APPLY IMAGE DATA GENERATOR FUNCTION

Image recognition in auto damage claim process:

INTRODUCTION:

Research background:

The current situation of car insurance industry has a huge waste on claims' leakage. Claim leakage is the overpay in claim payment due to visual inspection (IRJET 2020). And according to the research of McKinsey (2018), claims processing in 2030 remains a primary function of insurance carriers but head count associated with manual claims is reduced by 70-90% compared with 2018 levels. Therefore, the urgency for claim automation is rising in all aspects of insurance**.**

Concept and delimitations:

Image recognition, or computer vision, is the machine or computer's ability to detect an object, a feature or a useful information from an image or a sequence of images, such as video. Auto insurance, according to insurance information institution, is the agreement between the insurance company and its customer the protect him.

Research objectives and questions:

Thus, the foremost research question is how does image recognition revolutionize damage claim process in auto insurance. In order to give a comprehensive answer to the big question, there are supporting sub questions to be addressed within the area of image recognition and auto claim in the relativeness to the traditional method: -

What is image recognition and its mechanism?

What values does image recognition contribute?

What is key to success and future outlook of image recognition?

Image recognition:

As in image recognition, on the object-level types, there are seven popular possibilities or models according to the Claim Genius (2020) including.

- Classifying objects specifies the broad category of the object in the image.
- Identifying object recognizes the type of the object in the image.
- Verifying object confirms whether the object
- Object landmark detection presents the key points of the object.
- Segmentation of the objects determines the pixels that belong to the object.
- Object recognition: that names the objects are in the image.

Each of the above possible uses are applied in a certain areas of business processes. Some of the procedures will require the combination of a few above methods in order to maximize the functionality.

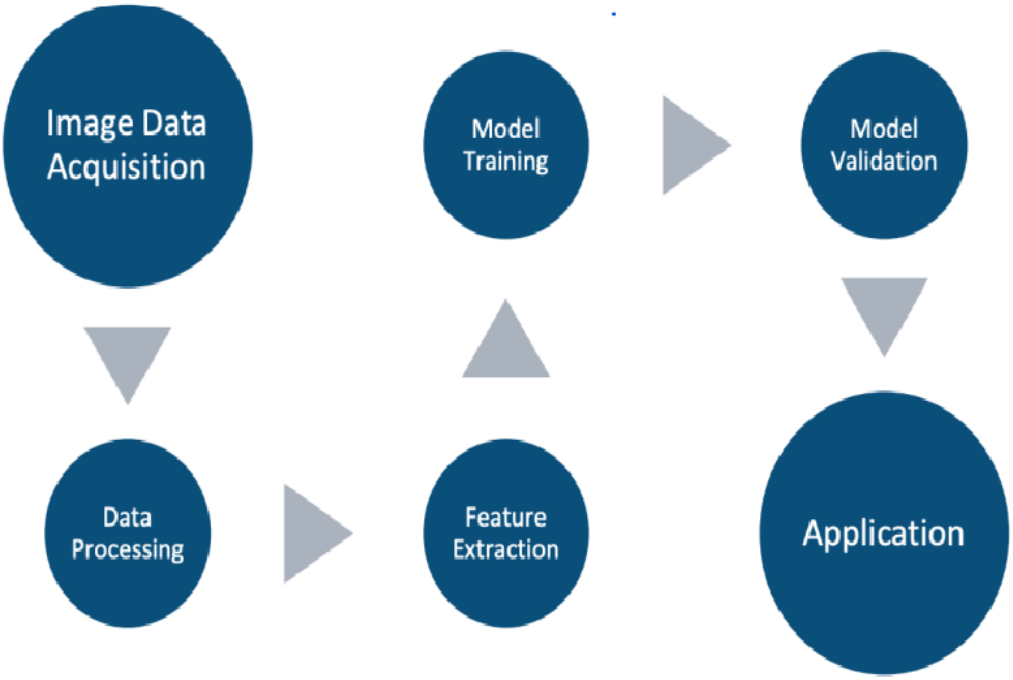


Figure 1 Image recognition process (Society of Actuaries 2018)

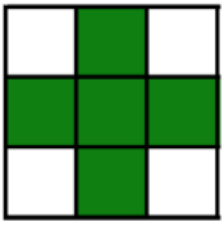
Though the usual process of image recognition application contains various steps of complexity, the overview is similar to traditional statistical analysis. From the figure 1, the flow of data is transparent from collecting with the aid of specialized popularly camera, drone or satellite.

Image data acquisition:

According to the article of Mishra & Kumar (2017), success of image acquisition depends purely on Hardware Process, that converts the reflected light from the object into electrons.

