2024.2367387](https://doi.org/10.1080/24751839.2024.2367387)

To link to this article: <https://doi.org/10.1080/24751839.2024.2367387>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 25 Jun 2024.



Submit your article to this journal [↗](#)



Article views: 1146



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

# Powering AI-driven car damage identification based on VeHIDE dataset

Van-Dung Hoang<sup>a</sup>, Nhan T. Huynh<sup>b,c</sup>, Nguyen Tran<sup>d</sup>, Khanh Le<sup>b,c</sup>, Thi-Minh-Chau Le<sup>a</sup>, Ali Selamat<sup>e,f</sup> and Hien D. Nguyen<sup>b,c</sup>

<sup>a</sup>HCMC University of Technology and Education, Ho Chi Minh city, Vietnam; <sup>b</sup>University of Information Technology, Ho Chi Minh city, Vietnam; <sup>c</sup>Vietnam National University, Ho Chi Minh city, Vietnam; <sup>d</sup>AISIA Research Lab, Ho Chi Minh city, Vietnam; <sup>e</sup>Malaysia-Japan International Institute of Technology (MJIT), Universiti Teknologi Malaysia Kuala Lumpur, Kuala Lumpur, Malaysia; <sup>f</sup>Faculty of Computing, Universiti Teknologi Malaysia & Media and Games Center of Excellence (MagicX), Universiti Teknologi Malaysia, Johor Bahru, Malaysia

## ABSTRACT

In the realm of automobile insurance, the imperative of automating car damage evaluation has surged, offering streamlined assessment processes and heightened accuracy. Deep learning techniques have notably influenced vehicle damage assessment, reshaping insurance procedures. However, the primary challenge remains in crafting robust models for damage detection and segmentation. This study presents a novel contribution through the development of the Vehicle Damage Detection (VehiDE) dataset, specifically tailored for comprehensive car damage assessment. The dataset, encompassing 13,945 high-resolution images annotated across eight damage categories, serves as a foundational resource for advancing automated damage identification methodologies. Notably, VehiDE offers support for multiple tasks, including classification, object detection, instance segmentation, and salient object detection, thereby fostering versatile research avenues. Through extensive experimental analysis, including the evaluation of state-of-the-art methodologies on VehiDE, this study not only highlights the dataset's efficacy but also unveils new insights into the challenging nature of car damage assessment. Moreover, the study pioneers the exploration of salient object detection techniques in this domain, showcasing their potential in addressing irregular damage types. By offering VehiDE to the research community, we aim to catalyze advancements in the field of car damage assessment, paving the way for more accurate and efficient automated systems.

## ARTICLE HISTORY



Received 18 January 2024  
Accepted 8 June 2024

## KEYWORDS

Automobile impairment;  
deep learning; CNN;  
instances segmentation;  
vehicle damage

## 1. Introduction

In the realm of automobile insurance, the automated evaluation of car damage has gained increasing importance. This technology not only simplifies the assessment of

**CONTACT** Hien D. Nguyen  [hiennd@uit.edu.vn](mailto:hiennd@uit.edu.vn)  University of Information Technology, Ho Chi Minh city, Vietnam; Vietnam National University, Ho Chi Minh city, Vietnam

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

damages but also enhances its accuracy, diminishing the need for extensive manual inspections. In recent times, there has been a notable upswing in the incorporation of deep learning techniques within the domain of vehicle damage assessment, transforming the way insurance firms handle claims and compensations.

The primary obstacle in this domain revolves around creating robust models for detecting and segmenting damages (Pham & Hoang, 2024). Effective solutions to this challenge have emerged through the utilization of deep neural networks (Nguyen et al., 2022; Nguyen et al., 2023a). Different models based on neural networks have played a crucial role in identifying various damage categories, while alternative architectures have demonstrated their significance in tasks such as object detection and instance segmentation (Xiang et al., 2024).

When vehicles encounter sharp objects, are involved in accidents, or undergo other incidents affecting their structure, they may incur various forms of damage altering their original shape and conditions (Medrano-Berumen & İlhan Akbaş, 2021). The increasing utilization of automated systems for detecting automotive damage is evident, with many consumers relying on these state-of-the-art technologies (Hoang et al., 2023; Huynh et al., 2022; Nguyen et al., 2023b). The prevalence of AI-driven vehicle damage detection is expanding, and its effectiveness is continually improving. Numerous approaches for detecting damage in automobiles are now accessible.

The development of automatic technologies for assessing automotive damage has shown to be very advantageous because it requires less manual labor and considerably increases the effectiveness of damage inspection. In recent years, the vehicle insurance business has seen a significant uptick in the integration of deep learning in car damage assessment (Zhang et al., 2020). When processing routine claims, insurers can save a lot of time, money, and manpower by automatically identifying small exterior damages like broken glass or broken glass. To correctly locate and categorize the aforementioned sorts of damage, considerable emphasis is being paid to research that focuses on damage assessment for vehicles.

Car damage assessment is closely intertwined with the task of visual detection and segmentation, which can be effectively addressed using deep neural networks, showcasing their impressive potential (He et al., 2017; Redmon et al., 2016). Convolutional neural network (CNN)-based models have been applied for damage category identification, while YOLO (Redmon et al., 2016) and Mask R-CNN (He et al., 2017) have been frequently referenced for object detection and instance segmentation tasks, respectively.

The research on vehicle damage assessment has been hindered by insufficient datasets for training such models. Although a publicly available dataset exists in Pham et al. (2023), it only handles classifying damage types, but car assessment needs both detection and segmentation. Existing works often rely on private datasets (Balci et al., 2019; de Deijn, 2018; Patil et al., 2017), making it difficult to fairly compare the performance of different approaches, further impeding the field's development.

To facilitate research in vehicle damage assessment, this paper introduces a novel dataset, called *VehiDE for Car Damage Detection*. The building dataset presents various challenges in car damage detection and segmentation (Waqas et al., 2020). It is the first publicly available dataset with the following features:

- *Damage class*: Broken glass, broken lights, scratch, lost parts, dents, torn, punctured, non-damaged.
- *Dataset quantity*: It contains over 13,945 high-resolution car damage images with their labelled instances of eight damage categories.
- *High-resolution images*: VehiDE offers significantly higher image quality compared to other datasets. The average resolution of VehiDE images is an impressive 684,231 pixels. This high-resolution feature enhances the dataset's ability to capture fine details and nuances in car damage.
- *Multiple task support*: VehiDE is designed to accommodate four distinct tasks, making it a versatile resource for researchers. The dataset supports classification, object detection, instance segmentation, and salient object detection (SOD). This multi-task support allows researchers to explore various aspects of car damage assessment using a single comprehensive dataset.
- *Fine-grained differentiation*: Unlike generic object detection and segmentation tasks, VehiDE focuses on the fine-grained differentiation of car damage types, such as distinguishing between dents and scratches. This level of detail is crucial for developing precise and accurate car damage assessment algorithms.
- *Diversity in object scales and shapes*: The images in VehiDE exhibit a greater diversity of object scales and shapes compared to general object datasets. This diversity arises naturally from the nature of car damage, where damages can occur on various parts of the vehicle with distinct shapes and sizes. This aspect challenges researchers to handle a wide range of car damage scenarios.

The VehiDE dataset stands out due to its high-resolution images, support for multiple tasks, emphasis on fine-grained differentiation, and the inherent diversity in object scales and shapes. These features make VehiDE a valuable resource for advancing research in the field of car damage assessment and computer vision. This study is an extended version of the dataset which was presented in Huynh et al. (2023) about enlarging the original dataset and doing more experiments on the state-of-the-art methodologies to prove the efficiency of the built dataset and the proposed method.

