IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005030|

Auto Insurance Claim Prediction Using CNN and XGBOOST Algorithm

C.Vijayalakshmi[1], R.Rishi[2] , G.Raghul[3], N.Nivash[4]

Assistant Professor, Dept. of CSE, Panimalar Engineering College, Tamil Nadu, India[1]

UG Student, Dept. of CSE, Panimalar Engineering College, Tamil Nadu, India[2,3,4]

ABSTRACT: One of the most important tasks in the automobile insurance and rental businesses is vehicle damage detection. These systems are used by the driver to identify the damage to a car after an accident, as well as by the insurance company to detect and decide a reasonable amount based on the damage, and by vehicle rental companies to alert customers about the damage to a vehicle. Object recognition is the main approach used here. So,after the accident owner of the vehicle just have to scan the image of the car and machine learning model will you the rough claimable amount ; The system will recognise the image once it has been uploaded. I examine the image for dent marks, damaged glasses, and other flaws. The car is then categorised into various severity classifications based on criteria such as the make, model, and year of manufacturing. The severity created by the damage image is then mapped to the cost rules, which are built using numerous vehicle parameters such as the make, model, and year of manufacture. Finally, the customer is informed of the degree of the damage and the usual cost at which the harm can be repaired.To look into this problem, we are utilising the notion of image analysis, which is utilised to obtain a more precise damage result for any outside portion of the vehicle.

I. INTRODUCTION

Claims leakage is wasting a lot of money in the vehicle insurance market right now. The gap between the actual claim payment made and the amount that should have been paid is known as claims leakage. To lessen these impacts, validation and visual examination have been applied. There have been too few attempts to minimize claim processing time by too few start-ups. Convolutional Neural Network (CNN)-based algorithms were used to classify different types of automotive damage. Glass shatter, door dent, bumper dent, tail light broken, head lamp broken, smash, and scrape are just a few examples of typical damage kinds. We generated our own dataset because there isn't one for car damage classification that is publically available. create your own dataset by manually annotating several photographs of various types found on the internet. Due to characteristics such as barely evident damages and high inter-class similarities, the classification task is difficult. The effectiveness of our future insurance solution architecture has been validated by experimental findings.

II. RELATED WORK

We went out and reviewed a lot of articles on car damage reduction that had already been published. The advantages and demerits of some of the papers are listed below."Convolutional Neural Networks for Automatic Car Damage Recognition." [1]."Image based automatic vehicle damage detection." is a concept that uses a CNN[2].it failed to predict the data from given data set This method necessitates the use of 3D computer-aided design (CAD) modes of the car in question in order to determine how it would appear if it were not damaged. However, because they were not able to get such 3D models, they turned to advanced computer vision applications such as convolutional neural networks.. Automotive damage classification based on deep learning- We used CNN-based algorithms to classify car damage. types. They looked at common damage types like glass shattering, bumper and door dents, broken headlights and taillights, smashes, and other scratches. When it comes to auto insurance, image analysis is being used. A unique application for triage [4].have employed a basic strategy for analyzing minor vehicle damage and claiming the procedure quickly because the model does not need to be trained.[5]

Damage classification in this paper we learned to classify where the damage has been done from this paper we can scan the damage from any part of the car [6]from reference 6 we took car monitoring in which it can scan the image of the car successfully when compared to image classification algorithm[7] From this we learned how to use convolutional neural network efficiently on image so that we got 80% accuracy on convolutional neural network[8] first we used residual neural network form image classification using this idea we used our image classification by

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005030|

convolutional neural network[9]

In this paper they used they detected percentage of damage in vehicle from this we used to detect whether the image is damaged or not[10]from this paper it will detect the damaged car from video it can used for our future development of the project[11]from this paper they used random forest which is less accurate we used XGBOOST to calculate the insurance of the company which is more accurate when compared to random forest[12]On road vehicle reduction gives us the insurance without the numerical data which is purely based on image classification predicting amount We have done this to verify how the data correlates with one another.

III. DATASET PROCESSING

Preprocessing image : In this module the collected images are resized and converted from RGB to BGA and stored to data frame

Preprocessing data : in this module the numerical data are scaled and non numerical data converted to numerical data and filling the missing data.

Split data : splitting 80% of dataset as training data and 20 % as testing data for both image and numerical dataset EDA(Exploratory Data Analysis) : We have plotted graphs to analyze features related to predicting amount We have done this to verify how the data correlates with one another.

Feature Selection : We have plotted a heatmap to check the correlation of features. To check which features hold more ground and which are not much useful and can be eliminated without severe change.