Data processing:

Society of Actuaries (2018) states that when the images are received, the very first step is to transform the image into the form that the computer could comprehend, which are numbers. A picture is technically made from pixels, the basic unit in digital imaging. Each pixel is defined by a colour, which is specified by a colour imaging system. The RGB colouring system is believed to be most popular and define any colour in the combination of three dimensions



=

(0,0,0)	(0,128,0)	(0,0,0)
(0,128,0)	(0,128,0)	(0,128,0)
(0,0,0)	(0,128,0)	(0,0,0)

Figure 2 Image data representation (Society of Actuaries 2018)

Additionally, besides RGB colouring system, there are other lights that can't be seen by human eyes or captured by normal cameras such as infrared, x ray, etc. Image recognition tasks are relatively similar to quantitative analysis when it deals with mostly numbers encoded from the pixels.

Figure 3 Original image (Hype 2016)

Another solution could be downsizing the original image to lower resolution. This transformation could reduce not only the storage requirement but also Figure 3 Original image (Hype 2016) 7 increase the strength of computer processing due to less details, which leads to improved computing speed and administrative burden.

Figure 5 Greyscale image (adapted from Hype 2016)

Figure 4 Compressed image (adapted from Hype 2016)

In other case image could be blurred by adjusting the pixels according to the values of its neighbours. The images could also be sharpened to enhance the edges' contrast. With any type of transformation may be applied, the images are meant for improving model input and the accuracy of image recognition can't be compromised.

Figure 7 Blurred image (adapted from Hype 2016)

Figure 6 Sharpened image (adapted from Hype 2016)

Beside changing the content, data augmentation could also be possible before entering the model. Usual augmentations include flipping, rotating, shifting, cropping, and brightening.

Flipped

Rotated

Cropped

Brightened

Figure 8 Image augmentation (adapted from Hype 2016)

For a more advanced method, feature can be extracted by measuring the changes from pixels. By detecting the edges, the machine will get an idea of patterns needed to generate meaningful recognitions. (Society of Actuaries 2018.)

2.3 Model training:

During this step, the data collected and processed are used as the input to train the model in order to retrieve the meaningful information. Mostly successful image Figure 8 Image augmentation (adapted from Hype 2016) 9 recognitions models are developed from

convolutional neural network (CNNs). Through image recognition models, there are many similarities to the traditional underwriting models.

Neural networks:

Neural networks are the tools of machine learning. Within the networks, the computer learns to perform a certain purpose by analysing examples as a part of training process. The examples in this context are images for recognition purposes. (MIT 2017.) Rather than statistical models, neural networks are believed to be more important in actuarial analysis

Convolutional neural network (CNNs):

Convolutional Neural Networks (CNNs) are advance specifying a certain area of interest in a certain image instead of spreading the analysis over the whole as Fully Connected Neural Network. CNNs are now widely applied for model training in image recognition by reflecting the local connectivity of pixels, imitating the human subconscious mind while looking for objects and clues within a big picture.

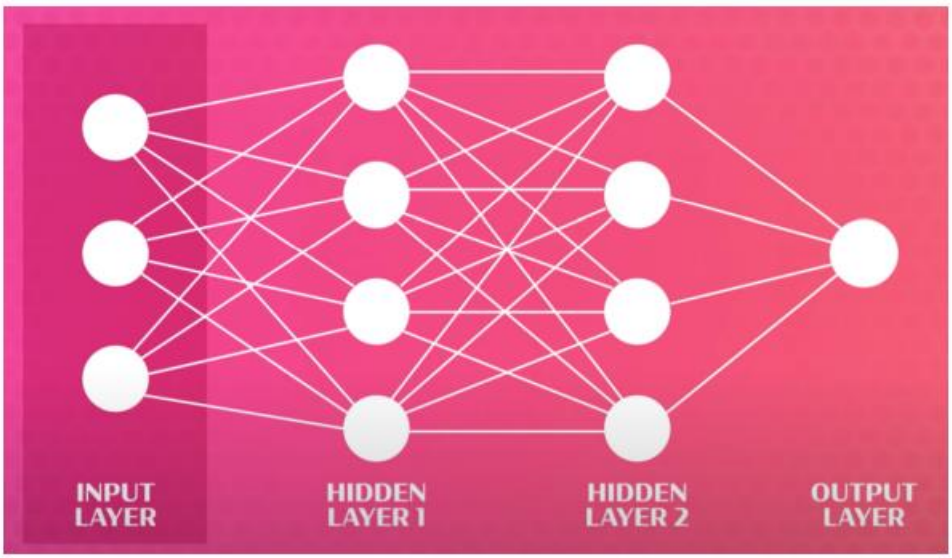
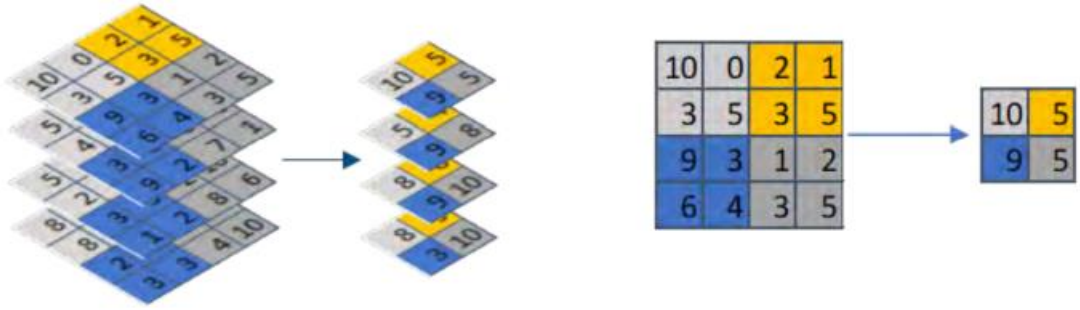