There are main signification contributions in this paper:

- *VehiDE dataset*: a publicly available dataset designed for car damage detection and segmentation. It not only facilitates the training and evaluation of algorithms for car damage assessment but also presents new opportunities to advance the field by providing comprehensive data for this specific task.
- *Comprehensive annotations*: VehiDE includes high-quality images annotated with detailed information, such as the type, location, and magnitude of the damage. These annotations offer practical and precise data, surpassing the existing datasets in terms of comprehensiveness.
- *Extensive experimental analysis*: We conduct in-depth exploratory experiments and thoroughly analyze the results, providing valuable insights to fellow researchers. The evaluation of state-of-the-art (SOTA) methods on VehiDE reveals the challenging nature of car damage assessment, thereby presenting new avenues for object detection and instance segmentation tasks. Additionally, we pioneered the utilization of SOD (Salient Object Detection) methods for this task, and our experimental outcomes

demonstrate the potential of SOD techniques, especially in dealing with irregular damage types. By sharing VehiDE, we aim to contribute to the vision-based car damage assessment domain and foster progress within the computer vision community.

## 2. Related work

The application of deep learning in the realm of intact vehicles has shown substantial progress, particularly in tasks like precise vehicle recognition (Xiang et al., 2020; Zhang et al., 2022), vehicle type and component classification (Huang et al., 2020; Jeon et al., 2022), and license plate recognition (Al-Shemarry et al., 2020; Zhang et al., 2021). On the basis of the time series, the investigation in Meißner and Rieck (2021) focuses on the interplay between accident characteristics and location and time. However, the recognition and evaluation of damaged vehicles and their components present unique challenges, with existing models trained on undamaged vehicle datasets proving ineffective for this specific task (Zhu et al., 2021). Additionally, the unconventional nature of damages, which are not easily classified as traditional ‘objects,’ adds complexity to the problem (Parhizkar & Amirfakhrian, 2022; Tran et al., 2022).

In response to these challenges, Scene Academy has committed to developing innovative approaches tailored to the recognition and assessment of damaged vehicles and their components. Through the customization of deep learning techniques for damaged scenes, the objective is to enhance the accuracy and efficacy of algorithms in this specialized domain.

### 2.1. Related methods for vehicle damage field

Early approaches to car damage detection relied on heuristics, such as comparing 3D CAD car models with photographs of damaged cars (Reddy et al., 2022). However, the limitations of these methods in real-world scenarios led to the adoption of deep learning techniques. Recent endeavors have employed CNNs to classify various types of car damage (Mahavir et al., 2021), but challenges remain, especially in handling instances with multiple types of damage and providing detailed information about the location and severity of the damage.

Advanced techniques, including YOLOv3 and Mask R-CNN, have been applied to detect damaged regions and localize various classes of car components and damages (Li et al., 2018; Singh et al., 2019). Despite progress, the difficulty of the car damage detection task is evident in the relatively low scores on metrics like mean Average Precision (mAP). Interestingly, there has been limited exploration of applying techniques from various domains, including salient object detection, to tackle the difficulties in identifying car damage.

Commercial applications, like the car damage assessment system from Ant Financial Services Group (Zhang et al., 2020), underscore the practical significance of incorporating object detection and instance segmentation techniques to accurately identify damaged components and assess their severity.

### 2.2. Related dataset

Table 1 summarized existing datasets for car damage assessment, exhibiting limitations in terms of size, task focus, and limited public availability. Notably, the majority of these

**Table 1.** Datasets for classification, detection, and segmentation of the vehicle damage.

Dataset	Task	Size	PA	Categories	Damage Categories
de Deijn (2018)	C	1,007	N	4	dent, glass, scratch, and hail
Balci et al. (2019)	C	533	N	2	damaged, and non-damaged
Waqas et al. (2020)	C	600	N	3	medium damaged, huge damage, and no damaged
Patil et al. (2017)	C	1,503	N	7	bumper dent, door dent, glass shatter, scratch, tail lamp broken, scratch, smash, and headlamp broken
Li et al. (2018)	D	1,790	N	3	crack, scratch, and dent
Patel et al. (2020)	D	326	N	N/A	N/A
Dhieb et al. (2019)	D, S	N/A	N	3	minor, moderate, and major
Singh et al. (2019)	S	2,822	N	5	scratch, minor dent, major dent, missing, and cracked
Wang et al. (2022)	C, D, S, SOD	4,000	Y	6	scratch, dent, glass shatter, crack, tire flat, and lamp broken
VehiDE	C, D, S, SOD	13945	Y	8	dents, broken glass, scratch, lost parts, punctured, torn, broken, lights, non-damage

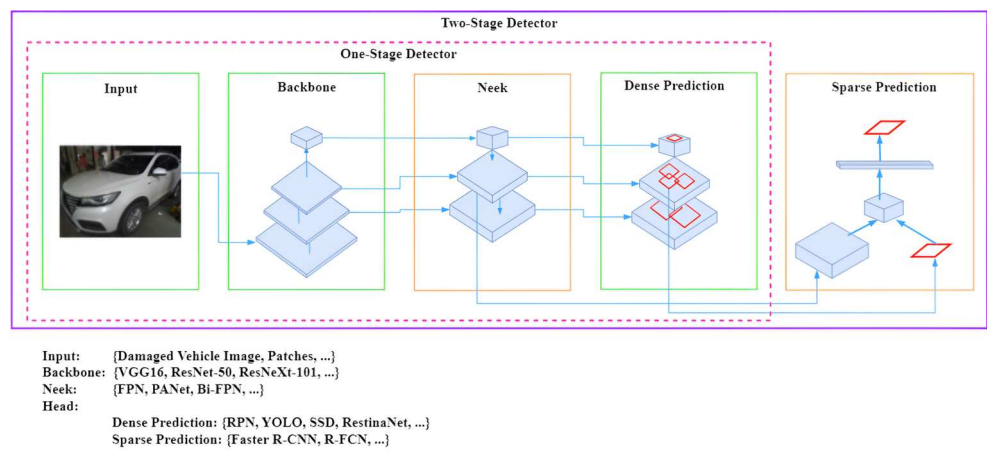
Note: The letters C, D, S, and SOD encompass damage classification, detection, segmentation, and salient object detection, resp. N/A: The authors omitted to include that information.

datasets are relatively small, hindering a comprehensive evaluation of deep learning techniques. Additionally, the focus on damage classification in many datasets may not fully address practical scenarios, prompting a shift towards more comprehensive tasks like damage detection, segmentation, and salient object detection.

The VehiDE dataset stands out as a comprehensive solution, offering 13,945 meticulously annotated high-resolution samples for various tasks, including damage classification, detection, segmentation, and salient object detection. Released specifically for academic research purposes, the dataset aims to propel advancements in the field of car damage assessment and facilitate fair comparisons between methodologies. The paper at hand explores novel ways to apply techniques on the VehiDE dataset and subsequently compares their effectiveness in addressing the challenges of damaged vehicle recognition and assessment.

