IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 8, August 2021

DOI 10.17148/IJARCCE.2021.10808

Car Damage Detection using Machine Learning

Girish N[1], Mohammed Aqeel Arshad[2]

Student, Department of Information Science and Engineering, RV College of Engineering, Bengaluru, India[1]

Student, Department of Information Science and Engineering, RV College of Engineering, Bengaluru, India[2]

Abstract: Vehicle insurance processing using images is a critical sector with a lot of room for automation. In this study, we look at the topic of car damage detection. Vehicle damage detection. Using images taken at the site of an accident can save time and money when filing insurance claims, as well as provide more convenience for drivers. Artificial Intelligence (AI) in the sense of machine learning and deep learning algorithms can assist in solving problems. A vehicle-damage-detection technique based on transfer learning and a mask regional convolutional neural network (Mask RCNN) are utilized to quickly handle accident compensation problems. The algorithms identify the damaged section of a car, determine its position, and then estimate the severity of the damage. Very satisfactory results have been produced using transfer learning to take advantage of available models that have been trained on a more generic object identification challenge.

Keywords: Car Damage Detection, Machine learning, Prediction, Mask RCNN, Transfer Learning, Deep Learning

INTRODUCTION

Today, one of the first businesses to invest in innovation, cutting-edge technology, and artificial intelligence (IA) is the insurance industry.[1] Car insurance companies spend millions of dollars each year owing to claims leakage in today's society, when the rate of car accidents is on the rise. In the insurance industry, Artificial intelligence (AI) based on machine learning and deep learning can assist with challenges including data analysis and processing, fraud detection, risk reduction, and claim automation. As a result, insurance companies have sought to reduce the time it takes to analyze damage and settle claims.

However, developing current applications to address such issues remains difficult, particularly when using deep learning to assess automotive damage. Deep learning is an effective method for tackling complicated problems, but it necessitates more resources for model building, i.e., deep learning demands a large dataset and takes longer to compute.[3]

There are no publicly available datasets for automobile damaged photos because car damage assessment is a specialized topic. The most difficult part of training a model is doing it with a small dataset. However, in this instance we use a dataset format called COCO format. Using a Mask RCNN algorithm on a small COCO dataset has also shown precise results. We use a pre-trained model from Facebook's open-source Python Library called 'detectron2'.

The Mask RCNN method is used in the paper to detect and segment automotive damaged areas. It has a lot of research value and a lot of application scenarios in the transportation area. Due to the complexity of automatic damage detection and segmentation, issues such as reduced segmentation and lower detection speed exist. The Mask RCNN is applied to the field of this research, and a model for detecting and segmenting a vehicle's damaged region in an accident is proposed. This model can also be used by insurance firms to process claims fast.

I. LITERATURE REVIEW

One of the key research topics in computer vision is object detection. On the instance level, it determines the category and position information of the object of interest in the image. RCNN [2], Fast RCNN [3], Faster RCNN [4], and SSD[5] are a some of the most popular target detection algorithms. These frameworks, on the other hand, necessitate a large quantity of training data and thus end-to-end detection is not possible. The detection frame's positioning ability is limited, and the gradient disappearance or gradient explosion is common when a feature is extracted as the number of convolution layers grows.

For these drawbacks, Author [6] proposed a residual network (ResNet) that uses the residual module to help the model converge, accelerates neural network training, and integrates it with the Mask RCNN target detection model to achieve object detection and segmentation, significantly enhancing model detection accuracy. Mask RCNN is the first deep learning model that incorporate target identification and segmentation in a single network.

Candidate regions are used in the majority of contemporary instance segmentation techniques. The Author [10] presented a DeepMask segmentation model that generates prediction candidate masks based on the instances in the input image, allowing each instance object to be segmented, although border segmentation accuracy is low. Full convolutional instance

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 8, August 2021

DOI 10.17148/IJARCCE.2021.10808

segmentation was proposed by Author [11] as the first end-to-end instance segmentation framework (FCIS). Both the bounding box and instance segmentation are predicted by FCIS by enhancing the position sensitive score map, but when processing overlapping object instances, it can only detect the approximate boundary of each instance object [12].

