

# **DSE-I2100 Applied Machine Learning Project**

Car Damage Severity Prediction

Data Augmentation, Deep Learning, Classification Machine Learning

Lizzette Salmeron, Nikhita Kannam, Rohit Ravindra Gulve, Jose Salas

## **I. Introduction**

The applications of Machine Learning (ML) techniques have integrated in everyday tasks and interactions with technology. With access to large amounts of data and images, Deep Learning (DL) algorithms can utilize the limitation of needing massive amounts of data as an advantage to common real-world problems. ML is an important tool that enables businesses to make informed decisions and allows us to look at problems in innovative ways. A crucial aspect of ML is image recognition, the process of using different algorithms and technologies to identify an object in an image or video to achieve different tasks like classifying images into certain categories and understanding what an object represents. Various computer vision methods as well as Deep Learning algorithms are used to achieve the required results for solving a certain problem. One application this process can be used for is for car damage detection.

## **II. Motivation**

Car accidents are very common and can cause various degrees of damage to a car, ranging from small crashes to totaled cars. Insurance companies and car repair businesses can predict the price of replacement and make better choices by accurately forecasting the degree of the damage. However, since visual inspections and validations for the most part are done manually, assessing the damage severity can be time-consuming and subjective to each inspector. Most insurance companies waste a lot of time on claim leakages, which is the difference between what an insurer spent to settle the claim vs what they should spend. Claim leakages determine the amount of money an insurance company would have to pay. An algorithm that automatically and fairly inspects the damage to the car can greatly speed up the process and reduce costs for inspections.

With this project, we are looking to explore the use of Deep Learning techniques by building a Damage Classifier capable of detecting three different types of damage (Minor, Moderate, and Severe) and correctly predicting the severity of car damage. This is crucial for insurance firms

and car repair shops. Accurately determining the extent of the harm can assist a vehicle's owner's claims with insurance.

### **III. Data**

The Car Damage Severity Dataset [1] from Kaggle contains over 1631 images of cars with various degrees of damage: minor, moderate, and severe. The data is also split into a training set and a validation set.

### **IV. Method**

This project is an application method that will apply Deep Learning techniques for image classification. The goal is to develop a DL algorithm that can accurately predict the severity of car damage based on images of damaged cars. The results from this project will have practical applications in the insurance and auto repair industries.

The experiment will be conducted as follows:

1. Using a dataset of images of damaged vehicles with varying degrees of severity, one can fine-tune CNN models that have already been trained.
2. Comparing the performance of different pre-trained models such as ResNet, VGG16, YOLOv3, and detectron2.
3. Explore how to expand the dataset and enhance the model's resilience by using data augmentation methods like random cropping, flipping, and rotation.
4. Examine how various hyperparameters, such as learning rate, group size, and optimizer, affect model success.

To further elaborate on the aforementioned steps, our group will begin by expanding our sample and strengthen the model's resilience. This will be done by employing data augmentation methods, such as Keras, to perform the image augmentation to the existing dataset, as the amount of images for each classification label is small. An image annotation tool, Labelmg, can be utilized to label the images based on the severity of the damage: minor, moderate, and severe. Here, convolutional neural networks (CNNs) can be utilized to classify the severity of the damage for a baseline ML model that can be trained using Imagenet, a massive database of images. Specifically, we will use transfer learning methods to enhance pre-trained CNN models like ResNet, VGG16, YOLOv3, and detectron2. Transfer learning enables us to apply the

features found by these models to other image classification tasks, enhancing the accuracy of our job to classify the severity of vehicle damage. By observing the effects of various hyperparameters on the model's success, including learning rate, group size, and optimizer, we can develop a well-trained model that can support small datasets and present exceptional results.

## **V. Evaluation**

We plan on evaluating our Deep Learning model using a combination of metrics such as accuracy, recall, precision, and F1-score. These measures will give us an idea of how well algorithms perform on image classification.

## **VI. References**

- [1] Bhamere, P. (2022, December 31). *Car damage severity dataset*. Kaggle. Retrieved from <https://www.kaggle.com/datasets/prajwalbhamere/car-damage-severity-dataset>
- [2] Parra, D., & Gonzalez, A. E. (n.d.). *Car damage assessment - UTRGV*. Retrieved April 6, 2023, from <https://faculty.utrgv.edu/dongchul.kim/csci4352/spring2019/report/R11.pdf>.
- [3] *A Complete Guide to Image Recognition*. Nanonets. (n.d.). Retrieved April 5, 2023, from <https://nanonets.com/image-recognition>