### 3. Vehicle damage prediction based on deep convolutional neural network approach

The data is quite noisy, making it quite difficult to detect physical damage. The inherent noise within the dataset presents a significant hurdle in discerning physical damage on the vehicle accurately. This challenge, if not adequately addressed, can adversely affect the overall performance and functionality of the system. In response to this issue, we have implemented vehicle damage recognition based on a comprehensive strategy leveraging the Object Detection YoLoV5 (Jocher, 2020) approach, as illustrated in Figure 1. This advanced model is specifically designed to identify objects belonging to the car class, allowing for the extraction of overarching features from the diverse and noisy dataset. Similar to the original approach, the feature backbone block constructed based on the CNN models such as VGG, ResNet, Darknet and so on; the Neck block is built by using FPN, PANet, Bi-FPN architecture; and the head block can be implement by using SSD, Faster-RCNN, Yolo approach. Consequently, this sophisticated approach plays a crucial



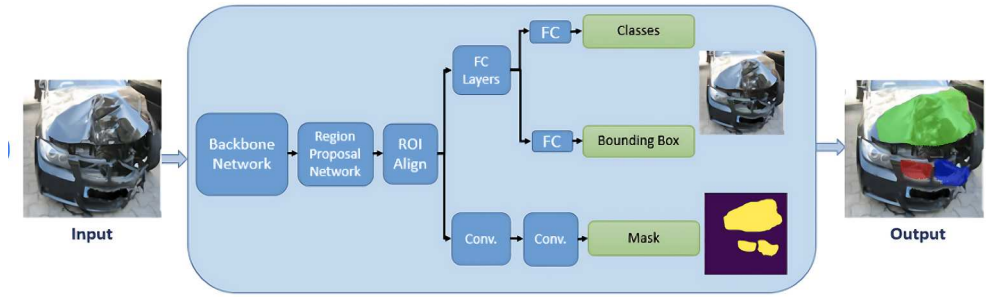
**Figure 1.** Architecture overview of the detection system based on YoLo approach.

role in not only detecting but also segmenting and pinpointing instances of physical damage on the targeted vehicle, contributing to a more nuanced and effective analysis.

In terms of vehicle damage detection, our approach involves the utilization of the ImageNet dataset and leveraging the pretraining capabilities of YoloV5 (Jocher, 2020). Following this initial step, we employ a meticulous process of layer filtering, wherein we selectively retain only the Car class. This targeted filtering is instrumental in facilitating object manipulation for subsequent analysis. Importantly, the background information is retained during this stage, providing the necessary context for comprehensive post-processing activities. This methodology not only streamlines the analysis of the Car class but also ensures the preservation of contextual elements for more robust post-processing techniques.

In another approach, aimed at achieving precise and comprehensive damage assessment on vehicles, are built upon these methodologies. The instance segmentation for object detection based on the Mask R-CNN approach is constructed. The overall of architecture of the Mask R-CNN is demonstrated in Figure 2.

Based on YOLOv5 model, this research embarks on an exhaustive exploration, engaging in a comparative analysis that juxtaposes the efficacy of it against alternative object detection models. The study unfolds as an intricate examination, delving into the nuanced intricacies of each model's predictive capabilities and generalization



**Figure 2.** Mask R-CNN model architecture.



across various object categories. This comprehensive approach aims to discern the models' adaptability and robustness in handling complex real-world scenarios.

The results garnered from this meticulous analysis illuminate YOLOv5's exceptional capabilities, unveiling its superior performance relative to its counterparts in the experimental setup (Figure ??). The model consistently achieves an impressive average mean Average Precision (mAP) score of 50.4%, underscoring its heightened accuracy and reliability in detecting objects across diverse categories within the COCO dataset.

These compelling findings not only affirm YOLOv5's standing as a high-performing model in the domain of object detection but also accentuate its adaptability to diverse datasets and scenarios. As a result, it emerges as a stalwart contender, offering a robust solution for applications that demand precise and efficient object recognition in a multitude of real-world contexts.

The binary classifier significantly streamlines the initial data-cleansing process. Its high recall score ensures the successful identification of a substantial percentage of images containing cars, allowing us to concentrate on pinpointing damaged cars from a comparatively smaller set of anticipated positives. This efficient approach accelerates the process of locating and isolating damaged vehicles in the dataset.

## 4. Experiments and evaluation results

The primary objective of automated automotive damage assessment is to accurately recognize and classify damages on vehicles, including visualizing their precise locations. The goals of instance segmentation and object detection align naturally with this objective. Hence, our initial focus involves exploring cutting-edge methods, specifically designed for tasks such as segmentation and object detection.

### 4.1. Experimental dataset

#### 4.1.1. Overview of VehiDE dataset

This segment introduces the development of an extensive and intricately annotated dataset for car damage, encompassing high-resolution images. The purpose of this dataset is to underscore the challenges linked to the detection and segmentation of car damage. Within this segment, we will offer insights into the procedures for gathering, choosing, and annotating the images. Additionally, a thorough examination of the dataset's characteristics and statistics will be provided.

VehiDE primarily focuses on capturing and analyzing visible external damages to vehicles using advanced computer vision technology (Huynh et al., 2023). When outlining the scope of VehiDE, two crucial factors are taken into account: the frequency of incidents and the clarity of definitions. Utilizing statistical data provided by the Insurance Company, VehiDE contains eight specific categories that encompass the frequently encountered forms of external damage: dent, scratch, broken glass, lost parts, punctured, torn, and broken lights. The choice of these categories is based on their frequent incidence and well-defined definitions, in contrast to the more ambiguous 'smash' category, which encompasses a combination of damages. This approach highlights VehiDE's dedication to precision and relevance in meeting the challenges of automated vehicle damage assessment.



To guarantee the presence of top-notch raw data, an exhaustive search for images was carried out on well-known platforms like Flickr and Shutterstock. These platforms provide an extensive collection of high-quality images, covering a variety of damage types and perspectives. The raw image dataset exhibits significant heterogeneity, characterized by variations in illumination, camera position, and capture angle. This inherent diversity strengthens the generalizability of the dataset, allowing it to effectively represent a wider spectrum of real-world conditions.

To automatically detect and eliminate redundant images, we utilized the robust tool Duplicate Cleaner, as outlined in Pham et al. (2023). Before executing the deletion process, we conducted a manual double-check of the identified duplicates to ensure precision. For optimal efficiency, we initially employed a set of manually selected car-damaged images (500 images) to train a binary classifier based on VGG16, as detailed in Simonyan and Zisserman (2015). This classifier underwent specialized training to distinguish whether an image features damaged cars or not. Subsequently, newly acquired images were subjected to classification by the trained classifier. Those identified as positive samples were then designated as candidate images.

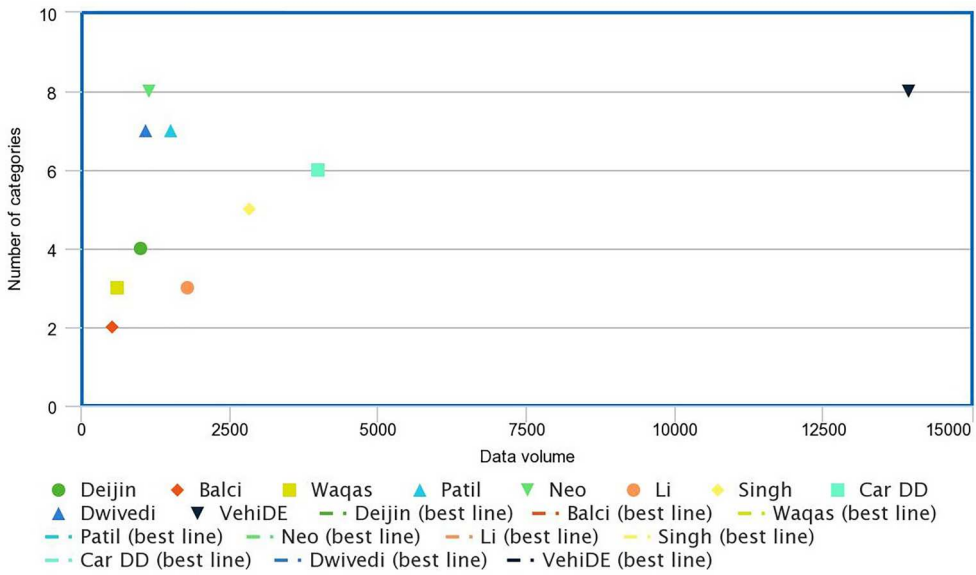
There are more than 42,000 candidate images for using the binary classifier. Some of these extended beyond the parameters of this study, including images of rusty, burned, or smashed cars. Opting for a forward-thinking approach, these images are chosen separately to incorporate into the dataset without annotations, anticipating their potential relevance in subsequent research initiatives. In the ultimate dataset, there are 13,945 samples requiring annotation for the damage assessment process, and these are categorized into eight distinct classes of damage.