The similarity in papers [19,20,21], is that used both transfer learning and ensemble learning to train a CNN model by comparing the results of fine-tuning in a pre-trained CNN model on an ImageNet dataset, focusing on damage detection accuracy. As a crucial method for increasing damage detection performance, the YOLO object identification model was employed by Author [22] to train and detect damage zones.

II. METHEDOLOGY

This section discusses the COCO dataset used for testing, validating and training the model. The details of COCO dataset and the reason to use this dataset is explained here. The basic composition of the COCO dataset and what basically a COCO dataset will be explained here. The dataset being used for this model is downloaded from Kaggle.

The dataset contains about 80 images (which is very less considering the fact that the model will not be accurate). In these 80 images – 59 images are separated for training (train dataset), 11 images are separated for validation (val dataset) and the remaining images are testing purpose (test dataset).

A. COCO Dataset Format

The most popular object detection dataset at the present is Microsoft's Common Objects in Context dataset (COCO). It is commonly used to evaluate the performance of computer vision algorithms. Because of the dataset's popularity, the COCO format for storing annotations is frequently used when constructing a new custom object detection dataset. Annotations for other tasks, such as segmentation, are also supported by the COCO dataset. The "COCO format" is a JSON structure that specifies how labels and metadata for an image collection are saved.

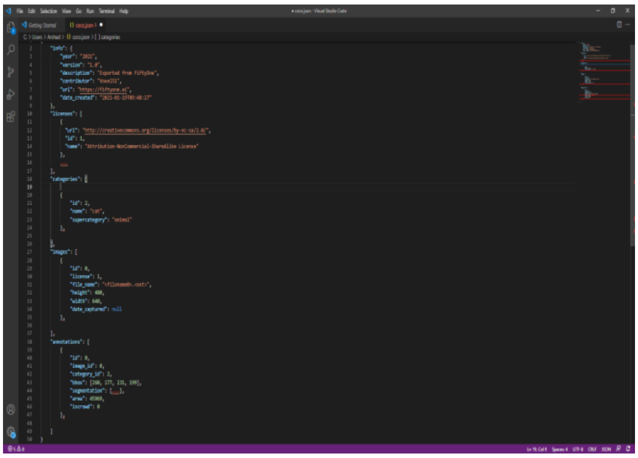

Figure 1: Basic COCO Format

- **Info:** Description and versioning information about the dataset.
- **Licenses:** List of licenses with unique IDs to be specified by your images.
- **Categories:** Classification categories each with unique ID. Optionally associate with supercategory that can span multiple classes.
- **Images:** List of images in the dataset and relevant metadata including unique image ID, filepath, height, width and optional attributes like license, URL, date captured, etc.
- **Annotations:** List of annotations, each with its own ID and the image ID to which it refers. This is where the information for later jobs is saved, such as segmentation/keypoint/other label information. This also holds the bounding area and 'iscrowd', which indicates a huge bounding box enclosing several objects of the same category for evaluation.
- **Data Visualization**

As we are using COCO dataset we do not need to preprocess the dataset. This because the data is already preprocessed and stored in the form of a JSON file. Therefore in order to visualize the data i.e inorder to check whether the data given data JSON file is correct or not. To Visualize the dataset we just simply import a random image from the

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 8, August 2021

DOI 10.17148/IJARCCE.2021.10808

validation dataset (val dataset) and use the data related to that image from the JSON file and plot the damaged regions. In the Figures below, it can be seen that first a random raw image is generated and then the damaged areas are plotted. This helps in concluding if the JSON data is valid or not before beginning the training process.l vu

Figure 2: A Random Image from Val dataset

Figure 3: Visualizing based on the JSON data.

III. IMPLIMENTATIONS

A Machine learning algorithm called Mask RCNN is used in the implementation of the model in this paper. The model is implemented using supervised learning. Figure 4 depicts a method for detecting and segmenting vehicle damage based on the Mask RCNN model. The JSON data containing the details of the images are passed to the Mask RCNN for classification prediction, feature extraction and segmentation masking, and the outcome of the vehicle damage detection is returned i.e., the image with damaged areas is the output. The damaged area is represented using bounding-boxes and segmentation masks.