Figure 8 Neural networks (Crash course 2019)

According to Crash course (2019), The figure 9 is a simple and classic illustration of convolutional neural networks. The input layer is the place where neural networks would receive the data in form of numbers as described during the data processing phase.

Activation function and Pooling:



. There are other pooling methods worth mentioning such as average pooling by extracting the average value or stochastic pooling which uses the random element from the set. (Society of Actuaries 2018.)

2.4 Validation:

The accuracy of the model depends solely on the dataset. The variety of data set in terms of angels, brightness, background and other variables, will contribute to the accuracy rate. On another hand, while the data is not sufficient, the accuracy of application also fail to perform. (Society of Actuaries 2018.)

3.Auto insurance applications

3.1 General evolution

In general, the application of image recognition would increase the proceeding of auto damage claim by reducing the steps and the processing time thanks to the technology. The figure 12 is the structured process of traditional claim:

Figure 12 Traditional damage claim process (Altoros 2020)

For the traditional method, the claim process will undergo total six steps to final insurance payment. The very first one is claim submission with images together with the receipt submission, showing the costs of repair. All the information is manually transferred into the system for analysis and records.

Figure 13 Applied image recognition in damage claim process (Altoros 2020)

In the assisted services, the insurer would connect with the customer and share the cause and severity of the damages, which will provide the most reliable and images as required. On another hand, the self-services rely completely on the submitted images which will lower the processing time to payment decision. (Altoros 2020.)

3.2 Information extraction

Continuing the advantages of image recognition in details, the below subsections will clarify how image recognition benefit the auto insurance industry in terms of technical aspect. Each of the functionality would play a role in the claim process, from retrieving auto basic information, detecting the damaged components and its extent.

Auto basic information According to Databricks (2020), starting from 2012, the Deep Learning Revolution enabled image recognition to classify car images into specific car models with steps described in the figure 14. The car models as basic information set the standard cost estimation of repair and replacement due to the noticeable difference in pricing.

