Medicon Engineering Themes
Volume 5 Issue 5 November 2023
Article Type: Research Article
ISSN: 2834-7218

Car Damage Detection Based on Mask Scoring RCNN

Farah Oubelkas[1]*, Lahcen Moumoun[1,2], Abdellah Jamali[1,2]

[1]*Hassan First University of Settat, Faculté Sciences et Technique, Settat, Morocco*

[2]*Laboratoire de Recherche Informatique, Réseaux, Mobilité et Modélisation, Settat, Morocco*

***Corresponding Author:** Farah Oubelkas, Hassan First University of Settat, Faculté Sciences et Technique, Settat, Morocco.

Received: July 19, 2023; **Published:** November 02, 2023

Abstract

The growth of the car industry is now closely linked to the rise of the number of car accidents. As a result, insurance companies must deal with several claims at the same time while also addressing claims leakage. To resolve these issues, we propose a car damage detection system based on Mask Scoring RCNN. The experiment first makes a dataset by collecting car damage pictures of different types and on different angels for pre-processing then use Mask scoring RCNN for training. It is envisaged that this method would assist insurance in correctly classifying the damage and reducing the time spent on damage detection. The test results demonstrate that the proposed system has better masking accuracy in the case of complex images, allowing the car damage detection duty to be completed swiftly and easily.

Keywords: Computer vision; Mask Scoring RCNN; Car-damage detection; Mask detection accuracy

Introduction

We live in the big data age, in which vast volumes of data are generated across all fields of research and industry. Nowadays, Deep learning is an innovative approach that is now gaining a lot of attention. Deep learning has also been effectively used to a variety of computer vision application challenges. One of the key research topis nowadays in computer vision is object detection. RCNN [1], Fast RCNN [3], Faster RCNN [10], Mask RCNN [1] are now the most common detection algorithms. However, they need a substantial quantity of training data, that is sometimes challenging to have. The detection frame's positioning ability is restricted, and as the number of convolution layers rises, gradient vanishing problem frequently occurs. In order to resolve all these drawbacks, Mask Scoring RCNN was proposed [12].

In this paper, we propose a car damage segmentation and detection system based on Mask Scoring RCNN algorithm to mark the car damaged areas. This paper improves the default model's architecture by optimizing the residual network (ResNet), adjusting the hyperparameters and the parameters of the anchor box in order to improve the accuracy of the model. This proposed system can be used by insurance companies to process claims rapidly.

Related work

There are two types of instance segmentation methods now in use. The first type is based on detection, and the second one is based on segmentation. Detection based approaches use detectors to get the region of each instance and then predict the mask for each region. An example of such detectors is Faster RCNN. Mask RCNN is an instance segmentation algorithm that was developed based on Faster RCNN. It is actually an improved version of Faster RCNN. It can be used in different types of problematics: Object detection, semantic segmentation or instance segmentation.

It was proposed first by He et al [5]. Based on this new algorithm, Chen et al. (Liang- Chieh Chen, 2018) [8] introduce MaskLab; A new approach that improves the result of Mask RCNN by using position-sensitive scores. However, the mask instance segmentation isn't perfect, and certain locations where damage isn't visible can't be segregated. The major problem of all these approaches sited above is that the score of the mask is based on the classification precision which is not accurate, resulting in a huge discrepancy in the prediction results.

Segmentation based methods first employ pixel-level prediction to predict the category label and then use a clustering algorithm to group the pixels into various instances. To cluster the pixels [13] adopt spectral clustering. On the other hand, Baiet al. [11] utilized watershed algorithm to categorize pixels and forecast pixel-level energy levels. All these methods use the average pixel-level classification score to measure the quality of the instance mask. It is possible that the detection results are great, the classification score is high but the quality mask is low. Background clutter and occlusion can also contribute to this discrepancy. That is when detection score correction methods appeared. Their main focus is to correct the classification score for the detection box. Cheng et al utilize a separated network to rectify the false positives samples' score. SoftNMS [2] corrects the low score box by using the overlap between two boxes. While Tychsen-Smith et al. [6] propose Fitness NMS, a method that uses the IOU between the identified bounding boxes and their ground truth to adjust the detection score. Huang Z, Huang L, Gong Y et al [12] propose Mask Scoring RCNN. The key distinction between the two methods is that the first one formulates box IOU prediction as a classification task while Mask Scoring RCNN formulates it as a regression task. Mask scoring RCNN is the method implemented in this paper in order to detect the car damage. It solves the problem we have with Mask RCNN which is using the classification confidence to measure the score of the mask. Instead, it proposes a new evaluation method by adding Mask IOU Head. It takes the ROI features and the predicted mask as an input to obtain the score of the model.