IJARCCE

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 8, August 2021

DOI 10.17148/IJARCCE.2021.10808

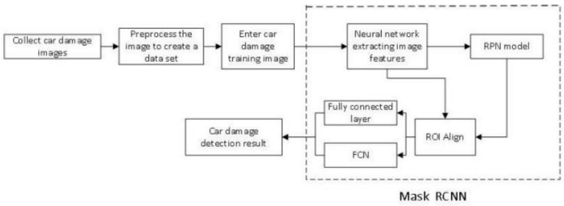

Figure 4: Car Damage segmentation framework.

A. Mask RCNN Algorithm

The Mask RCNN basically is a Faster RCNN-enhanced instance segmentation framework. It has two stages, in which the first scans the image and creates a proposal, while the second classifies it and creates the bounding box and mask.

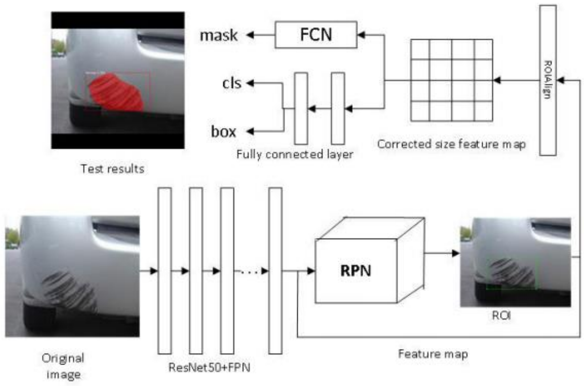

Figure 5: Mask RCNN network framework model.

The algorithm flow is as follows:

1) In order to extract features and create related feature maps, feed the image which has to be identified for damage into a previously trained FPN+RestNet50 model.
2) This feature map uses RPN to obtain a large number of candidate frames (i.e., the ROI), and then performs binary classification of foreground and background by using the SoftMax classifier, as well as frame regression in order to gather more precise information on the candidate-frame position and non-maximum suppression to filter out part of the ROI.
3) The feature map is passed to the RoIAlign layer, along with the last remaining ROI, resulting in a fixed-size feature map for each ROI.
4) Now the flow finally splits into two branches, of which one goes for frame regression and object classification to the fully connected layer and the other goes for pixel segmentation to the full convolution network.

B. Backbone Improvement

In general, Mask RCNN's backbone network uses ResNet101, which means there are 101 network layers. However, having too many layers reduces the network structure's rate significantly. The damage parts category developed in the study is quite straightforward and also RestNet50 is employed in this model with the intention to increase the algorithm's running performance.

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 8, August 2021
DOI 10.17148/IJARCCE.2021.10808

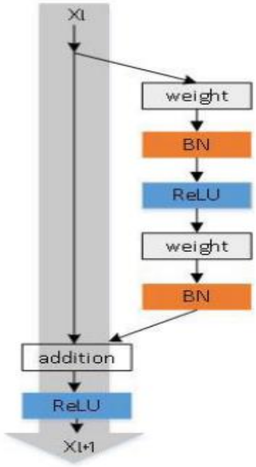

Figure 6: ResNet structure.

A single convolutional neural network will not be able to extract all of the image features effectively because the magnitude of the car damage in each image will vary. As a result, this model employs the ResNet50 backbone structure as well as an FPN feature pyramid network. Using a top-down hierarchy with lateral connections, FPN tackles the multi-scale difficulty of extracting target objects in the images and developing a network feature pyramid by starting with a single-scale input. This structure is quite resilient and adaptable, and it only requires a few parameters. Figure 6 above shows the ResNet structure.

C. RPN MODEL

The Feature Pyramid Networks structure is used, and the photos are resized to generate features that match to various sizes. Small targets can be distinguished by shallow characteristics, while large targets can be distinguished by deep features. The RPN receives the different-size feature maps generated by the FPN and extracts the RoI features from different levels of the feature pyramid depending on the size of the target object. As a result, the simple network structure changes without significantly increasing the computation amount, dramatically boosting the detection performance of small objects while also achieving good accuracy and speed improvements. Figure 7 shows the ROI generated by the RPN.