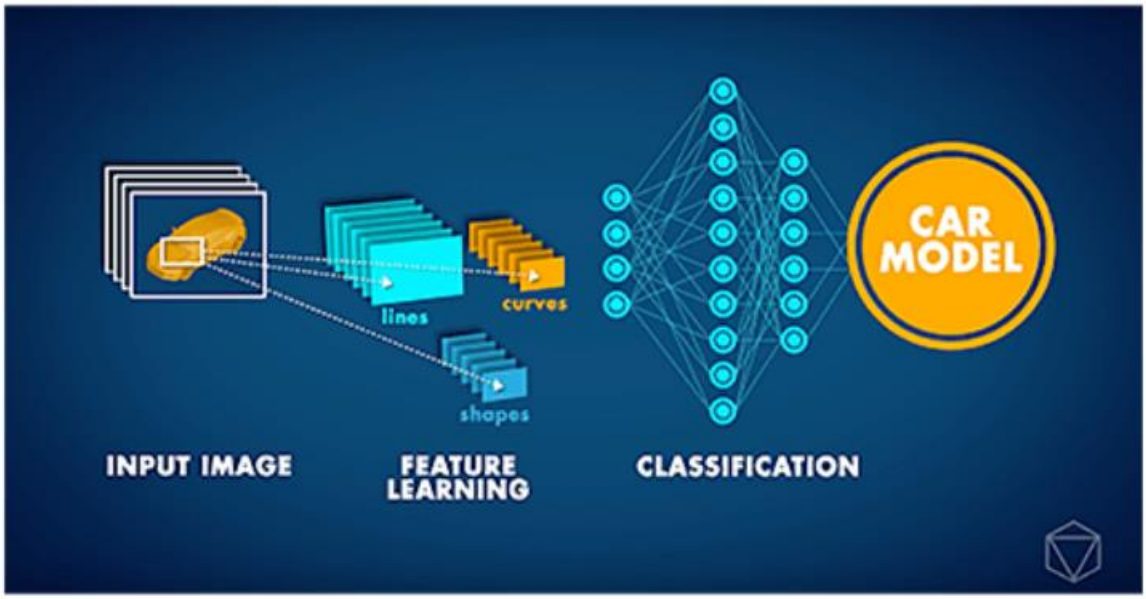


Figure 14 Neural network for car model classification (Databricks 2020)

The locations of damaged components:

The car images are processed within the neural network. The neutral network in this phase would be trained from Transfer Learning and Object Detection Algorithms. Transfer Learning is a machine learning method where a model is developed for one task and reused for a new but related task (Hussain et al. 2018). Object detection, as mentioned earlier, is the act of localizing a specific object in the image. In most cases, the damages would be manually proceeded for evaluation of damages (Towards data science 2019). However, the latest version of image recognition model is able to detect the damage location on the car body and estimate the severity of damages with little human intervention.

Loading our data and performing Data Augmentation

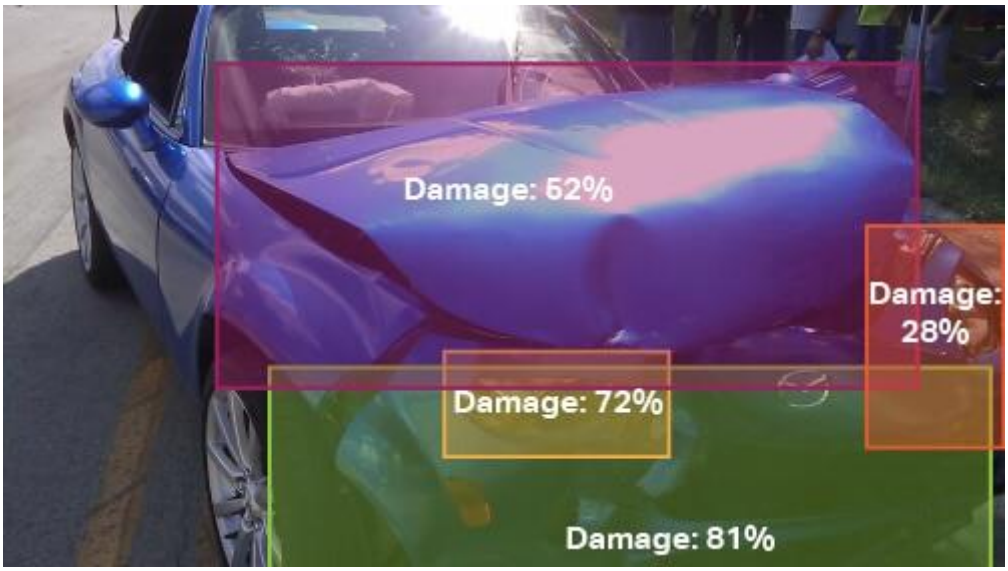
For Body Damage:

```
training_set = train_datagen.flow_from_directory(trainPath,
                                                 target_size = (224, 224),
                                                 batch_size = 10,
                                                 class_mode = 'categorical')

test_set = test_datagen.flow_from_directory(testPath,
                                            target_size = (224, 224),
                                            batch_size = 10,
                                            class_mode = 'categorical')
```

For the level of Damage:

```
training_set = train_datagen.flow_from_directory(trainPath,
                                                 target_size = (224, 224),
                                                 batch_size = 10,
                                                 class_mode = 'categorical')

test_set = test_datagen.flow_from_directory(testPath,
                                            target_size = (224, 224),
                                            batch_size = 10,
                                            class_mode = 'categorical')
```

We notice that 979 images are belonging to 3 classes for training and 171 images belong to 3 classes for testing purposes.



