**Capstone Project**

**Automated Car Damage Detection for Insurance Claims**

**By**

**Yan Zhang**

**Date: 20/08/2025**

Table of Contents

[Problem statement 3](#_Toc205655872)

[Industry/ domain 4](#_Toc205655873)

[Stakeholders 5](#_Toc205655874)

[Business question 5](#_Toc205655875)

[Data question 6](#_Toc205655876)

[Data 6](#_Toc205655877)

[Data science process 7](#_Toc205655878)

[Data analysis 7](#_Toc205655879)

[Modelling 8](#_Toc205655880)

[Outcomes 17](#_Toc205655881)

[Implementation 18](#_Toc205655882)

[Data answer 18](#_Toc205655883)

[Business answer 18](#_Toc205655884)

[Response to stakeholders 18](#_Toc205655885)

[End-to-end solution 19](#_Toc205655886)

[References 19](#_Toc205655887)

# Problem statement

**Opportunity / Problem:**

Manual car damage assessment in insurance claims is often inefficient, inconsistent, and prone to human error. Currently, claim adjusters are required to physically inspect vehicles or review images manually, which can lead to significant delays in processing claims. This not only increases operational costs but also opens the door for fraudulent claims to slip through the cracks.

**Value of Addressing the Problem:**

Addressing this problem can have several key benefits. First, it can significantly reduce the time required to process claims, cutting it from days to just minutes. Additionally, by reducing the need for manual inspections, operational costs will decrease. The consistency and accuracy of assessments can also be greatly improved, ensuring that claims are handled more objectively. Most importantly, automation will help reduce fraud by classifying damage types with greater precision, limiting the chances of errors or misinterpretations.

**Current State:**

At present, claim settlements often take several days due to backlogs and the time-consuming nature of manual reviews. This leads to customer dissatisfaction, as they experience slow resolutions to their claims. On the insurer's side, there are higher labor costs associated with these lengthy processes, and the potential for fraudulent claims to go undetected remains a significant issue.

**Desired State:**

The goal is to implement an automated, image-based assessment system with over 75% accuracy. This system would classify damage types in real-time from customer-uploaded photos, streamlining the process. By doing so, settlements could be completed much faster, reducing the likelihood of disputes and enhancing customer satisfaction.

**Prior Research:**

Previous research has explored the use of convolutional neural networks (CNNs) to detect vehicle damage, with accuracies ranging between 70% and 85%. However, most of these solutions were proof-of-concept models and have yet to see large-scale deployment. Our project builds on this prior work by experimenting with multiple architectures—including MobileNetV2, VGG-16, and custom CNNs—and plans to deploy the best-performing model for real-world application.

# Industry/ domain

**Industry Overview:**

This project is situated within the insurance and automotive repair industries, particularly focusing on the auto claims segment. The goal is to enhance the efficiency and accuracy of damage assessments, a crucial process within these industries.

**Current State of the Industry:**

The auto claims sector is facing increasing competition as insurers strive to improve the customer experience and streamline their operations. A key trend is the rising adoption of artificial intelligence (AI) in areas like fraud detection and claims automation, which help to reduce processing time and enhance accuracy. However, the industry also faces several challenges, including large volumes of claims, escalating repair costs, and ongoing regulatory scrutiny, all of which make the claims process increasingly complex.

**Value Chain Position:**

Automated damage detection technology fits strategically within the value chain, positioned between claim initiation (where customers upload images of the vehicle) and claim assessment (where the final payout decision is made). This technology serves as an essential pre-screening tool for adjusters, improving the accuracy and speed of damage assessments before they reach the final decision stage.

**Key Concepts:**

The project utilizes several advanced technologies to optimize the damage assessment process. Convolutional Neural Networks (CNNs) are deep learning models commonly used for image classification, making them ideal for this task. Transfer learning is another key concept, where pretrained models—such as those trained on ImageNet—are used to reduce the time and computational resources required for training. Additionally, image augmentation techniques are employed to improve the model’s ability to generalize, by artificially modifying the training images to expose the model to a wider variety of possible inputs.

