Automatic Car Damage Assessment System: Reading and Understanding Videos as Professional Insurance Inspectors

Wei Zhang, Yuan Cheng, Xin Guo, Qingpei Guo, Jian Wang,
Qing Wang, Chen Jiang, Meng Wang, Furong Xu, Wei Chu
Ant Financial Services Group, Z Space No. 556 Xixi Road Hangzhou, 310099, Zhejiang, China
{ivy.zw, chengyuan.c, bangzhu.gx, qingpei.gqp, bobblair.wj, wq176625, qichen.jc, darren.wm}@antfin.com,
{booyoungxu.xfr, weichu.cw}@antfin.com

Abstract

We demonstrate a car damage assessment system in car insurance field based on artificial intelligence techniques, which can exempt insurance inspectors from checking cars on site and help people without professional knowledge to evaluate car damages when accidents happen. Unlike existing approaches, we utilize videos instead of photos to interact with users to make the whole procedure as simple as possible. We adopt object and video detection and segmentation techniques in computer vision, and take advantage of multiple frames extracted from videos to achieve high damage recognition accuracy. The system uploads video streams captured by mobile devices, recognizes car damage on the cloud asynchronously and then returns damaged components and repair costs to users. The system evaluates car damages and returns results automatically and effectively in seconds, which reduces laboratory costs and decreases insurance claim time significantly.

Introduction

Capturing and uploading photos or videos by mobile devices is convenient today. If the domain knowledge of damage assessment can be formulated into a computer vision problem , photos or videos acquired by users can be processed automatically in car insurance field. In traditional ways, insurance companies have required inspectors to investigate at the accident scene, which takes hours or days long. We demonstrate an automatic car damage assessment system based on computer vision and deep learning, which completes claim processing in seconds and saves laboratory. The users without professional knowledge can also use our system to evaluate car damages when accident happens.

Recently, there are methods that formulate damage assessment as an image content analysis problem. Image classification utilized by (Jeffrey 2018) cannot handle tiny damages and multiple damaged components effectively. Retrieving similar damage photos from historical database(Tofte 2016) is limited to variety of database and tends to be affected by car appearances and shooting angle. Using 3D computer aided design model (Jayawardena 2013) to recognize components is hard to be applied to different kinds of cars with variant appearances. There are also companies announced that they can recognize damaged components and damage degree from images[1] [2] [3]. However, these methods are based on images not videos. On one hand photos taken distantly are hard to recognize damage details and closely are hard to recognize components, thus it requires users shooting more than once to get accurate results while our system using videos has less interaction with users.On the other hand videos containing much more frames can depress noises and improve recall compared with photos. Remote video conference systems (Huang et al. 2017) utilizing videos captured by smart phones and completing claims remotely, which still requires professional engineers to watch videos and make decisions at the service center. Our system evaluates damage from uploaded videos automatically.

We conduct experiments to compare damage assessment accuracy between acquiring videos and images with different app and more than 1000 cases each. The average accuracy of videos is 29.1% higher and the ratio of high quality shooting data on predefined criterion is also 20% higher.

System Framework and Technique Details

The challenges of automatic damage assessment and our solutions are in the following aspects:

- High quality videos are hard to acquire since unpredictable user behaviours lead to problems like defocus, arbitrary shooting angle and extreme car component scales. Our system acquires videos through a front end interaction module to guide users to shoot high quality and proper distance videos.
- Reflections, mud and covers on the car exterior can be misidentified as damages. With model fusion of multiple frames extracted from videos, noise is depressed and more accurate results are recalled.
- Car components and damages need to be segmented in pixel level to accurately localize damages. A challenge of segmentation is that annotations are much more costly

[1]tractable.ai/products/car-accidents
[2]tonkabi.com/artificial-intelligence
[3]www.altoros.com/car-damage-recognition

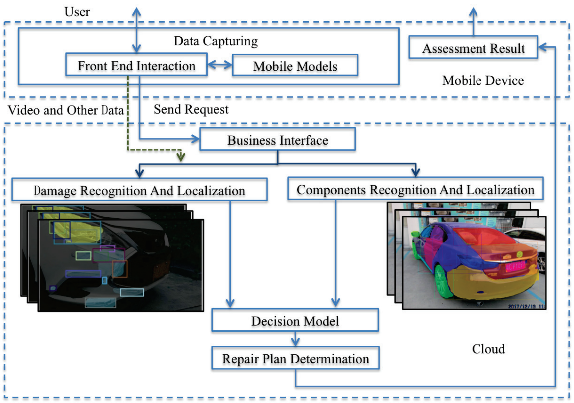

Figure 1: Overview of System Framework

than object detection. A weakly supervised segmentation model is proposed to improve damage localization accuracy using large scale bounding box annotations.