Compared to other traditional detection methods, Mask scoring RCNN is extremely effective. It has been used to detect and classify domestic garbage [9], in the medical Field to detect Breast tumor [7], in Agriculture Field to detect apple flowers [14]. But it never been used in the field of car damage detection before. In this paper, we applied Mask Scoring RCNN to detect and segment the area of the damage. This can be useful for insurance companies to automate the process.

Proposed method

The car damage detection and segmentation model proposed based on Mask scoring RCNN implemented in this paper is shown in the figure 1 bellow.

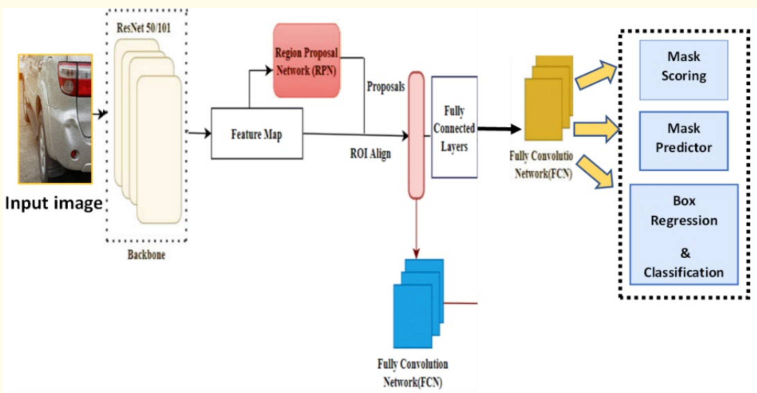

Figure 1: The flow of the car damage detection and segmentation system.

The first step of the flow process is the collection of the image. The image is labelled in Coco json format using LabelMe annotation tool. The image is fed to the Mask scoring RCNN for feature extraction, classification and segmentation. As an output, we get the car damage detection and the mask predicted.

Mask Scoring RCNN

Mask Scoring RCNN is an instance segmentation framework, a Mask-RCNN with an additional MaskIOU head that predicts the Mask-IoU score.

Figure 2 depicts the four steps of our proposed method based on Mask Scoring RCNN. The first step, known as feature extraction, employs ResNet-50+FPN backbone architecture to obtain the corresponding feature map. The second step, known as regions of interest (RoIs) generation, extracts RoIs by RPN. The third step extracts RoI features via RoIAlign then executes frame regression, classification using softmax and a predicted mask by FCN. The last step known as MaskIoU head, aims to regress the IoU between the predicted mask and its ground truth mask.

The goal of MaskIoU head is to regress the IoU between the predicted and the ground truth mask. In order to make the predicted mask have the same special size as the ROI feature, we employ a max pooling layer with a kernel size of 2 and a stride of 2 while concatenating. The Mask IoU head is made up of four convolutions layers and three fully connected layers. For the four convolution layers, we use Mask head and set the kernel size and filter number to 3 and 256 for each convolutional layer. For the three fully connected layers, we used the RCNN head and set the outputs to 1024 for the first two fully connected layers while the output of the final fully connected to five because we have 5 classes.