**Cross-Industry Relevance:**

The technology developed in this project is not limited to the insurance industry. It has broader applications across various sectors, including manufacturing quality control, used car evaluations, and vehicle rental inspections. In these fields, automated image classification could streamline processes, enhance accuracy, and reduce human error, offering valuable benefits beyond auto claims.

# Stakeholders

The primary stakeholders in this project include insurance companies, claim adjusters, policyholders, repair shops, and regulators. Insurance companies stand to benefit by reducing operational costs and improving the speed of claim processing. Claim adjusters will be able to focus on more complex cases, as routine inspections are automated, leading to greater efficiency. For policyholders, the key advantage is faster and fairer settlements, ensuring quicker resolution of their claims. Repair shops will also experience quicker approval for repair jobs, streamlining their workflow. Finally, regulators are important stakeholders, as they will need to ensure that the process remains fair and compliant with industry standards.

The motivations driving these stakeholders are diverse. Insurers are primarily motivated by the potential for cost savings and the reduction of fraud through more accurate, automated assessments. Customers, on the other hand, are looking for faster payouts and a more efficient claims process. Regulators expect the system to maintain transparency and fairness, ensuring that all claims are handled according to established guidelines and regulations.

In terms of expectations, there is a clear demand for an automated system that achieves at least 75% accuracy in damage classification. Additionally, the system must seamlessly integrate with existing claim management infrastructure and be easy to deploy for both customers uploading their images and adjusters processing the claims.

# Business question

**Main Business Question:**

The central business question revolves around whether an AI model can automatically and accurately classify different types of car damage from images, thereby speeding up and standardizing the insurance claim processing. By leveraging AI, the goal is to make the process more efficient and reliable, while reducing human error.

**Business Value:**

The potential business value of implementing this AI-driven solution is substantial. If manual assessments currently cost $20 per claim, and AI is able to handle 50% of the 1 million annual claims at a large insurer, the estimated cost savings could exceed $10 million per year. Additionally, by automating the process, claim processing time could be reduced dramatically—from an average of three days to less than one hour for simpler cases—leading to faster resolution and improved customer satisfaction.

**Required Accuracy & Risks:**

The AI model’s target accuracy is set at 75% or higher to ensure effective classification of damage. However, there are inherent risks involved. False positives—where the model detects damage that doesn't actually exist—could result in unnecessary payouts, impacting profitability. Conversely, false negatives—where the model misses actual damage—could lead to customer dissatisfaction, legal disputes, or increased fraud, undermining the trust and effectiveness of the system. Balancing these risks will be crucial for the successful deployment of the solution.

# Data question

The key data question is: Which damage category, from the eight predefined classes, is present in a given car image? To answer this, labelled car images are required that cover all the defined damage types, as well as examples of "unknown" cases where damage is not present or not identifiable.

# Data

The data was sourced from Kaggle, with a provided dataset containing approximately 1,500 RGB images (224×224 pixels) across eight classes:

* + - * bumper\_dent
      * bumper\_scratch
      * door\_dent
      * door\_scratch
      * glass\_shatter
      * head\_lamp
      * tail\_lamp
      * unknown

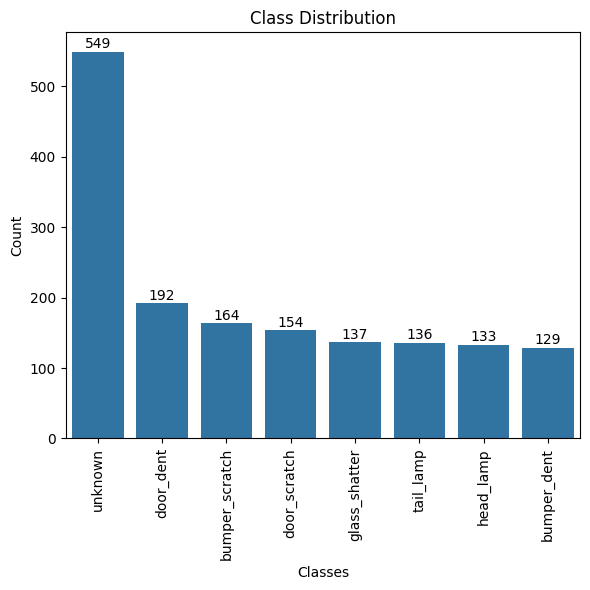
The dataset includes around 1,500 images, with attributes such as RGB pixel values and category labels. The images were collected under varying lighting, angles, and resolutions, simulating real-world conditions for more reliable model training. While there is an imbalance in some classes, this has been addressed through data augmentation techniques.