V. METHOD

Convolutional Layer:

CNN has three types layer in that convolutional layer is more important it is where more number of computation will takes place.it requires input data which is image ,a filter and feature map among other thing. Lets pretend our image is colour image, which is 3D matrix of pixels which means It has three dimensions height, width and depth which match to the RGB. Kernel will move across the image's receptive fields, checking the features Filter is generally 3x3 matrix which can vary in size which change the size of the fields after that filter is applied to a fields of the image and in between input pixels there will be dot product and filter is calculated repeating the process until kernel going across the image

XG Boost:

XG Boost is tree based machine learning algorithm which is scalable machine learning system for tree boosting.

XG boost (extreme gradient boosting).We used XG boost because it is better than random forest algorithm

BOOSTING : N new data sets are formed by random sampling instead of original dataset

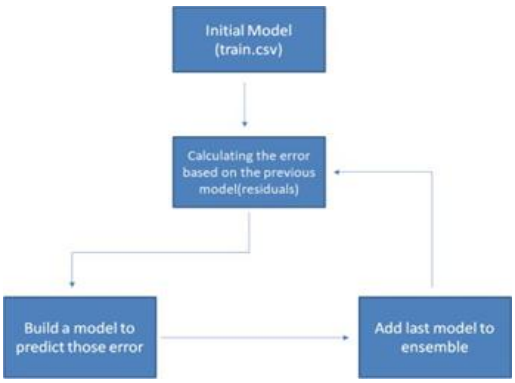

Figure1:Block diagram represents how XGboost algorithm works using the numerical dataset provided

In Figure 1:

i) It is used to calculate the amount using given numerical data.

ii) Initially build the first model and calculate the error for each observation in the dataset

iii) Then we built a new model to predict those residuals(errors)

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005030|

iv) Then you add prediction from this model to the ensemble of models

VI. IMPLEMENTATION

In training model we have to preprocess the image before training it. The preprocessed image size 150,150,3 i.e.(width height and channel) respectively. We have used (RGB)scale images to train the models then this image is turned into CNN mode

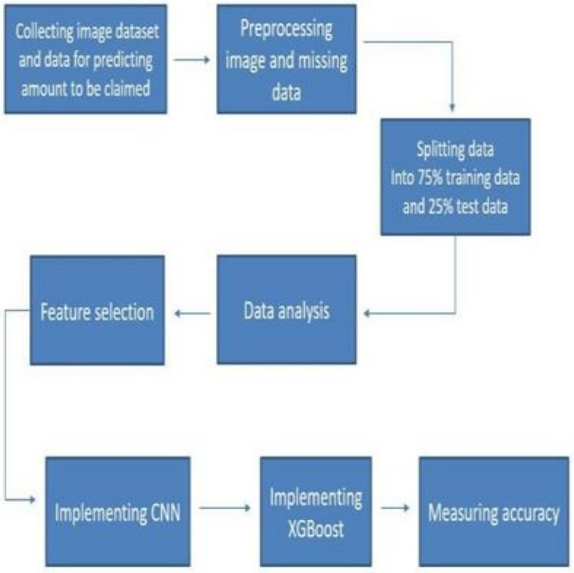

Figure 2:represents how this machine learning works

In figure 2 is how our machine learning model will work .we have two types of dataset which is image dataset and numerical dataset we have to preprocess the image as well as numerical data and we have to split the data into training and testing in which we are using 75% data as testing data and 25% as training data first applying convolutional neural network we have to use image of the car it will detect whether the vehicle is damaged or not after this algorithm We will get 2 outcomes 1 or 0 if its 1 using our numerical data we can find the insurance of the company if it is 0 the insurance claimable will be null

VII. RESULTS

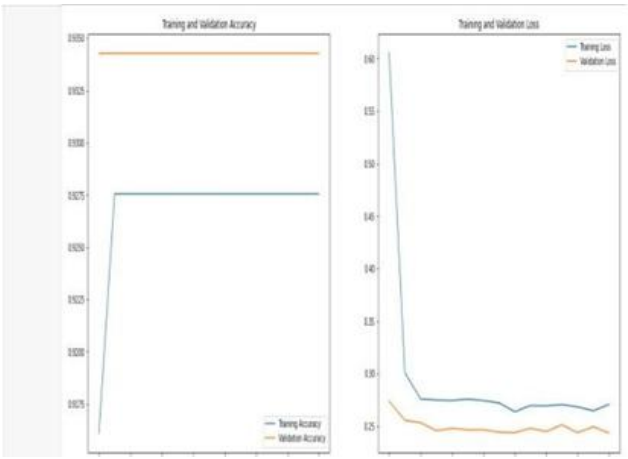

Figure 3: (i)graph 1 relation between training and validation accuracy

(ii) Graph 2 relation between training and validation loss

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005030|

In figure 3

i. graph 1 after some training from epoch 3 we got stable accuracy which is 93.8%

ii. graph 2 after some training we got stable validation loss less than 0.5%

TABLE I

MODEL	ACCURACY
Convolutional neural network	80%
XGboost	93.8%

In Table 1:We are interested in calculating the accuracy of the model so we ran 15 epoch for both convolutional neural network as well as XGBoost from this we calculated the accuracy of the model