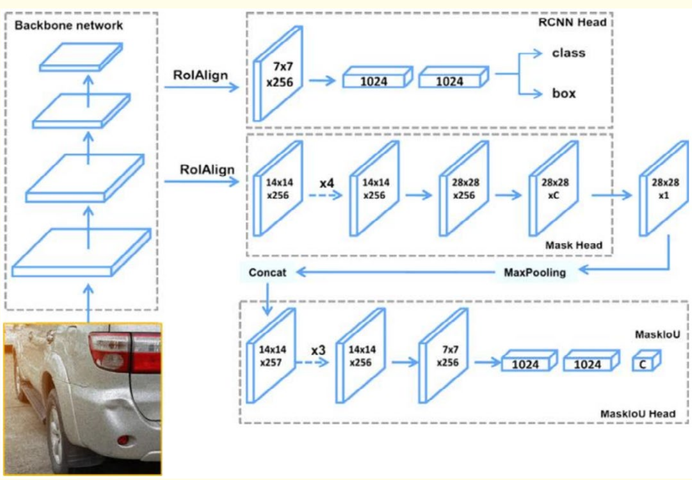

Figure 2: Mask Scoring RCNN architecture.

Transfer learning

In this paper, we have used the transfer learning method to shorten training time and avoid overfitting of the model. In addition to that, it was difficult to collect enough training data of different type of damages. We had a small dataset compared with the giant open-source datasets we can find online.

Transfer learning is to transfer the knowledge gained from training a model in the source domain to new models to solve a target task. In this paper, the ResNet backbone network, a model trained on Coco Dataset, is used. We remove the last layers of the pretrained model and replace them with untrained layers.

The default backbone network of Mask RCNN is ResNet101. However, if there are too many layers, the rate of the network structure will severely be slow. Instead, in this experiment, we used ResNet-50 based FPN network, that is, the number of layers: 50 in order to improve the running speed of the algorithm and the generalization performance of the model.

The improved Resnet backbone architecture as shown in the figure 3 is now instead of using post-activation, it focuses more on using the pre-activation of weights layers. This has two advantages. First, back-propagation satisfies the requirement, and information transmission is unaffected. Second, the BN layer serves as a pre-activation layer, with the term 'pre' referring to the weight of the convolutional layer. This improves the model's regularization, as well as its generalization performance.

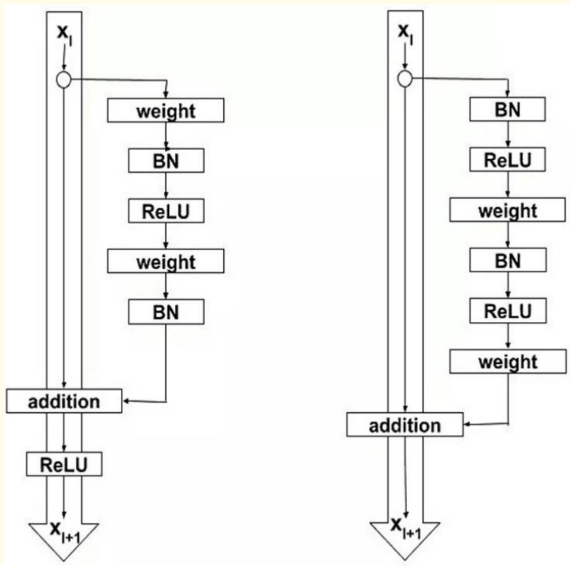

Figure 3: ResNet V1 and ResNet V2 structure.

Results and Discussion

Dataset

528 damaged car images were collected (for train and for test) from Google and daily photographs. Each image has a maximum of 5 damage types and a minimum of 1 type of damage. We tried to have different angles and different type of damages (headlamp, hood, rear bumper, front bumper and door). However, rear bumper damage is rare and has a few kinds, therefore it makes up just around 14% of the overall dataset. We had images of different sizes. In order to feed them to Mask Scoring RCNN, we had to normalize them using a script to 1024 * 1024 pixels.

To train Mask Scoring RCNN, we didn't just need the images but also the corresponding masks. The annotation of the dataset is done using LabelMe, it's an open-source image annotation software, to label the damage area and make a mask to get the segmentation information of the damage in the car, as shown in the figure 4.

Training platform

The model training is done in Ubuntu 64bit and a hardware environment of NVIDIA GeForce MX250, memory 16Go. The virtual environment must be as follows:

- PyTorch ≥ 1.3.

Figure 4: Labelme annotation image.

- Python ≥ 3.6.
- torchvision that matches the PyTorch installation.
- OpenCV.
- Pycocotools.
- gcc & g++≥ 4.9.

475 images were used for training and 52 images for testing. To better represent the benefits of Mask Scoring RCNN, in this paper, we used the same dataset to compare the results obtained using Mask Scoring RCNN and its precursor Mask RCNN.