For ongoing availability, the dataset would need to be integrated with the insurer's photo submission systems, ensuring continuous data collection for further model refinement.

# Data science process

## Data analysis

The dataset includes a CSV file and a set of images. The CSV file contains the file paths for each image along with its corresponding class label. Below is the distribution of the image classes:



The processing pipeline involves several key steps: data loading and cleaning, followed by a train-test split to facilitate proper model evaluation. To improve generalization, address class imbalances, and prevent overfitting, data augmentation techniques such as rotation, zoom, and flipping were applied.

During exploratory data analysis (EDA), it was observed that some damage categories, such as *bumper\_dent* and *bumper\_scratch*, are visually similar, which makes classification more challenging. On the other hand, the "unknown" class is highly diverse, enhancing the model’s robustness for real-world scenarios.

The pipeline has been designed for reusability, enabling it to efficiently process future image data in the same format, ensuring scalability and smooth integration as the project evolves.

## Modelling

Several models were tested during the project:

* + MobileNetV2 (using transfer learning)
  + VGG-16 (using transfer learning)
  + Custom CNN (built from scratch)

For feature engineering, the images were normalized and augmented to improve model performance.

The models were evaluated based on key performance metrics including accuracy, precision, recall, and F1-score.

For each model, we experimented with different dropout rates—0.2, 0.3, and 0.5—to assess their impact on model performance.

MobileNetV2\_1 (dropout rate: 0.2)A screenshot of a computer

AI-generated content may be incorrect.

A graph of different colored lines

AI-generated content may be incorrect.

MobileNetV2\_2 (dropout rate: 0.3)

A screenshot of a computer

AI-generated content may be incorrect.

A graph of different colored lines

AI-generated content may be incorrect.

MobileNetV2\_3 (dropout rate: 0.5)A screenshot of a computer

AI-generated content may be incorrect.

A graph of different colored lines

AI-generated content may be incorrect.

VGG-16\_1 (dropout rate: 0.2)A screenshot of a computer

AI-generated content may be incorrect.

A graph of different colored lines

AI-generated content may be incorrect.

VGG-16\_2 (dropout rate: 0.3)A screenshot of a computer

AI-generated content may be incorrect.

A graph of a graph of a graph

AI-generated content may be incorrect.

VGG-16\_3 (dropout rate: 0.5)A screenshot of a computer

AI-generated content may be incorrect.

A graph of different colored lines

AI-generated content may be incorrect.

Custom\_CNN\_1 (conv\_dropout\_rate**:**0.2, dense\_dropout\_rate**:**0.2)A screenshot of a computer

AI-generated content may be incorrect.

A graph of a variety of lines

AI-generated content may be incorrect.

Custom\_CNN (conv\_dropout\_rate**:**0.5, dense\_dropout\_rate**:**0.5)A screenshot of a computer

AI-generated content may be incorrect.

A graph of a line and a line graph

AI-generated content may be incorrect.

Custom\_CNN\_3 (conv\_dropout\_rate**:**0.25, dense\_dropout\_rate**:**0.5)A screenshot of a computer

AI-generated content may be incorrect.

A graph of different colored lines

AI-generated content may be incorrect.

## Outcomes

**Summary of the results:**A screenshot of a computer

AI-generated content may be incorrect. A graph of blue bars

AI-generated content may be incorrect.