Data Capturing The front end interaction sub-module is designed to get high quality videos. It uses mobile inference engine to invoke lite deep learning models deployed on mobile devices to guide the shooting process. Firstly plate and VIN code should appear in the video and the contents are recognized automatically to verify the identity of car owners and filter frauds. Secondly users are guided to shoot videos at a farther distance and then closer to capture damage details. The distance is determined by a mobile classification model. During shooting, videos and other information are uploaded to the cloud asynchronously for further process.

Damage Recognition Car damage recognition can be formulated as multi-class detection or segmentation. A large quantity of variant samples are annotated to train detection models. This is unpractical for segmentation whose annotation costs 5 times more than detection. Moreover, it is hard to define boundaries of damage textures in many cases. We adopt a two-stage detection model (Liu et al. 2019) and add a weakly supervised semantic segmentation branch to get accurate pixel level results from bounding box annotations. The predictions of detection and segmentation branches are fused further. Furthermore, multiple frames instead of single image are analyzed to utilize time consistency and content complementarity of videos (Wang et al. 2018) to filter noises which look like damages and improve recall.

Damaged Component Localization Frames shooted at a farther distance are better for car component recognition while closer ones are better for damage detail recognition. Thus both kinds of frames are extracted from videos automatically by a frame selection algorithm instead of shooting multiple photos by users. We improve MASK R-CNN(He et al. 2017) to detect damaged components and segment accurate component boundaries simultaneously. he results of multiple frames are fused to accurately localize damaged components.

Decision and Repair Plan Determination Given both damage and component recognition results as the input, decision module learns and predicts the final damaged components and damage severity. The recognition and decision modules are not designed to be end-to-end, since in each single stage multiple models can be ensembled to improve accuracy. Then the system transforms algorithm results into repair plan based on pricing algorithm.

Use Cases

When a car accident occurred, car owner can assess damages using our system. Specifically, the user following shooting guidance should scan the plate and then VIN code to verify the consistency of car identity and its damages. Then the user would shoot a video of damaged components at both farther and nearer distances.During the capturing process, the video is uploaded to cloud to recognize damages. Finally damaged results are returned to the user. The damaged results are also presented with estimated prices and the user can decided to make a claim or repair personally.

Conclusions

We demonstrate an automatic car damage assessment system based on computer vision techniques. We analyze the challenges of acquisition of high quality vision data , promotion of damage recognition accuracy and reduction of large scale annotation costs and give solutions respectively. Practical applications have shown that our system is effective and can reduce the time cycle of making claims significantly.

Acknowledgments

The system is developed by the Vision Cognition Team since 2017, which is supported by Cognition Computing and Knowledge Graph Group of Artificial Intelligence Department and Insurance Business Group of the Ant Financial Services Group.

References

He, K.; Gkioxari, G.; Dollár, P.; and Girshick, R. 2017. Mask r-cnn. In *ICCV 2017*, 2980–2988.

Huang, Z.; Wu, T.; Wang, J.; and Xiao, J. 2017. A method and system of car insurance claims. CN patent 1062203644A.

Jayawardena, S. 2013. *Image based automatic vehicle damage detection*. Ph.d. diss., Australian National University.

Jeffrey, D. 2018. Automatic Car Damage Recognition using Convolutional Neural Networks. Technical report, MSc Business Analytics.

Liu, L.; Ouyang, W.; Wang, X.; Fieguth, P.; Chen, J.; Liu, X.; and Pietikäinen, M. 2019. Deep learning for generic object detection: A survey. *IJCV 2019*.

Tofte, N. 2016. System and method for assessing a claim using an inspection vehicle. US Patent 9262789.

Wang, S.; Zhou, Y.; Yan, J.; and Deng, Z. 2018. Fully motion-aware network for video object detection. In *ECCV 2018*, 557–573.