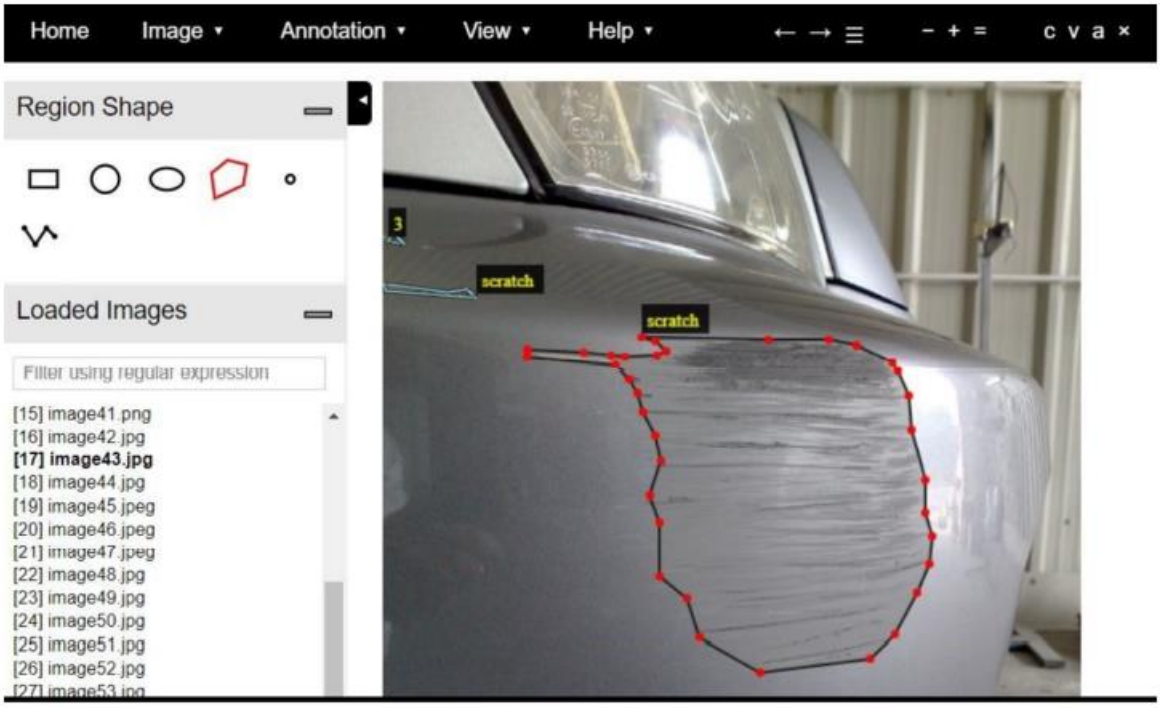


Figure 15 Sample of damaged component (Towards data science 2019)

The figure 15 has shown a vivid example of a Convolutional neutral network, especially VGG image Annotator (VIA).

The author has submitted images and chosen the mask shape on the top left corner to mark the edges of the damaged part for training purposes.

The case model later is input with more datasets including both images and annotations for training and validation purposes.

The continuation of repetitive training would enhance the accuracy of the model.

The figure 15 has shown a vivid example of a Convolutional neutral network, especially VGG image Annotator (VIA).

The author has submitted images and chosen the mask shape on the top left corner to mark the edges of the damaged part for training purposes.

The case model later is input with more datasets including both images and annotations for training and validation purposes.

The continuation of repetitive training would enhance the accuracy of the model.