Evaluation Metric

To evaluate the performance of the proposed model, we use the mean intersection over union (MIoU).

$$\text{Precision} = \frac{TP}{TP+FP} \quad , \quad \text{Recall} = \frac{TP}{TP+FN}$$

Where TP is true positive, FP is false-positive and TN is true negative.

Results

52 car damage pictures were used to evaluate the model. Here is an example of the test results of Mask Scoring RCNN as shown in Figure bellow:

The damage to be detected in the image has been precisely circled and labelled with the category of the damage.

In the figure 5 Result 1, we have only one type of damage. It is correctly identified and marked. Result 2 is more challenging, there are different types of damage, and the two types (head lamp and front bumper) are intertwined. However, the segmentation and the identification may be demonstrated to be quite excellent. The two types of damage are correctly classified with a high accuracy score.

After the analysis of the results of test dataset, it can be seen that the more complex the image, the lower the segmentation result. When the image's background is perfect, even though we have intertwined different type of damages, the segmentation is good around 96%. Figure 6 identifies the accuracy of test dataset.

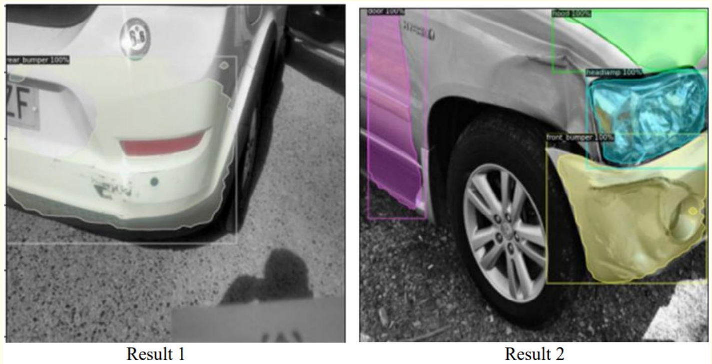

Figure 5: Mask Scoring RCNN test results.

Mask Scoring RCNN Accuracy

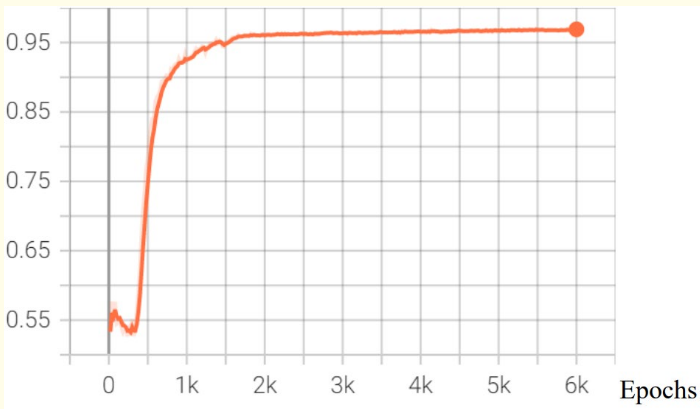

Figure 6: Accuracy of Mask Scoring RCNN.

Comparison with Mask RCNN

In order to reflect the high detection performance of Mask Scoring RCNN, we compared its results with its predecessor Mask RCNN. The different types of damages to be identified in the image is overlapping.

The result of Mask Scoring RCNN in Figure 7 shows that, while the different types of damages that exists in the image are not totally separated, the damage still can be detected and the classification is correct. While using the Mask RCNN, we can clearly notice that the classification and the segmentation was more challenging. The segmentation recognition becomes worse and the confidence score of classification is lower. Therefore, Mask Scoring RCNN is better.

In general, we can see that Mask RCNN method has great identification accuracy but it can't handle overlapping targets very well, while Mask Scoring RCNN is capable of completing both tasks: instance segmentation and classification perfectly, and has the most optimal output.