VehiDE, in line with established datasets like ImageNet, COCO, and LFW (Deng et al., 2009; Lin et al., 2014), utilizes images under Flickr and Shutterstock licenses. Copyright of these images remains with the original platform providers. Researchers can access the dataset on the condition that they agree to comply with the terms specified in the licenses of Flickr and Shutterstock. Researchers bear full responsibility for their utilization of VehiDE, restricting it solely to non-commercial research and educational purposes. Furthermore, to ensure user privacy, any pictures containing identifiable user information, such as human faces or license plates, have been either mosaicked or completely removed from the dataset.

#### **4.1.2. Labelling dataset**

The primary hurdles faced during the annotation of car damage images involve managing multiple instances and potential uncertainties. To overcome these challenges, we've formulated annotation guidelines grounded in the insurance claim standards set by the Insurance Company. During the annotation process of VehiDE, three crucial rules hold the utmost significance.

Multi-type damage frequently presents in car damage assessments, especially when dents, scratches, and cracks coexist. To address this complexity, a prioritized annotation scheme is implemented. This prioritization ensures crucial information is captured first, followed by ancillary details. For instance, in Figure 3, the image is annotated following the priority order of broken glass > dent > scratch. This prioritization logic draws inspiration from existing car damage repair standards. According to these standards, cracks often necessitate welding operations, and if the crack involves a missing part, that



**Figure 3.** Comparison of VehiDE statistics and other datasets.

particular surface needs to be replaced. Consequently, cracks are considered the most severe form of damage among these three types. Scratches, on the other hand, typically require paint repair, while dents usually involve both metal restoration and paint repair. This distinction leads to a higher priority being assigned to dents over scratches. The implementation of a priority-based annotation rule ensures a faithful representation of both severity and type for each damage present. This strategy aligns with established industry best practices, streamlining accurate assessment and subsequent repair procedures in cases involving multi-type car damage. This prioritization fosters efficient processing and analysis of the data, thereby enhancing the efficacy of subsequent tasks such as damage detection and segmentation algorithms.

In alignment with the established car insurance claim regulations set forth by the Insurance Company, each damaged component incurs separate billing. For instance, consider a scenario where a car exhibits scratches on both the front and rear doors, as depicted in Figure 3. In such cases, an auto body repair shop would typically invoice for the repair of each door individually. This financial segmentation necessitates the annotation process to treat damages spanning multiple car components as distinct instances, with annotations placed along the boundaries of each affected component. This methodology ensures the accurate recording and consideration of individual damages on different car components during the assessment and repair process. By annotating them as distinct instances, we adhere to the insurance company's guidelines, facilitating a precise breakdown of the incurred damages. This, in turn, enables the appropriate allocation of costs and streamlines efficient repair procedures for each affected component. Figure 4 shows some samples of annotated images in VehiDE dataset.

Within the domain of car insurance claims, contiguous instances of the same damage type on a singular component are aggregated into a single annotation, regardless of the damage's individual footprint. This practice stems from the uniform pricing structure



**Figure 4.** Samples of annotated images in the VehiDE dataset. The first row presents the images of damaged vehicle; the second row shows the annotated labels; the last row demonstrates the masks of vehicles.

employed by insurers for such occurrences on a single component. For example, even if discernible demarcations separate dents on the car's carriage, manifested as surface-level wrinkles, the entirety of the carriage's dents would be annotated as one unified instance. This annotation approach aligns with the pricing policies employed in car insurance claims. By consolidating adjacent damages of the same type into one instance, the annotation process ensures consistent billing for similar damages on a particular component. This practice allows for a simplified assessment and estimation of repair costs, upholding fairness and efficiency in the car insurance claim procedure.

To ensure the creation of precise ground truths, a team of 10 individuals was assembled. This team comprised 5 specialists in car damage assessment from the Insurance Company, and the remaining 15 annotators were selected from a pool of candidates. The selection process included imparting professional training to all candidates on the rules of car insurance claims and the previously mentioned annotation guidelines.

Candidate proficiency and annotation accuracy were evaluated through a dedicated quality-testing set comprising 100 car damage images, annotated by domain experts from the Insurance Company. Each candidate then undertook an 'annotation test' involving labelling this test set. To ensure a high level of expertise, only candidates exceeding a 90% recall score were deemed qualified to participate in subsequent annotation tasks. This threshold aligns with recognized benchmarks for reliable annotation quality in similar domains.

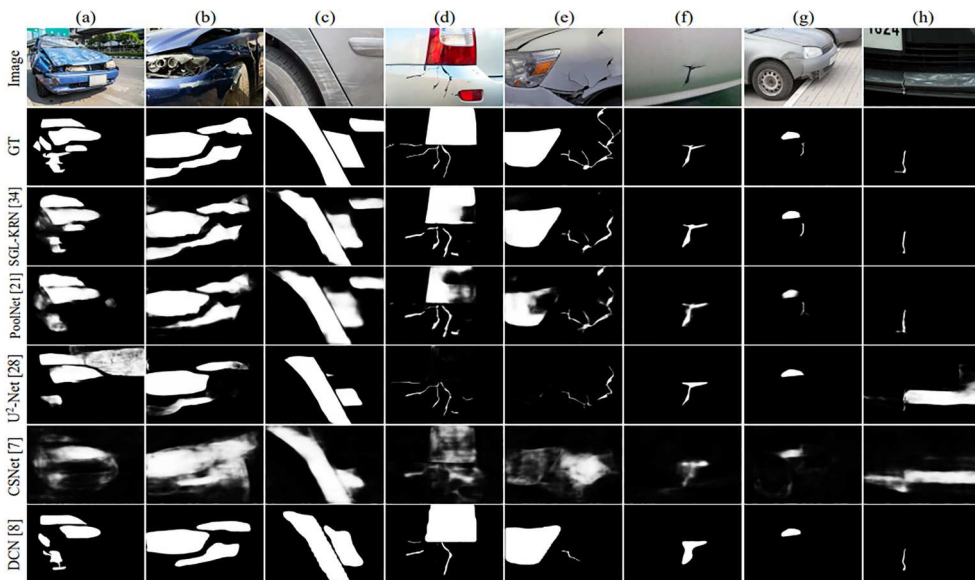
The annotators utilized 27-inch display devices equipped with the ability to zoom in ( $\times 1-8$ ) for enhanced precision. It took approximately 3 minutes to annotate each picture. To enhance precision and uniformity, expert annotators systematically revisited any areas exhibiting inconsistencies, subjecting them to subsequent re-annotation. Through the implementation of this exhaustive process, our goal was to attain dependable and accurate annotations, with a commitment to continual enhancement via expert verification, feedback iterations, and strict adherence to professional training and guidelines.