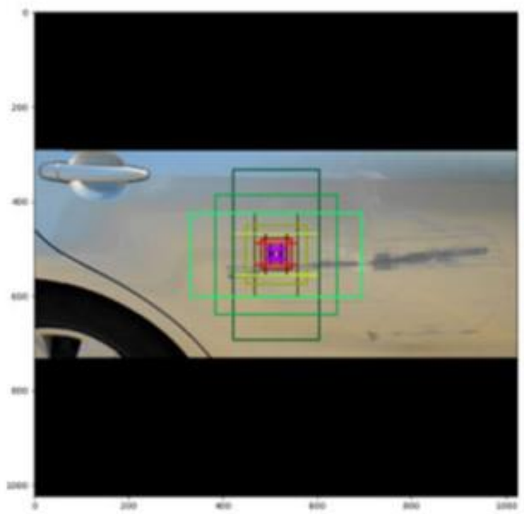

Figure 7: ROI generated by the RPN

A sliding-window-based classless target detector is equivalent to RPN. It has a convolutional neural network-style structure. Anchor frame anchors are created by scanning the sliding frame. To cover as many images as feasible, a recommended region can generate numerous anchors of varied sizes and aspect ratios that overlap.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 8, August 2021

DOI 10.17148/IJARCCE.2021.10808

B. ROIAlign Model

The mask branch of the Mask RCNN network structure must detect whether or not a given pixel is part of the target, with precision at the pixel level. The scale of the image has changed after it has been substantially convolved and pooled. The image target object cannot be precisely positioned when segmentation at pixel level is performed directly, therefore on the basis of Faster RCNN the Mask RCNN is improved. Also, the RoI Pooling layer is turned into the interest region alignment layer (RoIAlign). The spatial information on the feature map is preserved by the bi-linear interpolation, which significantly eliminates the inaccuracy which is produced by the two quantization's of the feature map in the RoI Pooling layer and resolving the picture object regional mismatch problem. As a result, pixel-level detection segmentation is possible.

CI. Results and Analysis

The results were quite accurate even though a small dataset was used. We first run the algorithm through two random images from the val dataset to check the accuracy of the model, without considering the JSON file data. This can be compared with the actual image data by visualizing those particular images from the val dataset as we have done earlier. Even so the results were quite accurate. The results are can be seen below. The framework of the Mask RCNN network model is shown in Figure 5.

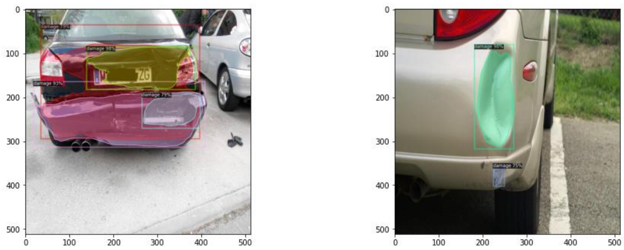

Figure 8: Two random images from Val dataset

Then we took a random image from the test dataset to check the damage detection. The results were very good. The test image can be seen below. It can bee seen that even though the wheel cap is detached from the car, the model still counts it as a damaged part and also the other small parts fallen around the car. This is very surprising considering the fact that a small dataset (of 59 images) was used for training the model. This is because of the pre-trained model which was used by us from the detectron2 library.

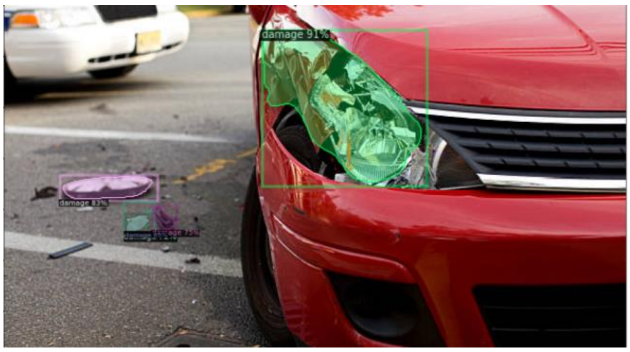

Figure 9: Random image from test dataset.

The accuracy of the model can be seen by the graphs plotted considering the Total Loss, Bounding Box Average Precision and Segmentation Average Precision. The graphs are shown below.

ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 8, August 2021

DOI 10.17148/IJARCCE.2021.10808

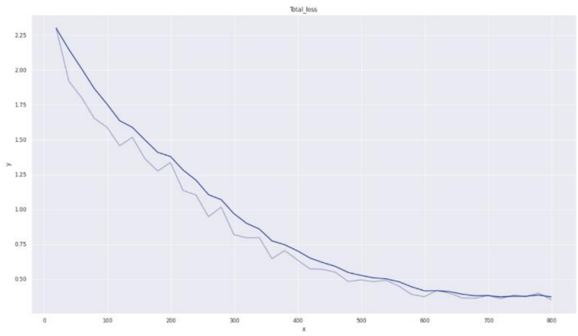

Figure 10: Total Loss Graph

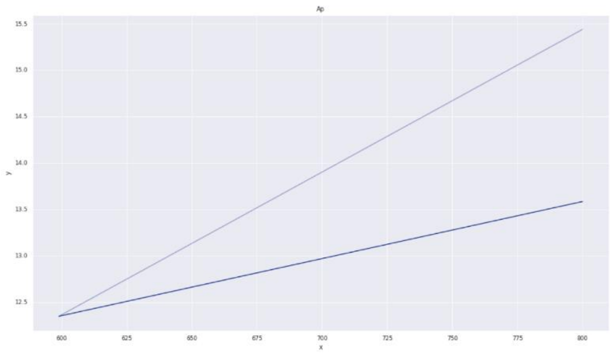

Figure 11: Bounding Box Average Precision Graph

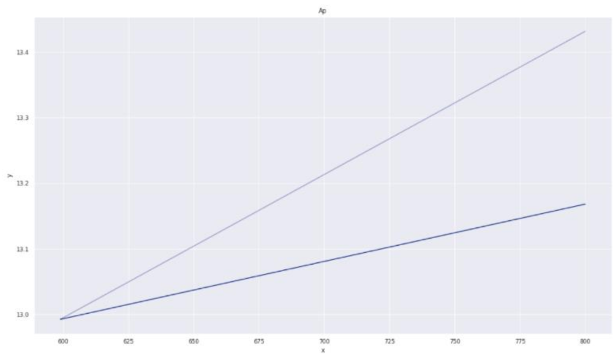

Figure 12: Segmentation Average Precision Graph

CII. CONCLUSION

To deal with the compensating problem of damaged autos, the model proposed here employs, a deep learning-based detection technique for vehicle-damage identification. The suggested approach of Mask RCNN and transfer learning-based damage detection of the vehicle is generic after testing, and can also better adapt to the diverse elements of damaged car images. Even though the model was trained on a very small dataset, good results were achieved.

Data extension can be done in the future to raise the dataset's size, gather additional automobile damage images under various degrees of illumination and weather conditions, enrich the data, the edge-contour enhancement of images can be improved and the damaged parts of the car can be masked more accurately. Also, the model can be further enhanced to predict the repair price of the damaged area by extracting the predicted part details like the segmented mask area.

The heading of the Acknowledgment section and the References section must not be numbered. Causal Productions wishes to acknowledge Michael Shell and other contributors for developing and maintaining the IEEE LaTeX style files which have been used in the preparation of this template.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 8, August 2021