The main findings and conclusions from the test-set metrics are as follows:

* **MobileNetV2 Performance**:

MobileNetV2 models consistently performed well across different configurations, with accuracy ranging from approximately 78% to 79%. The precision, recall, and F1-score values were all relatively balanced, with the best-performing configuration being *MobileNetV2\_3*, which achieved an accuracy of 78.99%, precision of 79.05%, recall of 78.99%, and F1-score of 78.47%. This indicates that MobileNetV2 provides a good balance between accuracy and generalization.

* **VGG-16 Performance**:

VGG-16 models showed a more varied performance. The accuracy ranged from 64.58% to 73.67%, with *VGG-16\_2* being the highest performer at 73.67%. However, precision, recall, and F1-scores for VGG-16 were generally lower than those of MobileNetV2. The model *VGG-16\_3* had significantly poorer performance with an accuracy of only 64.58%, highlighting its limitations in this task.

* **Custom CNN Performance**:

The Custom CNN models performed poorly compared to the other architectures, with accuracy values ranging from 38.56% to 42.01%. Precision, recall, and F1-scores were all substantially lower, with the highest F1-score reaching just 29.88%. This suggests that the custom model struggled with both classification and generalization, likely due to its architecture or insufficient training data.

**Overall Conclusion**:

* + **MobileNetV2** clearly outperforms the other models in both accuracy and consistency, making it the most promising candidate for this task.
  + **VGG-16** has decent performance, but it lags behind MobileNetV2 and shows significant variability across different configurations.
  + **Custom CNN** underperformed, indicating that more sophisticated or better-tuned architectures like MobileNetV2 and VGG-16 are more suitable for the given task.

Given these findings, the best model to proceed with would be **MobileNetV2**, especially the *MobileNetV2\_3* configuration– accuracy 79%, which offers the most balanced and reliable performance across all metrics.

## Implementation

The best-performing model has been deployed through a Streamlit web application. This allows customers or adjusters to upload an image and receive an instant prediction of the damage category. There is potential for seamless integration of this system into the insurer’s backend claims process for further automation.

# Data answer

The data question was answered satisfactorily, as the model is capable of classifying 8 damage types with an accuracy of approximately 79%. In terms of confidence, it is medium-high, as the accuracy is promising but could be further improved with additional training data and fine-tuning of the model.

# Business answer

The business question was answered satisfactorily, as the solution has the potential to significantly reduce both claim processing time and operational costs. In terms of confidence, it is high for using the system as a decision-support tool, but moderate for fully automating the process without any human oversight, as certain complexities may still require manual review..

# Response to stakeholders

To effectively implement the solution, it is recommended to start with a pilot program to validate its performance on live claims. This will provide insights into its real-world effectiveness and allow for any necessary adjustments. Additionally, ongoing collection of labeled data will be crucial to further improve the model’s accuracy over time. For high-value claims, a hybrid approach that combines AI with human verification should be adopted to ensure accuracy and reduce risk. Lastly, periodic retraining of the model will be important to keep it up-to-date with new damage patterns and evolving trends in the insurance industry.

# End-to-end solution

The end-to-end solution follows a streamlined workflow. First, the customer uploads car images through the claims portal. These images are then preprocessed and passed through the trained MobileNetV2\_3 model, which predicts the damage category. The prediction is sent to the claim adjuster for verification, ensuring accuracy before final approval. As a result, the claim is processed much faster, with significantly reduced reliance on manual inspections, improving both efficiency and turnaround time.

# References

<https://www.kaggle.com/code/pavansanagapati/a-simple-cnn-model-beginner-guide>

<https://medium.com/analytics-vidhya/types-of-convolutional-neural-networks-lenet-alexnet-vgg-16-net-resnet-and-inception-net-759e5f197580>

<https://www.geeksforgeeks.org/computer-vision/what-is-mobilenet-v2/>

<https://machinelearningmastery.com/how-to-configure-the-number-of-layers-and-nodes-in-a-neural-network>

<https://www.kaggle.com/datasets/hamzamanssor/car-damage-assessment/data>

<https://docs.streamlit.io/deploy>