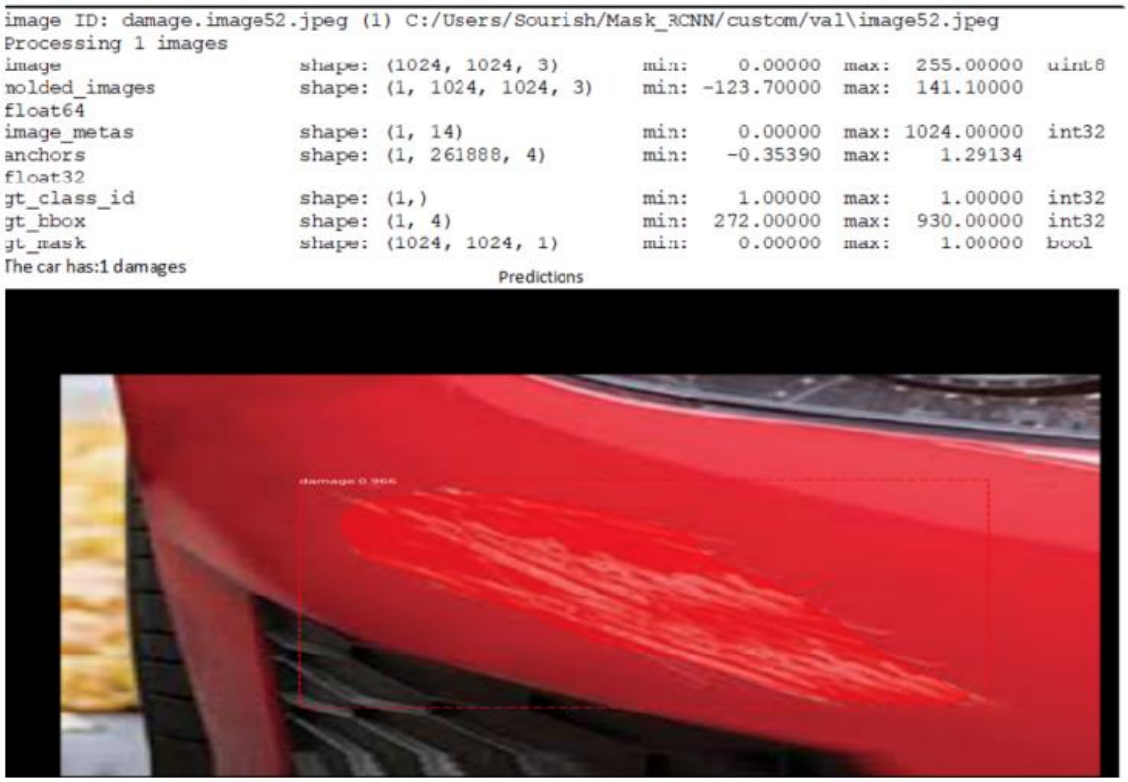


Figure 16 Prediction result of sample case (Towards data science 2019)

The final phase as demonstrated in figure 16 is testing model predictability before application. The result was acceptable as it could mark the area of scratches on the hood of the auto and also quantify the severity extent of scratches. (Towards data science 2019.)

Fraud detection in submitted images:

In the growth of technology, images are processed more by the computer vision in a large amount. Together with the trend, there is also the simultaneous rise of fraudulent claims in insurance as a deceptive act. The section would exemplify the solution by examining the case study “An Anti-fraud, System for Car Insurance Claim Based on Visual Evidence” from the University of 1otre Dame (Li et al. 2018). In brief, the solution is about capturing the images from the dataset, originating from annotations and augmentation provided.

Estimate the repair costs based on the type of damage recognised:

The chain of image recognition, estimation of repair costs predicted. Image recognition is able to collect the data, about whether it is a scratch or a dent, the size and shape of the damage, the car basic details, etc in order to determine whether it could be fixed or replaced. According to Data Reply (2020), a statistical evaluation is made based on the earlier mentioned factors, refers from the similar car model from the database. As the process of estimation involves various elements of the car images, from damages from particular angels, the vehicle basic details, combined with the insurance policy. Claim Genius (n.d) declares that its flagship estimation system could generate instant decisions on total loss and damage estimates after the customer upload the submitted images and videos

through its app called Genius APP. The figure 17 is a thorough example on how the estimates are extracted from the image.