VehiDE provides a powerful resource for advancing research in car damage assessment. Its ample size, exceptional image quality, and thorough annotations empower researchers to develop and evaluate robust algorithms and models with greater confidence. It is spanning 13,945 images with over 32,000 meticulously labelled instances, the VehiDE dataset comprehensively captures eight distinct categories of car damage. [Figure 3](#) provides a comparison of existing car damage datasets based on two key factors: dataset size (i.e. the

number of images) and the number of categories. It is imperative to highlight that the specifications of these datasets are extracted directly from the respective references, as they are not publicly accessible, as mentioned in Section 2. As discernible from the comparative analysis, the VehiDE dataset distinguishes itself through its notable abundance of samples and a judicious assortment of categories, thereby rendering it advantageous for the training of deep learning models. The baseline experiments demonstrate that the identification of broken glass lost parts, punctured, torn, and broken lights categories have yielded promising results, underscoring the sufficiency of the dataset's volume for meeting the requirements of deep learning approaches.

In the realm of image quality, the proposed dataset markedly outperforms the GitHub dataset (Pham et al., 2023). The GitHub dataset attains its highest resolution at a modest  $684 \times 684$  pixels, in contrast to our dataset's lowest resolution of  $1,000 \times 413$  pixels. Notably, the VehiDE dataset exhibits an impressive average resolution of 684,231 pixels, approximately 13.6 times higher, in comparison to the GitHub dataset's average resolution of 50,334 pixels. Moreover, the average file size of VehiDE is 739 KB (around 82.1 times larger) in contrast to the 9 KB file size of the GitHub dataset. It's noteworthy that the images in our dataset are free from watermarks, ensuring that damages are accurately identified without confusion with watermark artifacts. This stark contrast in image quality underscores the superior visual fidelity and richness of information provided by the VehiDE dataset.

As shown in Figure 6, we use a dataset of corresponding binary maps produced from the annotations of instance segmentation for the salient object detection (SOD) task. The training, validation, and test sets continue to be split in the same proportion as the detection and segmentation experiments. This makes it possible to evaluate the SOD methods fairly and precisely by guaranteeing consistency in the dataset distribution across all tasks.



**Figure 6.** Visual comparison of diverse Salient Object Detection (SOD) methods.

This is accomplished by employing the Duplicate Cleaner tool explicitly to identify and remove near-duplicate images from the dataset (Pham et al., 2023). This meticulous approach ensures the integrity of the training and validation sets while maintaining a representative distribution of instances across different categories. Figure 5 is the structure of the Mask R-CNN network with a parallel semantic segmentation branch.

#### 4.2. Evaluation metrics

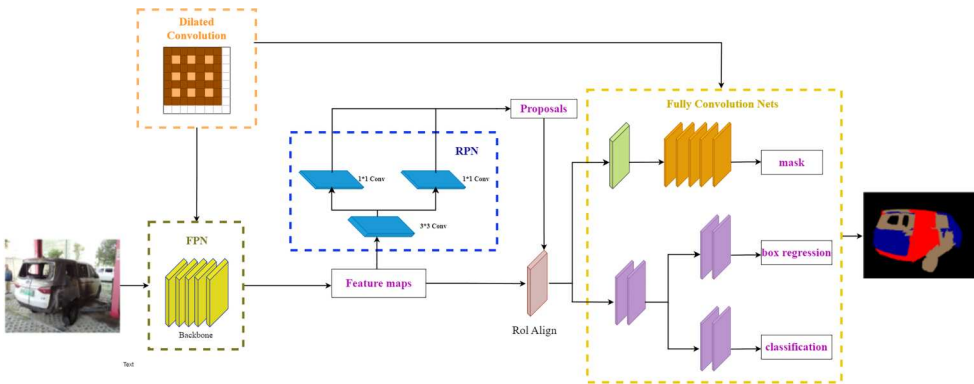
The probability maps generated by deep salient object detection (SOD) techniques maintain the same spatial resolution as the input images. Each pixel in the predicted saliency maps is assigned an integer between 0 and 1 (or [0, 255]). In contrast, ground truths are typically represented as binary masks, where each pixel is either 0 or 1 (or 0 and 255). Within these masks, pixels indicating foreground salient objects have a value of 1, while background pixels have a value of 0. Evaluation metrics for SOD are computed based on the salience maps produced by the models and the corresponding ground truths. These metrics offer a quantitative assessment of the model's effectiveness and performance in accurately identifying and categorizing salient objects within the images. For providing a quantitative evaluation of the performance of the salient object detection (SOD) methods on the VehiDE dataset, there are five widely used metrics are utilized, including F-measure ( $F_\beta$ ), weighted F-measure ( $Fw_\beta$ ), S-measure ( $S_m$ ), E-measure ( $E_m$ ), and mean absolute error (MAE) (Fan et al., 2018).

(1) *F-measure*: F-measure, denoted as  $F_\beta$ , is computed by the weighted harmonic mean of the precision and recall:

$$F_\beta = (1 + \beta^2) \times \frac{\text{Precision} \times \text{Recall}}{\beta^2 \times \text{Precision} + \text{Recall}} \quad (1)$$

where  $\beta^2$  is set to 0.3 as done in previous work to weight precision more than recall.

(2) *Weighted F-measure*:  $F_\beta^w$  is commonly used as an additional metric to address potential biases introduced by interpolation flaws, dependency flaws, and equal-importance flaws (Margolin et al., 2014). It serves as a supplementary measure to  $F_\beta$  and helps overcome any unfair comparisons that may arise due to these flaws. Its



**Figure 5.** Mask R-CNN network with parallel semantic segmentation branch.

calculation is defined as follows:

$$F_{\beta}^w = \frac{(1 + \beta^2) \times Precision^w \times Recall^w}{\beta^2 \times Precision^w + Recall^w}. \quad (2)$$

(3) *S-measure*: denoted as  $S_m$  is employed to assess the structural similarity between the predicted non-binary saliency map and the ground truth. The *S-measure* is calculated as a weighted sum of two measures, which are region-aware structural similarity ( $S_r$ ) and object-aware structural similarity ( $S_o$ ), to capture both the overall structural consistency at the region level and the specific object-level structural similarity:

$$S = (1 - \alpha)S_r + \alpha S_o \quad (3)$$

where  $\alpha$  is usually set to 0.5.

(4) *E-measure*: The Enhanced-alignment measure, referred to as  $E_m$ , integrates local pixel values with the mean value of the image at the global level: creating a unified metric that simultaneously incorporates image-level statistics and local pixel matching information:

$$Q_{FM} = \frac{1}{w \times h} \sum_{x=1}^w \sum_{y=1}^h \phi_{FM}(x, y) \quad (4)$$

where  $h$  and  $w$  are the height and the width of the map, respectively.

(5) *MAE*: This measure represents the average difference per pixel between a predicted saliency map and its corresponding ground truth mask. The MAE score indicates the similarity between a predicted saliency map, denoted as  $P$ , and the ground truth map, denoted as  $G$ :

$$MAE = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H |P(x, y) - G(x, y)| \quad (5)$$

where  $W$  and  $H$  denote the width and height of the saliency map, respectively.

### 4.3. Salient object prediction results

The difficult instances categorized as ‘dent,’ ‘scratch,’ and ‘broken glass’ pose substantial challenges for the existing state-of-the-art (SOTA) methods employed in tasks such as segmentation and object detection, as indicated by the result analysis in Section V-B. The irregular shapes and flexible boundaries of these instances make it challenging to accurately detect and segment them. Furthermore, the resemblance in contour and color features increases the likelihood of confusion between the ‘scratch’ and ‘broken glass’ classes. We propose exploring the application of salient object detection (SOD) methods to address these challenges, as they may offer more suitable solutions for two key reasons.