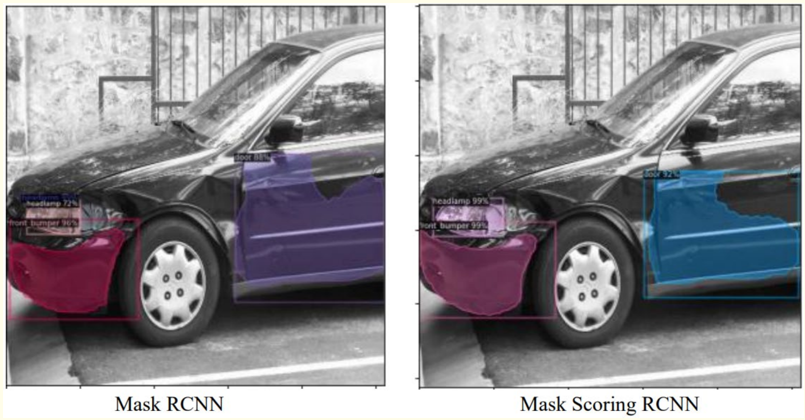

Figure 7: Comparison of test results between Mask RCNN and Mask Scoring RCNN.

It can be seen from Table 1 that the proposed Method of Mask Scoring RCNN outperforms Mask RCNN. As the table shows, the mask accuracy of Mask Scoring RCNN is 96.9% which is higher than Mask RCNN mask accuracy by 2.03%. In addition to that, we can notice that Mask Scoring RCNN is way faster than Mask RCNN which makes it a better choice than Mask RCNN.

Algorithm	***Mask Accuracy (MIoU) (%)***	***Total loss (%)***	***Running speed (fps)***
Mask RCNN	95.93	0.198	4.27
Mask Scoring RCNN	96.9	0.181	4.71

Table 1: Comparison of test results (Mask RCNN and Mask Scoring RCNN).

The accuracy of test results was encouraging; we had an accuracy rate of 70% in headlamp, 67,5% in door, 68% in hood, 64.6% in rear bumper, 71% in front bumper. Therefore, an average of 70% on the whole dataset.

Conclusion

To resolve traffic accident compensation quickly, a detection algorithm based on deep learning and transfer learning is used. After testing, we demonstrate that the proposed car damage classification system based in Mask Scoring RCNN achieved great segmentation, classification and recognition results in different scenarios. However, there is always a room for improvement. In future work, having more data can definitely improve the training results and show the power of this chosen algorithm.

References

1. Abdulla W. "Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow". Github (2017).
2. Davis NB. "Soft-NMS -- Improving Object Detection with One Line of Code". Proceedings of the IEEE International Conference on Computer Vision (2017).
3. Girshick R. Fast R-CNN. International Conference on Computer Vision ({ICCV}) (2015).
4. Girshick Ra. Rich feature hierarchies for accurate object detection and semantic segmentation (2014).
5. K He GG. "Mask RCNN". ICCV (2017): 2980-2988.
6. Lachlan Tychsen-Smith, L. P. (n.d.). "Improving Object Localization with Fitness NMS and Bounded IoU Loss". arXiv, p. 1711.00164.
7. Lei YHX. "Breast tumor segmentation in 3D automatic breast ultrasound using Mask scoring R-CNN". Medical Physics (2021).

8. Liang-Chieh Chen AH. “MaskLab: Instance Segmentation by Refining Object Detection with Semantic and Direction Features”. CVPR (2018).
9. S Li MY. “Garbage object recognition and classification based on Mask Scoring RCNN”. International Conference on Culture-oriented Science & Technology (ICCST) (2020).
10. Sun SR. “Faster {R-CNN}: Towards Real-Time Object Detection”. Advances in Neural Information Processing Systems ({NIPS}) (2015).
11. Urtasun MB. “Deep watershed transform for in-stance segmentation”. IEEE Conference on Computer Vi-sion and Pattern Recognition (2017): 2858-2866.
12. Wang ZH. “Mask Scoring R-CNN”. CVPR (2019).
13. X Liang YW. “Proposal-free network for instance-level object segmentation”. In arXiv preprint arXi (2015): 1509.02636.
14. Yunong Tian GY. “Instance segmentation of apple flowers using the improved mask R-CNN model”. Biosystems Engineering (2020): 264-278.
15. ZhaBian SCX. “Vehicle-Damage-Detection Segmentation Algorithm Based on Improved Mask RCNN”. IEEE Access (2020): 6997-7004.

Volume 5 Issue 5 November 2023