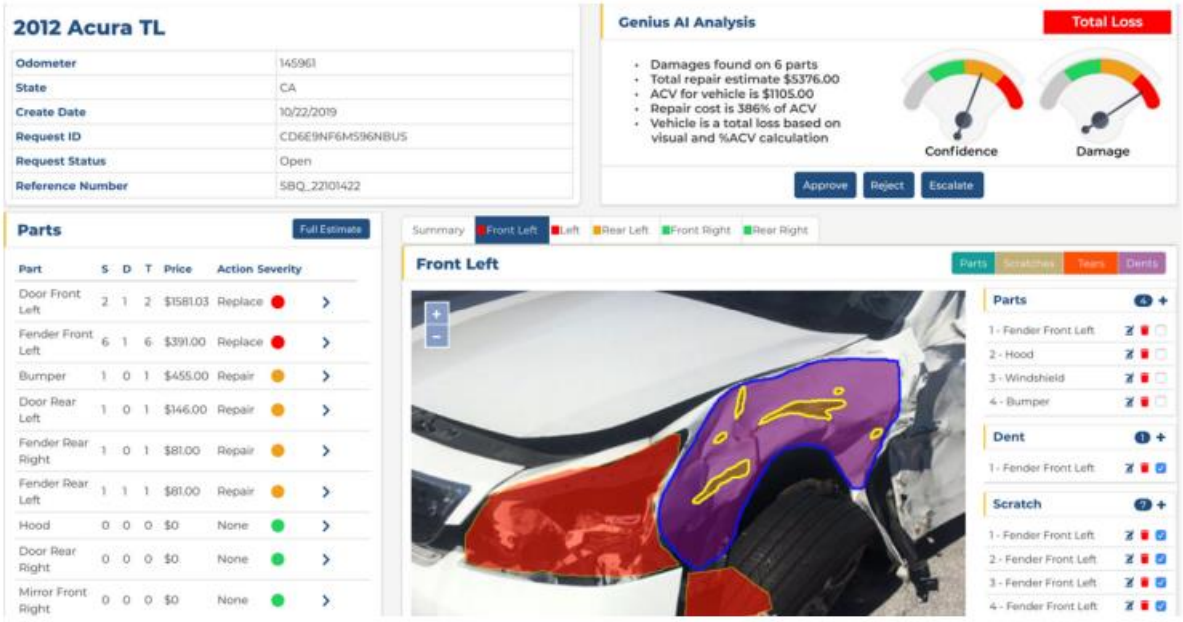


Figure 17 Thorough example of damage evaluation (Claim Genius n.d)

3.3 Business competence:

With the integration of image recognition, the automation and digitalization has created huge positive impacts on the industry and the groups of interest.

Complete and real-time information

The real time information would be highly consistent and followed during the claim flow, as well as recording purpose.

Tailored strategy repayment:

Image recognition would also serve the customers better with tailored plan for repayment. The personalization of repayment would satisfy the customers' current financial status at the moment of accident.

Shortened processing time:

The human force into the process would process the claim with longer time depending mostly on the personal experiences, amount of claims and complexity of cases. The reduction in dealing auto claim would be obvious thanks to the computer vision and its algorithms. Thus, the efficiency is highly enhanced as well as general operations of all parties involved. The absence of car brings customers a lot of inconvenience during their daily commute. With the added value, the car could be faster sent back to the customers. The insurer would expect less of processing time.

Accurate damage evaluation:

There are cases that the car is taken to the garage and decided as not repairable. With the assistance of image recognition, the computer owns a bigger data and careful inspection to define whether the car is repairable without the presence of the repair shop representative.

Code efficiency:

Image recognition could reduce the insurance carrier's cost by virtually assess the images instead of sending the adjusters to the scene. The automated process would save the company unnecessary expenditure and a decent amount of workforce.

4.Key to success of image recognition:

As image recognition is still a newly introduced term in the insurance industry, the success of this technology requires strategic and technical considerations. Most concerned issues could be listed as business added values and maintenance, data strategy, and customer privacy.

4.1 Resource allocation

From existing business and its current premises and infrastructures, there is a huge transformation into the image recognition-based solution. Specifically, the insurance company needs to identify which parts of the business that the computer vision can bring the value and also the return on investment.

4.2 data strategy:

The suppliers usually provide demos, proofs of concept, and pilot deployment. (Tech See 2019.) Employing the technology means also a data team to deliver, maintain and refine customer's experiences.

5. Research conduction:

This section will explain in details how the empirical data is conducted. Thus, it will firstly present the research methods and the reason for choosing. Further, the data collection methods together with questionnaire design are explained in details.

5.1 Research method:

There are knowingly qualitative and quantitative research methods. The quantitative involves a large of respond and procedures to gather specific forms of data in order to generate a numerical result. This method is best suitable for measuring, ranking and identifying patterns. Knowing the definition of two research methods and their features, the qualitative research method is chosen for this is as the purpose is to gain the in-depth insights about image recognition and in which aspects that it impacts the auto damage claim services.

5.2 Data collection: This section will specify the data types and how each type of data is collected for analysis by the thesis author.

Types of data:

There are generally known two types of data: primary and secondary. Secondary data is defined as the information collected by other researchers. Sources can be published books, magazines, scientific research, journals, academic journals, etc.

6. Research data analysis:

This section includes the analysis of researcher on the data collected from both primary and secondary source to ensure the validity and reliability. The structure of analysis follows the framework of research questions, sub questions, questionnaire. Specifically, the analysis revolves around the auto damage claim, first exploring the necessity of image recognition, its contribution to the current process, and lastly the outlook of image recognition.

6.1 Image recognition and its mechanism:

Image recognition is a totally different concept of automating the auto claim process, which emulates the human's triaging process. These are specified above chronologically. The flow is applied in the analysis of the first question. For the common agreement, thesis places its focus on the business view when addressing the issue. And the specific technology as the competitive edge is hardly disclosed only on the business practices instead of public understanding

Image acquisition:

Within the perspective of insurance, specifically auto damage claim, the image recognition starts with the series of photos the customer uploads regarding the car damaged following the accidents. Another source of images applicable for analysis is standard CCTV cameras, which is utilized by Ravin.AI beside the traditional smartphone's photos. The company utilized the frequency of cars passing through the camera-installed sites such as gas stations, parking lots, building entrances, garages, etc.