Four state-of-the-art (SOTA) methods specifically designed for the salient object detection (SOD) task are applied to the VehiDE dataset: CSNet (Cheng et al., 2022), PoolNet (Liu et al., 2019), U<sup>2</sup>-Net (Qin et al., 2020), and SGL-KRN (Xu et al., 2021). As utilizing a new dataset, the released code is provided by the respective authors and trains all the models from scratch on the VehiDE dataset. This approach ensures that the models are

**Table 2.** SOD method hyper-parameters.  $[i, j]$  denotes the interval between the  $i$ th and  $j$ th epochs.

	SGL-KRN	PoolNet	U <sup>2</sup> Net	CSNet
Weight Decay	5e-2	5e-2	0	5e-1
Batch Size	6	6	12	16
Maximun Epochs	100	100	600	300
Learning Rate	[1, 50]: 5e-2 [51, 100]: 5e-4	[1, 50]: 5e-2 [51, 100]: 5e-4	2e-3	[1, 200]: 1e-4 [201, 250]: 1e-5 [251, 300]: 1e-6

trained specifically on the characteristics and nuances of the VehiDE dataset, rather than fine-tuning pre-trained models provided by the authors.

The experiments result perform specific to each damage category to evaluate the performance of the salient object detection (SOD) methods. This involves training and testing the models on the VehiDE dataset, where the binary maps activate masks for only one damage class at a time. Through these class-wise experiments, we can assess how well the SOD methods effectively identify and segment individual damage categories. This approach provides valuable insights into the strengths and limitations of the methods for each specific category.

Table 2 provides a summary of the hyperparameters used during the training process. The Adam optimizer is used for all the models (Kingma & Ba, 2015). It's crucial to remember that SGL-KRN and PoolNet only allow for batches of 6. About the U<sup>2</sup>-Net model, we train it using a batch size of 12 for a total of 140,800 iterations. This is the same as training the VehiDE dataset using the approximately 6,278 images in the training set for 600 epochs.

#### 4.4. Comparison results and analyzes

The comparison between CSNet, U<sup>2</sup>-Net, PoolNet, and SGL-KRN uses the proposed evaluation metrics, including  $F_\beta$ ,  $F_\beta^w$ ,  $S_m$ ,  $E_m$ , and  $MAE$ , as shown in Table 3. SGL-KRN among these models demonstrated superior performance across most metrics, except the E-measure, where PoolNet and MAE stood out as the top performers. Table 4 evaluates SGL-KRN's performance in each category. It is important to note that, in comparison to the outcomes from segmentation and detection tasks, the performance gap between the easy and hard classes has significantly decreased.

In these results, SGL-KRN is better than others in all measures. Table 4 is the detailed results of quantitative salient object detection for each class by SGL-KRN.

The visualization of various generated saliency maps is shown in Figure 6. The chosen scenarios are designed to showcase the models' proficiency in various challenging situations, including those with a high number of objects (columns (a), (b), and (c)), slender objects (columns (d) and (e)), and small objects (columns (f), (g), and (h)). It's important to note that in each image, only masks corresponding to a single type of damage are activated to ensure clarity in the demonstration. Each instance is assigned a unique ID in the

**Table 3.** Results of CSNet, U<sup>2</sup>-Net, PoolNet, and SGL-KRN's quantitative salient object detection.

	$F_\beta \uparrow$	$F_\beta^w \uparrow$	$S_m \uparrow$	$E_m \uparrow$	$MAE \downarrow$
CSNet (Cheng et al., 2022)	0.634	0.652	0.720	0.770	1.120
U <sup>2</sup> -Net (Qin et al., 2020)	0.696	0.701	0.809	0.810	0.730
PoolNet (Liu et al., 2019)	0.823	0.775	0.834	0.885	0.069
SGL-KRN (Xu et al., 2021)	<b>0.832</b>	<b>0.805</b>	<b>0.842</b>	<b>0.938</b>	<b>0.067</b>



**Table 4.** Results of SGL-KRN's quantitative salient object detection for each class.

	$F_\beta \uparrow$	$F_\beta^w \uparrow$	$S_m \uparrow$	$E_m \uparrow$	$MAE \downarrow$
dents	0.689	0.672	0.770	0.879	0.058
broken glass	0.801	0.823	0.834	0.822	0.051
scratch	0.731	0.781	0.758	0.831	0.062
lost parts	0.832	0.855	0.882	0.921	0.032
punctured	0.849	0.782	0.863	0.895	0.043
torn	0.984	0.936	0.992	0.935	0.025
broken lights	0.885	0.925	0.873	0.938	0.035
All classes	<b>0.820</b>	<b>0.761</b>	<b>0.836</b>	<b>0.912</b>	<b>0.046</b>

annotation files for reference. These visualizations offer a straightforward way to assess the models' performance in handling these complex scenarios.

According to the visualizations' observations, SGL-KRN and PoolNet both perform well, successfully capturing and highlighting important objects in these scenarios. However, in some scenarios, U<sup>2</sup>Net and CSNet show flaws by either missing objects or producing false positives.

The segmentation outputs of DCN are converted into saliency maps in order to compare the detail-processing capabilities of SOD methods with instance segmentation methods. The saliency maps' highlighted objects have higher than 0.5 prediction probabilities. SOD techniques like SGL-KRN and PoolNet are more effective at precisely locating salient objects. Even in difficult situations involving small and slender objects, SOD methods can clearly distinguish the boundaries of salient objects. The visual outcomes show the SOD methods' superiority in detail processing and offer a fresh approach to estimating auto damage.

## 5. Conclusions and future work

Automated Car Damage Assessment (CDA) systems hold significant potential for societal and economic benefits. A well-designed CDA system can greatly enhance online evaluation efficiency and substantially reduce human effort. This eliminates the need for vehicle owners to physically transport their cars to insurance companies, allowing quick access to evaluation results. This not only alleviates traffic congestion but also contributes to overall cost reduction for society. Unfortunately, the lack of comprehensive datasets for analysis and study has hindered research progress in vehicle damage assessment.

In this study, the VehiDE dataset is proposed and expanded from Huynh et al. (2023). This is a large-scale car damage dataset publicly available with detailed annotations for damage detection and segmentation. The built dataset is located at<sup>1</sup>. VehiDE comprises 13,945 high-resolution images of damaged cars, featuring over 32,000 meticulously annotated examples across eight different damage categories. The dataset supports four main tasks: classification, object detection, instance segmentation, and salient object detection. Various state-of-the-art models are assessed on the VehiDE dataset, providing both quantitative and graphical analyzes of the results. The analysis underscores the challenges in accurately identifying and categorizing categories with diverse small and irregular shapes. The potential causes for these challenges are also explored and proposed emerging remedies to address them.

VehiDE is a useful addition to segmentation tasks already in use, such as COCO (Lin et al., 2014) and Pascal VOC (Everingham et al., 2010), which primarily concentrate on frequently encountered objects. Additionally, algorithms created using the VehiDE dataset can advance other related tasks like anomaly detection, object recognition for irregular

shapes, and fine-grained segmentation thanks to the inclusion of multiple types of fine-grained annotations. Moreover, gathering and annotating more pictures of auto damage will be developed and added to the VehiDE Dataset in the future. The goal is for VehiDE to significantly advance the field of research, advancing not only the understanding of vision-related tasks but also the quality of ordinary life.

## Note

1. <https://github.com/christopherohit/CarInspectionDamage>

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work belongs to the project grant No: T2024-162 funded by Ho Chi Minh City University of Technology and Education, Vietnam.