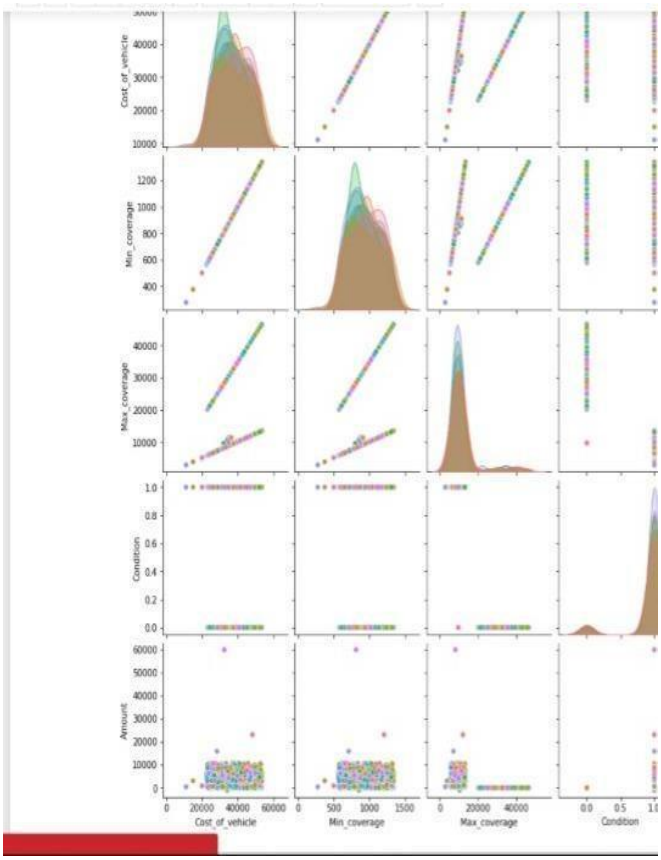

Figure 4 : this graphs represents the relation between all objects used in XGboost algorithm

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165 |

|| Volume 10, Issue 5, May 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1005030|

In figure 4: we can represent a relation between cost of the vehicle, amount , min coverage , max coverage and condition for 9 different types of the insurance company

VII. CONCLUSION

The paper outlines the amount and damage prediction of the car which are uploaded in our machine learning model.to verify this we used two types of algorithm convolutional neural network and xg boost
. we are using CNN to detect the image and to show output whether the car is damaged or not and we are using xg boost to predict the insurance cost from the given numerical data set

REFERENCES

[1] Sanket Doshi. “Traffic Sign Detection using Convolutional Neural Network.” 2019.
[2] de Deijn, Jeffrey. "Automatic Car Damage Recognition using Convolutional Neural Networks." (2018).
[3] Manju, More E. "RESEARCH PAPER ON INTELLIGENT FARMING FOR FARMERS USING CONTROL SYSTEM IN IOT." International Journal of Advanced Research in Computer Science
[4] Patil, Kalpesh, et al. "Deep learning based car damage classification." 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE, 2017.
[5] Harshani, WA Rukshala and Kaneeka Vidanage. "Image processing based severity and cost prediction of damages in the vehicle body: A computational intelligence approach." 2017 National Information Technology Conference (NITC). IEEE, 2017.
[6] Kinsella, Graham. Car damage monitoring system: Final project report analysis and design.Diss.Dublin,National College of Ireland, 2017.
[7] Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[8] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[9] Jayawardena, Srimal. "Image based automatic vehicle damage detection." (2013).
[10] Image Based Automatic Vehicle Damage Detection, Marcus Hutter, Stephen Gould, Richard Hartley, Hongdong Li, 2013.
[11] Li, Ying, and Chitra Dorai. "Applying Image Analysis to Auto Insurance Triage: A Novel Application." 2007 IEEE 9th Workshop on Multimedia Signal Processing. IEEE, 2007.
[12] Sun, Zehang, George Bebis, and Ronald Miller."On-road vehicle detection: A review." IEEE transactions on pattern analysis and machine intelligence 28.5 (2006): 694-711.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details