DOI 10.17148/IJARCCE.2021.10808

REFERENCES

[1] Qinghui Zhang, Xianing Chang, and Shanfeng Bian, ”Vehicle-Damage-Detection Segmentation Algorithm based on Improved Mask RCNN”, IEEEAccess , doi:10.1109/ACCESS.2020.2964055.
[2] R. Girshick, J. Donahue, T. Darrell, and J. Malik, ‘‘Rich feature hierarchies for accurate object detection and semantic segmentation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, vol. 13, no. 1, pp. 580–587
[3] R. Girshick, ‘‘Fast R-CNN,’’ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015. pp. 1440–1448.
[4] S. Ren, K. He, R. Girshick, and J. Sun, ‘‘Faster R-CNN: Towards realtime object detection with region proposal networks,’’ IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/tpami.2016.2577031.
[5] W. Liu, D. Anguelov, and D. Erhan, ‘‘SSD: Single shot multibox detector,’’ in Proc. IEEE Eur. Conf. Comput. Vision, Jun. 2016, pp. 21–37.
[6] K. He, X. Zhang, S. Ren, and J. Sun, ‘‘Deep residual learning for image recognition,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.
[7] K. He, G. Gkioxari, P. Dollar, and R. Girshick, ‘‘Mask RCNN,’’ in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 2980–2988.
[8] N. Kumar and R. Verma, ‘‘A multi-organ nucleus segmentation challenge,’’ IEEE Trans. Med. Imag., vol. 11, no. 1, pp. 34–39, Oct. 2019, doi: 10.1109/TMI.2019.2947628.
[9] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. Rodrigues, ‘‘Identifying pneumonia in chest X-rays: A deep learning approach,’’ Measurement, vol. 145, pp. 511–518, Oct. 2019, doi: 10.1016/ j.measurement.2019.05.076
[10] P. Pinheiro and R. Collobert, ‘‘Learning to segment object candidates,’’ in Proc. Adv. Neural Inf. Process. Syst., 2015, pp. 1990–1998.
[11] W. Tang, H.-L. Liu, L. Chen, K. C. Tan, and Y.-M. Cheung, ‘‘Fast hypervolume approximation scheme based on a segmentation strategy,’’ Inf. Sci., vol. 509, pp. 320–342, Jan. 2020, doi: 10.1016/j.ins.2019.02.054
[12] Y. Li, H. Qi, J. Dai, X. Ji, and Y. Wei, ‘‘Fully convolutional instanceaware semantic segmentation,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 4438–4446
[13] X. Rong, C. Yi, and Y. Tian, ‘‘Unambiguous scene text segmentation with referring expression comprehension,’’ IEEE Trans. Image Process., vol. 29, pp. 591–601, Jul. 2019, doi: 10.1109/tip.2019.2930176.
[14] Y. L. Qiao, M. Truman, and S. Sukkarieh, ‘‘Cattle segmentation and contour extraction based on mask R-CNN for precision livestock farming,’’ Comput. Electron. Agricult., vol. 165, Oct. 2019, Art. no. 104958, doi: 10. 1016/j.compag.2019.104958
[15] C. Shuhong, Z. Shijun, and Z. Dianfan, ‘‘Water quality monitoring method based on feedback self correcting dense connected convolution network,’’ Neurocomputing, vol. 349, pp. 301–313, Jul. 2019, doi: 10.1016/j.neucom.2019.03.023
[16] J. Yang, L. Ji, X. Geng, X. Yang, and Y. Zhao, ‘‘Building detection in high spatial resolution remote sensing imagery with the U-rotation detection network,’’ Int. J. Remote Sens., vol. 40, no. 15, pp. 6036–6058, Aug. 2019, doi: 10.1080/01431161.2019.1587200.
[17] VEDALDI, A. AND FULKERSON, B., 2010. Vlfeat: An open and portable library of computer vision algorithms. In Proceedings of the international conference on Multimedia, 1469–1472. ACM.
[18] Najmeddine Dhieb, Hakim Ghazzai, Hichem Besbes, and Yehia Massoud. 2019. Extreme Gradient Boosting MachineLearning Algorithm For Safe Auto Insurance Operations. In2019 IEEE International Conference of Vehicular Electronicsand Safety (ICVES). IEEE, 1–5.
[19] Najmeddine Dhieb, Hakim Ghazzai, Hichem Besbes, and Yehia Massoud. 2019. A very deep transfer learning modelfor vehicle damage detection and localization. In2019 31st International Conference on Microelectronics (ICM). IEEE,158–161.
[20] Kalpesh Patil, Mandar Kulkarni, Anand Sriraman, and Shirish Karande. 2017. Deep learning based car damageclassification. In2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 50–54
[21] Pei Li, Bingyu Shen, and Weishan Dong. 2018. An anti-fraud system for car insurance claim based on visual evidence.arXiv preprint arXiv:1804.11207(2018).
[22] Mahavir Dwivedi, Malik Hashmat Shadab, SN Omkar, Edgar Bosco Monis, Bharat Khanna, and Satya Ranjan. [n.d.].Deep Learning Based Car Damage Classification and Detection. ([n. d.]
[23] Ranjodh Singh, Meghna P Ayyar, Tata Sri Pavan, Sandeep Gosain, and Rajiv Ratn Shah. 2019. Automating Car InsuranceClaims Using Deep Learning Techniques. In2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM).IEEE, 199–207.