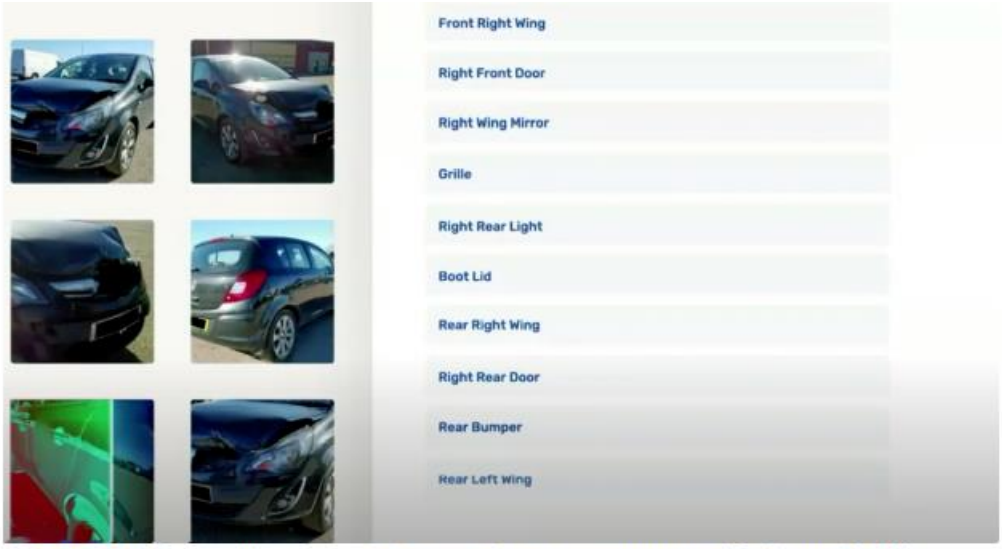


Figure 18 Scanning for car's exterior components (Cohen 2019)

For the concept of Ravin AI, the advantages of acquiring CCTV photos allow Ravin AI to deliver the on-going record of the cars' conditions. Extracting features from processed images, Ravin AI could inform the owners in advance and propose solutions that can help to save the cost for more severe damages in the future. Besides, the right of customers' privacy is highly respected within the company.

Key to success and future outlook:

The key to success is an abstract term regarding to several factors required from the clients. As image recognition is provided as services from the third party, this is a business model. Therefore, the technology is dealing directly with the carriers in term of building the system that works best for them.

7. Conclusion:

This chapter highlight the outcome of the research answering directly the research questions stated from the beginning with comments from the author. The findings are briefly presented following the flow of questions stated earlier. Research limitations and recommendations for further researches are mentioned at the last section.

7.1 Research findings

Through the research from several companies which provides the services, there are several findings that are valuable to encapsulate the whole research. Brief concept, image recognition is the technology that extracts the solid information from the photos. The system is created by two elements: models and data. The models are created by engineers as the frame to guide the system at the very initial stage. In car insurance, the data as photos of damaged cars are fed into the system for useful feature extraction and simultaneously, as training material. With the input of more photos, the more the machine learns and increase its accuracy.

Efficient resourcing and satisfied stakeholders

The image recognition deployment has revolutionized the resourcing plan with three main impacts: timely and accurate information, streamlined process, high processing capacity. With transition from human-based to technology-based business, the processing cycle is smoothened and shortened. There are less mistakes created and increasingly high processing capacity. The policy holders are in control of their vehicles' conditions, expected shorter processing time and procedures. All the values are integrated to improve customer experience, customer retention and eventually the carrier's competitiveness in the market.

Limitations and recommendation for further research:

The main motivation of this research is to examine the hype of image recognition with in a specific practice, auto claim process from the business perspective.

- Vehicle inspection requires transportation UVE

- A few advanced steps in model training to enhance the accuracy including:

Data normalization, regularization and calibration.

Within the related industry that the similar services can be utilized, there are several research that can be conducted: -

Image recognition in the transformation of the industries (such as car dealership, vehicle transportation, and others requiring car inspection services).

- Establishment of competitive analysis between image recognition and human force in auto insurance industry.

- Deployment plan of image recognition for insurance carriers.

- Extracting information from image recognition information for insurance