## Notes on contributors

**Van-Dung Hoang** received the Ph.D. degree in electrical and computer engineering from the University of Ulsan, Ulsan, South Korea, in 2015. After a year of experience with Telecom Sudparis as a Postdoctoral Research Fellow, he joined the Faculty of Information Technology, HCMC University of Technology and Education, Ho Chi Minh City, Vietnam, where he is currently a Professor and the Head of the Artificial Intelligence Department. He has a total academic teaching experience of 18 years with more than 85 publications in reputed international conferences, journals, and book chapter contributions (Indexed By: ISI, WoS, Scopus, ACM DL, and DBLP). His research interests include computer vision, robotics, medical image processing, autonomous vehicle, and ambient intelligence.

**Nhan T. Huynh** receive B.S from the University of Transportation and Community, Currently. He is studying Master at University Information Technology, VNU-HCM from 2023. His research interests are computer vision, deep learning. He also received Best Theory Paper Award at SOMET 2022.

**Nguyen Tran** received his B.S. degree in Computer Science from Saigon Technology University in Ho Chi Minh City, Vietnam, in 2019. In 2024, he obtained his M.S. degree in Data Science from the University of Science, VNU-HCM, Vietnam. His research interests are mostly focused on data analysis, data processing, and computer vision, particularly medical image processing. Notably, in 2022, he received the Best Theory Paper Award at SOMET 2022.

**Khanh Le** is currently a third-year honors student majoring in Computer Science at the University of Information Technology, VNU-HCM, Vietnam. His research interests is computer vision. He had a paper accepted at the MediaEval 2023 Multimedia Benchmark Workshop in February 2024.

**Thi-Minh-Chau Le** graduated from university in Infomaticsin 2005 at the Open University and received a master's degree in computer science in 2012 at Ho Chi Minh City University of Technology. She is currently working at the Faculty of Information Technology, HCMC University of Technology and Education, Vietnam. Her research interests include security, privacy, Big Data AI, and related technologies.

**Ali Selamat** has received a B.Sc. (Hons.) in IT from Teesside University, U.K. and M.Sc. in Distributed Multimedia Interactive Systems from Lancaster University, U.K. in 1997 and 1998, respectively. He has received a Dr. Eng. degree from Osaka Prefecture University, Japan, in 2003. His research

interests include data analytics, digital transformations, knowledge management in higher education, key performance indicators, cloud-based software engineering, software agents, information retrievals, pattern recognition, genetic algorithms, neural networks, and soft-computing. He received best paper awards from ACM for the 37th International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems 2024 (IEA/AIE 2024).

**Hien D. Nguyen** received his B.S. and M.S. degrees from the University of Sciences, VNU-HCM, Vietnam, in 2008 and 2011, respectively. He received his Ph.D. degree from the University of Information Technology, VNU-HCM, in 2020. He is a senior lecturer at the Faculty of Computer Science, University of Information Technology, VNU-HCM, Vietnam. His research interests include knowledge representation, knowledge-based systems, and knowledge engineering, especially intelligent and expert systems. He received the Best Paper Awards at CITA 2023, SOMET 2022, and ICOCO 2022, Best Student Paper Awards at KEOD 2023 and KSE 20202, Incentive Prizes of the Technological Creation Awards of Binh Duong province in 2021 and VIFOTEC in 2016.

## References

- Al-Shemarry, M. S., Li, Y., & Abdulla, S. (2020). An efficient texture descriptor for the detection of license plates from vehicle images in difficult conditions. *IEEE Transactions on Intelligent Transportation Systems*, 21(2), 553–564. <https://doi.org/10.1109/TITS.6979>
- Balci, B., Artan, Y., & Elihos, A. (2019). *Front-view vehicle damage detection using roadway surveillance camera images*. Proceedings of the 5th International Conference on Vehicle Technology and Intelligent Transport Systems – Volume 1: Vehits (pp. 193–198), Crete, Greece, Scitepress.
- Cheng, M. M., Gao, S. H., Borji, A., Tan, Y. Q., Lin, Z., & Wang, M. (2022). A highly efficient model to study the semantics of salient object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11), 8006–8021. <https://doi.org/10.1109/TPAMI.2021.3107956>
- de Deijn, J (2018). Automatic car damage recognition using convolutional neural networks.
- Deng, J., Dong, W., Socher, R., Li, L. J, Li, K, & Fei-Fei, L. (2009). *ImageNet: a large-scale hierarchical image database*. In 2009 IEEE Conference on Computer Vision and Pattern Recognition (pp. 248–255), Miami, Florida, USA, IEEE.
- Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2019). *A very deep transfer learning model for vehicle damage detection and localization*. In 2019 31st International Conference on Microelectronics (ICM) (pp. 158–161), Cairo, Egypt, IEEE.
- Everingham, M., Gool, L. V., Williams, C., Winn, J., & Zisserman, A. (2010). The pascal visual object classes (VOC) challenge. *International Journal of Computer Vision*, 88(2), 303–338. <https://doi.org/10.1007/s11263-009-0275-4>
- Fan, D. P., Gong, C., Cao, Y., Ren, B., Cheng, M. M., & Borji, A. (2018). *Enhanced-alignment measure for binary foreground map evaluation*. Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI 2018) (pp. 698–704), Stockholm, Sweden, ACM.
- He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). *Mask R-CNN*. Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, IEEE.
- Hoang, V. D., Vo, X. T., & Jo, K. H. (2023). Categorical weighting domination for imbalanced classification with skin cancer in intelligent healthcare systems. *IEEE Access*, 11, 105170–105181. <https://doi.org/10.1109/ACCESS.2023.3319087>
- Huang, Y., Liu, Z., Jiang, M., Yu, X., & Ding, X. (2020). Cost-effective vehicle type recognition in surveillance images with deep active learning and web data. *IEEE Transactions on Intelligent Transportation Systems*, 21(1), 79–86. <https://doi.org/10.1109/TITS.6979>
- Huynh, A. T., Hoang, V. D., Vu, S., Le, T. T., & Nguyen, H. D. (2022). *Skin cancer classification using different backbones of convolutional neural networks*. Proceeding of 35th International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA/AIE 2022) (pp. 160–172), Kitakyushu, Japan, Springer.
- Huynh, N. T., Tran, N., Huynh, A., Hoang, V. D., & Nguyen, H.D. (2023). *VehiDE dataset: New dataset for automatic vehicle damage detection in car insurance*. In Proceedings of 15th International

- Conference on Knowledge and Systems Engineering (KSE 2023) (pp. 322–327), Ha Noi, Vietnam, IEEE.
- Jeon, H. J., Nguyen, V. D., Duong, T. T., & Jeon, J. W. (2022). A deep learning framework for robust and real-time taillight detection under various road conditions. *IEEE Transactions on Intelligent Transportation Systems*, 23(11), 20061–20072. <https://doi.org/10.1109/TITS.2022.3178697>
- Jocher, G. (2020). Ultralytics YOLOv5 [Computer software manual]. <https://github.com/ultralytics/yolov5>.
- Kingma, D. P., & Ba, J. (2015). *Adam: A method for stochastic optimization*. In The 3rd International Conference on Learning Representation (ICLR 2015), San Diego, CA, USA, ACM.
- Li, P., Shen, B., & Dong, W. (2018). An anti-fraud system for car insurance claim based on visual evidence. Preprint [arXiv:1804.11207](https://arxiv.org/abs/1804.11207).
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., & C. L. Zitnick. (2014). *Microsoft coco: Common objects in context*. In Proceedings of 13th European Conference Computer Vision (ECCV), Part V 13 (pp. 740–755), Zurich, Switzerland, Springer.
- Liu, J. J., Hou, Q., Cheng, M. M., Feng, J., & Jiang, J. (2019). *A simple pooling-based design for real-time salient object detection*. In Proceedings of The IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3917–3926), Long Beach, CA, USA, IEEE.
- Mahavir, D., Hashmat Shadab, M., Omark, S. N., Monis, E. B., Khanna, B., S. R. Samal, Tiwari, A., & Rath, A. (2021). Deep learning-based car damage classification and detection. In C. Niranjana F. Takanori (Eds.), *Advances in artificial intelligence and data engineering* (pp. 207–221). Springer Nature Singapore.
- Margolin, R., Zelnik-Manor, L., & Tal, A. (2014). *How to evaluate foreground maps?* In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 248–255), Columbus, OH, USA, IEEE.
- Medrano-Berumen, C., & M. İlhan Akbaş (2021). Validation of decision-making in artificial intelligence-based autonomous vehicles. *Journal of Information and Telecommunication*, 5(1), 83–103. <https://doi.org/10.1080/24751839.2020.1824154>
- Meißner, K., & Rieck, J. (2021). Multivariate forecasting of road accidents based on geographically separated data. *Vietnam Journal of Computer Science*, 8(3), 433–454. <https://doi.org/10.1142/S2196888821500196>
- Nguyen, H., Tran, N., Nguyen, H. D., Nguyen, L., & Kotani, K. (2023a). KTFEv2: Multimodal facial emotion database and its analysis. *IEEE Access*, 11, 17811–17822. <https://doi.org/10.1109/ACCESS.2023.3246047>
- Nguyen, H., Tran, V. P., Pham, V. T., & Nguyen, H. D. (2022). Designing a learning model for mobile vision to detect diabetic retinopathy based on the improvement of mobilenetv2. *International Journal of Digital Enterprise Technology*, 2(1), 38–53. <https://doi.org/10.1504/IJDET.2022.124987>
- Nguyen, H. D., Truong, D., Vu, S., Nguyen, D., Nguyen, H., & Tran, N. (2023b). Knowledge management for information querying system in education via the combination of rela-ops model and knowledge graph. *Journal of Cases on Information Technology (JCIT)*, 25(1), 1–17. <https://doi.org/10.4018/JCIT>
- Parhizkar, M., & Amirfakhrian, M. (2022). Recognizing the damaged surface parts of cars in the real scene using a deep learning framework. *Mathematical Problems in Engineering*, 2022, 5004129. <https://doi.org/10.1155/2022/5004129>
- Patel, N., Shinde, S. R., & Poly, F. (2020). Automated damage detection in operational vehicles using mask R-CNN. In *Advanced computing technologies and applications* (pp. 563–571). Springer.
- Patil, K., Kulkarni, M., Sriraman, A., & Karande, S. S. (2017). *Deep learning based car damage classification*. Proceedings of 16th IEEE International Conference on Machine Learning and Applications (ICMLA 2017) (pp. 50–54), Cancun, Mexico, IEEE.
- Pham, T. A., & Hoang, V. D. (2024). Chest X-ray image classification using transfer learning and hyperparameter customization for lung disease diagnosis. *Journal of Information and Telecommunication*, 1–15. <https://doi.org/10.1080/24751839.2024.2317509>.
- Pham, V., Nguyen, H., Pham, B., Nguyen, T., & Nguyen, H. (2023). Robust engineering-based unified biomedical imaging framework for liver tumor segmentation. *Current Medical Imaging*, 19(1), 37–45. <https://doi.org/10.2174/1573405617666210804151024>

- Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., & Jagersand, M. (2020). U2-net: Going deeper with nested U-structure for salient object detection. *Pattern Recognition*, 106, 107404. <https://doi.org/10.1016/j.patcog.2020.107404>
- Reddy, D. A., Shambharkar, S., Jyothsna, K., Kumar, V. M., Bhoyar, C. N., & Somkunwar, R. K. (2022). *Automatic vehicle damage detection classification framework using fast and mask deep learning*. Proceedings of 2nd International Conference on Computer Science, Engineering and Applications (ICCSEA) (pp. 1–6), Gunupur, India, IEEE.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You only look once: Unified, real-time object detection*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, IEEE.
- Simonyan, K., & Zisserman, A. (2015). *Very deep convolutional networks for large-scale image recognition*. Proceedings of International Conference on Learning Representation (ICLR 2015), San Diego, USA.
- Singh, R., Ayyar, M. P., Pavan, T. S., Gosain, S., & Shah, R. R. (2019). *Automating car insurance claims using deep learning techniques*. Proceedings of 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM) (pp. 199–207), Singapore, IEEE.
- Tran, N. N., Nguyen, H. D., Huynh, N. T., Tran, N. P., & Nguyen, L. V. (2022). Segmentation on chest CT imaging in COVID-19 based on the improvement attention U-Net model. In *New trends in intelligent software methodologies, tools and techniques* (pp. 596–606). IOS Press.
- Wang, X., Li, W., & Wu, Z. (2022). Cardd: A new dataset for vision-based car damage detection. *IEEE Transactions on Intelligent Transportation Systems*, 24(7), 7202–7214. <https://doi.org/10.1109/TITS.2023.3258480>
- Waqas, U. A., Akram, N., Kim, S. Y., Lee, D., & Jeon, J. Y. (2020). *Vehicle damage classification and fraudulent image detection including moiré effect using deep learning*. In 2020 . Proceedings of IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (pp. 1–5), Edmonton, AB, Canada, IEEE.
- Xiang, Y., Fu, Y., & Huang, H. (2020). Global topology constraint network for fine-grained vehicle recognition. *IEEE Transactions on Intelligent Transportation Systems*, 21(7), 2918–2929. <https://doi.org/10.1109/TITS.6979>
- Xiang, G., Yao, S., Deng, H., Wu, X., Wang, X., Xu, Q., Yu, T., Wang, K., & Peng, Y. (2024). A multi-modal driver emotion dataset and study: Including facial expressions and synchronized physiological signals. *Engineering Applications of Artificial Intelligence*, 130, 107772. <https://doi.org/10.1016/j.engappai.2023.107772>
- Xu, B., Liang, H., Liang, R., & Chen, P. (2021). *Locate globally, segment locally: A progressive architecture with knowledge review network for salient object detection*. Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35 (pp. 3004–3012), Virtually.
- Zhang, W., Cheng, Y., Guo, X., Guo, Q., Wang, J., Wang, Q., & Chu, W. (2020). *Automatic car damage assessment system: Reading and understanding videos as professional insurance inspectors*. Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, No. 9, pp. 13646–13647, New York, USA.
- Zhang, L., Wang, P., Li, H., Li, Z., Shen, C., & Zhang, Y. (2021). A robust attentional framework for license plate recognition in the wild. *IEEE Transactions on Intelligent Transportation Systems*, 22(11), 6967–6976. <https://doi.org/10.1109/TITS.2020.3000072>
- Zhang, X., Zhang, R., Cao, J., Gong, D., You, M., & Shen, C. (2022). Part-guided attention learning for vehicle instance retrieval. *IEEE Transactions on Intelligent Transportation Systems*, 23(4), 3048–3060. <https://doi.org/10.1109/TITS.2020.3030301>
- Zhu, Q., Hu, W., Liu, Y., & Zhao, Z.. (2021). *Research on vehicle appearance damage recognition based on deep learning*. Proceedings of the 5th International Conference on Machine Vision and Information Technology (CMVIT 2021), Virtually, IOP Science.